Resolution Enhancement in Magnetic Resonance Imaging by Frequency Extrapolation

by

Gregory Mayer

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 2008

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis focuses on spatial resolution enhancement of magnetic resonance imaging (MRI). In particular, it addresses methods of performing such enhancement in the Fourier domain.

After a brief review of Fourier theory, the thesis reviews the physics of the MRI acquisition process in order to introduce a mathematical model of the measured data. This model is later used to develop and analyze methods for resolution enhancement, or "super-resolution", in MRI.

We then examine strategies of performing super-resolution MRI (SRMRI). We begin by exploring strategies that use multiple data sets produced by spatial translations of the object being imaged, to add new information to the reconstruction process. This represents a more detailed mathematical examination of the author's Master's work at the University of Calgary. Using our model of the measured data developed earlier in the thesis, we describe how the acquisition strategy determines the efficacy of the SRMRI process that employs multiple data sets.

The author then explores the self-similarity properties of MRI data in the Fourier domain as a means of performing spatial resolution enhancement. To this end, a fractal-based method over (complex-valued) Fourier Transforms of functions with compact spatial support, derived from a fractal transform in the spatial domain, is explored. It is shown that this method of "Iterated Fourier Transform Systems" (IFTS) can be tailored to perform frequency extrapolation, hence spatial resolution enhancement.

The IFTS method, however, is limited in scope, as it assumes that a spatial function $f(x)$ may be approximated by linear combinations of spatially-contracted and range-modified copies of the entire function. In order to improve the approximation, we borrow from traditional fractal image coding in the spatial domain, where *subblocks* of an image are approximated by other *subblocks*, and employ such a block-based strategy *in the Fourier domain*. An examination of the statistical properties of subblock approximation errors shows that, in general, Fourier data can be *locally self-similar*. Furthermore, we show that such a block-based self-similarity method is actually equivalent to a special case of the auto-regressive moving average (ARMA) modeling method.

The thesis concludes with a chapter on possible future research directions in SRMRI.

Acknowledgements

I would first like to thank my supervisor, Edward R. Vrscay for his consistent support, inspiration, and encouragement to explore new ideas throughout the duration of my studies at the University of Waterloo. I would like to thank Dr. M. Ebrahimi, Dr. H. Peemoeller, and Dr. C. Lemaire, for many helpful conversations. Many other colleagues, members of the university community and friends, too many of which there are to name, are gratefully acknowledged for lending their encouragement and perspective throughout the course of my studies at UW.

I would also like to thank Dr. J. Ross Mitchell, Dr. Bradley Goodyear and Dr. Louis Lauzon of the Seaman Family MR Research Center, Foothills Hospital, University of Calgary, Alberta, Canada, for many helpful conversations during my Masters program, and for providing MR data that was used in this thesis. I also thank Dr. Hongmei Zhu, formerly at the Seaman Center and now at York University, Toronto, Ontario, Canada, for many helpful conversations during my Master's program. The ideas that these individuals contributed during my Master's program helped formulate the subject of this thesis.

I would like to thank my external examiner, Dr. Terry Peters (Robarts Research Institute, London, Ontario, Canada), whose invaluable suggestions and comments greatly improved this thesis. I also want to thank my other readers, Dr. S. Lipshitz, Dr. H. Peemoeller, Dr. J. Orchard, and Dr. A. Hajian, for their insight and feedback that also contributed to this thesis.

I also want to thank Dr. C. Lemaire and Dr. H. Peemoeller for providing MR data used in this thesis. Data from the University of Waterloo was obtained from a Bruker micro-magnetic resonance imaging spectrometer at UW which was purchased with funds provided by the Canada Foundation for Innovation (CFI) and the Ontario Innovation Trust. This spectrometer provides the focal point for an interdisciplinary project involving the collaboration of UW researchers from physics, engineering, and mathematics.

This project has also been funded in part by the NSERC Strategic Grant, "Fusion and Analysis of Multiscale MRI and SEM Imaging Data of Porous Media".

This research was also supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) in the form of a Postgraduate Scholarship, and in part by the Multiple Sclerosis Society of Canada in the form of a Studentship, which are gratefully acknowledged.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Magnetic resonance imaging[1] (MRI) is a powerful, noninvasive imaging modality capable of extracting a wide variety of information for clinical studies and scientific research. However, the efficacy of MRI in any application is limited by the image quality the machine hardware can produce. One measure of image quality is the spatial resolution[2], Δx, which quantifies the extent to which different features of an imaged object can be distinguished. The focus of the research presented throughout this thesis is to investigate new ways of improving the spatial resolution in MRI, thereby enhancing the versatility of this prevalent imaging modality.

Basic MRI Theory

To facilitate a discussion on MRI and its limitations, it will be helpful to first provide a brief description of how a typical MR image would be obtained. Chapter 3 will provide a more thorough description of the image acquisition process.

Typically in MRI, a set of magnetic fields is applied to a given object and the object's response to these fields is measured to generate data. Three different sets of fields are applied. A *main magnetic field* is used to produce a net magnetization that may be manipulated to produce the MR signal. A set of *linear magnetic field gradients* are also applied to the object. These gradients can provide a mechanism for encoding spatial information of proton concentrations in the sample. The gradients do this by creating a linear relationship between magnetic field strength and spatial position. Finally, another applied field, the *radiofrequency (RF) field*, applies RF energy that the object absorbs,

[1] A partial list of abbreviations used in this thesis is provided for the reader in Appendix C.

[2] A partial list of mathematical symbols used in this thesis is provided for the reader in Appendix D.

tipping the net magnetization away from its equilibrium position. The RF field is applied for a brief period of time, and the magnetization relaxes back to its equilibrium position after it is turned off. The hardware that applies the RF pulse is also responsible for measuring the MR signal. The RF coil then measures the signal as a continuous entity, which is subsequently digitized and stored on a computer.

Typically, the RF filed is designed to excite and be sensitive to a thin slice (or volume) of the object being imaged (see Fig. 1.1a). This allows the MR scanner to measure information from this thin volume, reducing the data acquisition and reconstruction process to a two-dimensional problem. Data in the remaining two coordinate directions are acquired in the *Fourier domain*[3], or k-space. Samples are (often) acquired along straight lines, parallel to one of the coordinate axes referred to as the *frequency encoding* (FE) axis, and subsequently sampled. In Fig. 1.1b, the FE direction coincides with the k_x-axis. The k_y direction is referred to as the *phase encoding* (PE) direction. After a sufficient number of data samples have been recorded, the image is obtained using an *inverse discrete Fourier Transform*[4] (see Fig. 1.1c).

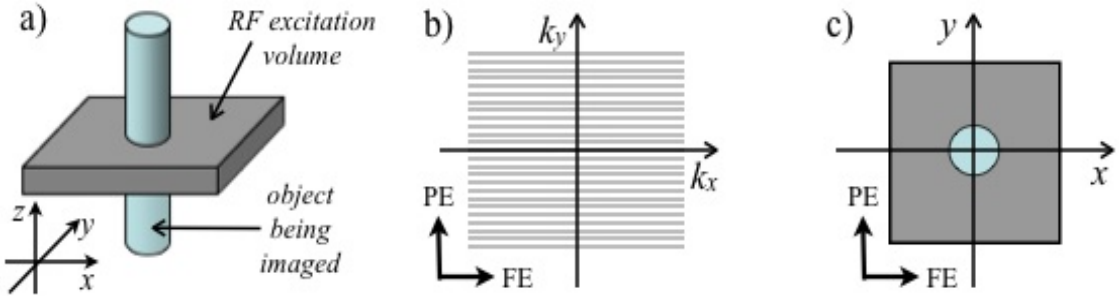


Figure 1.1: An object (the blue cylinder) being imaged and the region that the RF coil excites and is sensitive to (the grey volume) is shown in (a). Fig. (b) displays a schematic of how raw MR data is often acquired. Fig. (c) shows a schematic of the image obtained by taking a 2D inverse Fourier Transform of the raw MR k-space data.

Scan Time and SNR in MRI

Physical processes such as noise and T_2 decay as well as the scan time limit the spatial resolution in MRI. Each of these factors determines the particular acquisition strategy employed by the scanner operator.

Noise in MRI arises from thermal fluctuations in the sample, and electronic noise

[3]Fourier theory is *briefly* reviewed in Chapter 2. The interested reader may wish to refer to [8, 9], or [76] for a more thorough review of Fourier theory.

[4]The inverse discrete Fourier Transform is introduced in Section 2.2.

in the radiofrequency coil[5]. The *amount* of noise present in a given data set is often characterized via the signal-to-noise-ratio (SNR). In MRI, we often consider how much noise is present per voxel, which may be defined as the signal amplitude of a particular voxel divided by the noise standard deviation ([74] page 163):

$$\text{SNR per voxel} = \frac{\text{signal amplitude of a given voxel}}{\text{noise standard deviation}}. \tag{1.1}$$

In MRI, SNR and the spatial resolution depend upon several acquisition parameters. For example, the spatial resolution is connected to SNR in such a way that increasing the k-space extent (thereby increasing the spatial resolution) can decrease the SNR per voxel. Super-resolution MRI (SRMRI), which is the focus of this thesis, approaches attempt to circumvent this difficulty by improving resolution without increasing the SNR.

The SNR per voxel depends upon several acquisition parameters. To highlight the dependencies upon these parameters in conventional imaging, we use the following expression for the SNR (see [34], page 340)[6]

$$\text{SNR per voxel} \propto \frac{\Delta x \Delta y \Delta z \sqrt{N_x N_y N_z}}{\sqrt{BW_{read}}} \sqrt{N_{acq}}. \tag{1.2}$$

N_{acq} is the number of averaged acquisitions. For example, imaging the same object twice and averaging the result would yield a $\sqrt{2}$ improvement in SNR. N_x, N_y, and N_z are the number of samples in the respective x, y and z directions. BW_{read} is the readout bandwidth,

$$BW_{read} = 1/\Delta t, \tag{1.3}$$

and Δt is the sample spacing in the k_x direction. BW_{read} is ultimately bounded below by the spatial extent of the object being imaged. This parameter is introduced here, as it will play an important role in our discussion of multiple-acquisition SRMRI in the frequency encoding direction in Chapter 5.

The total scan time, T_{scan}, is given by

$$T_{scan} = N_y N_{acq} TR. \tag{1.4}$$

The parameter TE is the *echo time*, and is the amount of time required to measure one line in k-space.

There are several ways of increasing the spatial resolution in MRI through adjusting the acquisition parameters. For example, from Eq. 1.2, decreasing the voxel volume ($\Delta x \Delta y \Delta z$) by a factor of two would decrease the SNR by a factor of two. To compensate,

[5]Noise in the raw k-space MR data comes from a standard normal distribution. Noise in the spatial domain magnitude image is *Rayleigh distributed* (see [34], page 876).

[6]This equation assumes that frequency, phase, and slice encoding schemes are used in the x, y, and z directions, respectively.

one could increase the SNR by increasing N_{acq} by a factor of 4, but this would also increase the total scan time by a factor of four!

Indeed, the spatial resolution, T_{scan}, and the SNR are interconnected in MRI in such a way that increasing one implies a decrease in one or both of the others. This relationship is demonstrated by Eqs. 1.2 and 1.4, which play an important role in defining parameters that determine the obtainable resolution for a given amount of time and SNR. The motivation behind the research presented in this thesis is to explore multiple acquisitions and post processing strategies as an alternate means of enhancing the spatial resolution.

1.2 Problem Statement

Data in MRI is typically acquired in the Fourier domain over a finite region centered about the k-space origin. SRMRI techniques attempt to decrease Δx by increasing the frequency extent, $2k_{max}$. Specifically, the spatial resolution in a one-dimensional experiment is given by

$$\Delta x = \frac{1}{2k_{max}}. \tag{1.5}$$

This formula demonstrates the fundamental relationship between the spectral extent and the capacity of the image to represent features of the object over different spatial scales. This thesis explores methods of improving the spatial resolution by estimating frequencies beyond $|k_{max}|$.

The extrapolation problem is fundamentally tied to the constrained reconstruction problem through Eq. 1.5. Given that the achievable spatial resolution for a given SNR and scan time is limited, accurate estimation of higher frequencies can yield an improved resolution without an increase in scan time. Indeed, extrapolation techniques have been successfully used to provide a more effective means of producing an image in MRI [53, 55].

In the signal processing literature, extrapolation has normally been studied with signals that are functions of time. Although in this thesis we are interested in functions of frequency, the ideas that were developed for time dependent functions can easily be extended to k-space functions[7].

Signal extrapolation has been explored in the literature using continuous or discrete signals which may or may not be spatially limited. To distinguish between these different situations, we consider three extrapolation problems which are relevant to MRI.

Problem 1 - Continuous Signal Extrapolation: Consider a function $s_{true}(k) \in L^2(\mathbb{R})$ measured over a single, finite interval of length $2k_{max}$, centered at the origin, to

[7]The MR signal can actually be viewed as a function of time with the appropriate change of variables (see Eq. 3.51).

produce $s_{k_{max}}(k), k \in [-k_{max}, +k_{max}]$:

$$s_{k_{max}}(k) = \Pi(\frac{k}{2k_{max}}) s_{true}(k),\ \Pi(u) = \begin{cases} 0 & |u| > \frac{1}{2} \\ 1 & |u| \leq \frac{1}{2} \end{cases} . \tag{1.6}$$

Given $s_{k_{max}}(k)$, the general continuous extrapolation problem is to estimate $s_{true}(k)$ for $|k| > k_{max}$.

Problem 2 - Continuous Spatially Limited Signal Extrapolation: Now consider Problem 1 with the additional constraint that $s_{true}(k)$ is spatially limited with spatial extent R. Thus, given $s_{k_{max}}(k)$, the spatially limited continuous extrapolation problem is to estimate $s_{true}(k)$ for $|k| > k_{max}$. In MRI, R is the spatial extent of the object being imaged, and therefore may always be estimated.

Problem 3 - Discrete Spatially Limited Extrapolation: Consider Problem 2, with the added constraint that $s_{k_{max}}$ is sampled N_x times at regularly spaced intervals of length Δk, to produce $s_{d,N_x}(n)$. Given $s_{d,N_x}(n)$, we wish to estimate values of $s_{true}(k)$ at a set of points outside of the measured interval.

The solution to Problem 1 is non-unique, and without additional constraints on what values $s_m(k)$ should take, estimating $s(k)$ beyond $|k_{max}|$ represents an ill-posed problem. Additional information to constrain the set of possible solutions is invaluable. Such information may be that the inverse Fourier Transform of the underlying function that we wish to estimate has finite spatial support, which describes Problem 2 (in the continuous case). Several approaches to Problem 2 have been addressed that implement the use of this additional constraint, including the Prolate Spheroidal Wave Function (PSWF) algorithm [91], and the Papoulis Gerchberg (PG) algorithm [77]. For Problem 2, these algorithms have been shown to converge to a unique solution.

Problem 3 is a closer representation of an extrapolation problem found in MRI. There are widely studied techniques in MRI which solve the two-dimensional discrete extrapolation problem by using a series of such one-dimensional discrete extrapolations (see for example [4, 35, 68, 92]). Unfortunately however, the solution to Problem 3 is also non-unique. Appendix A.1 provides a brief mathematical analysis of this case.

Problem 4 - Two-Dimensional Discrete Extrapolation: Consider a continuous, complex valued function, $s_{true}(k_x, k_y) \in L^2(\mathbb{R}^2)$, defined for all values of the real variables k_x and k_y, and which has compact support in the spatial domain. $s_{true}(k_x, k_y)$ is measured over a finite set of values to produce

$$\begin{aligned} s_d(m_x \Delta k_x, m_y \Delta k_y) &= s_{true}(m_x \Delta k_x, m_y \Delta k_y) \\ m_x &= -N_x/2 + 1, \ldots, N_x/2 - 1 \\ m_y &= -N_y/2 + 1, \ldots, N_y/2 - 1, \end{aligned} \tag{1.7}$$

for N_x and N_y even. Given s_d, the two-dimensional extrapolation problem attempts to determine the values outside of the measured domain. In MRI, we may wish to only extrapolate in the k_x or k_y direction, depending on which acquisition scheme is used and how the data is acquired.

1.3 Methodology

An enormous amount of literature can be found on the use of *a priori* knowledge in *constrained reconstruction*, or *super-resolution* (SR) algorithms in MRI to decrease Δx through a variety of post-processing strategies (see [53, 55] for literature reviews of research on these algorithms). This thesis explores both algorithms that employ multiple images and algorithms that employ only a single image to perform resolution enhancement.

1.3.1 Multiple Image SRMRI

A relatively recent approach to enhancing the spatial domain resolution in MRI utilizes *multiple* low-resolution images of the object, thereby introducing new *a priori* conditions to the reconstruction [11, 30, 31, 32, 33, 46, 47, 62, 66, 70, 71, 72, 75, 79, 80, 81, 82, 86, 90]. The underlying concepts behind multiple acquisition SRMRI (MASRMRI) was presented earlier outside of the MRI literature, in the area of super-resolution image processing. In this area of research, the input data are images of a moving scene, possibly taken with a moving imaging apparatus [7, 13].

Another related approach to SRMRI is applied to computed tomogrophy (CT). Spiral CT [43] uses a continuously rotating scanner during a single "breath hold" while the object is shifted at a constant rate through the gantry. Depending upon the rate at which the object is translated, the object may be imaged multiple times. This redundancy reduces noise and motion artifacts in the image in the data.

The earliest published results of SR applied to MRI (to the best of the author's knowledge) can be found in a year 2000 conference proceedings [86]. In the algorithms that have so far been proposed, repeated imaging of an object with known *spatial* shifts is employed. After N_{img} images, or channels, have been acquired, a deconvolution algorithm is used to produce a synthesized image with higher resolution than that of the low-resolution images. To implement the deconvolution algorithm, a point-spread function (PSF) is required that models the relationship between the low-resolution data and a desired high-resolution image. By incorporating information from multiple channels and the PSF, these approaches attempt to produce an image of higher image quality than that obtained with a conventional IDFT of the measured frequency data. A potential

drawback however, is an increase in the total scan time, T_{scan}, by requiring multiple images to be acquired.

Indeed, we would expect (or hope) that under sufficient conditions, using more data would imply that more information would be incorporated in the reconstruction algorithm. If this were always true, then MASRMRI may have a significant edge over single acquisition techniques. Otherwise, there would be no adequate justification to incur a scan time penalty by acquiring more data. The *conditions under which collecting more data sets yields more information are therefore of importance.*

Controversy over the validity of implementing a MASRMRI approach recently appeared in the MRI literature [82, 90]. Verifying that acquiring multiple acquisitions even constitutes a viable approach to enhance spatial resolution in MRI had not been investigated[8], and rests on the following three issues.

1. *How the spatial shift is produced:* In the FE and PE directions, this shift has only been produced by multiplying phase ramps to the measured data. The order in which this processing step is added to other processing steps (such as sampling and filtering) affects the efficacy of SRMRI.

2. *The extent to which the PSF can be determined:* Most SRMRI research assumes a Gaussian or boxcar, $\Pi(x)$, form for the PSF without basing this choice on the underlying physics or mathematical properties of the acquisition process.

3. *Most importantly, the nature of the new information present in each acquisition:* If a significant amount of new information is present in each acquisition after the first, exactly what new information about the object being imaged is obtained.

These points are explored in Chapter 5 to consider the validity of using multiple acquisitions to enhance spatial resolution in MRI.

Sensitivity Encoding (SENSE)

An altogether different approach that yields high spatial resolution data in MRI incorporates an *array* of RF coils [84]. Sensitivity encoding (SENSE) is based on the fact that receiver sensitivity generally has an encoding effect that can be exploited to dramatically reduce the total scan time while maintaining a high spatial resolution[9]. This research area requires specialized hardware, and as such, was not explored in this thesis. A specific reconstruction strategy that use a superreceiver array is mentioned Chapter 10 that may be of interest for future research.

[8] Prior to the start of the author's PhD programme in January 2004.

[9] This reduction in scan time is achieved by reducing of the number of Fourier encoding steps (by increasing the space between sample positions in k-space). The local coil sensitivity profiles provide a means of combining information from multiple acquisitions from individual receiver coils.

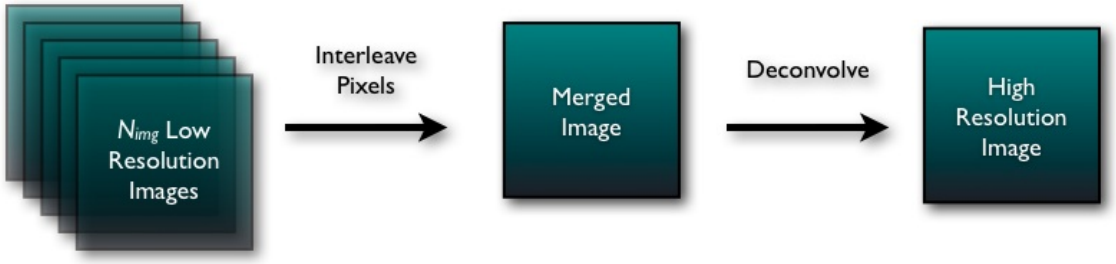


Figure 1.2: The classic SR algorithm involves the acquisition of N_{img} images, with a field-of-view that translates between each acquisition. The pixels of each acquisition are interleaved to create the merged data, which is used as an input to a deconvolution algorithm that calculates a high resolution image.

1.3.2 Single Image SRMRI

One such SRMRI technique that has been studied by a number of research groups is the projection over convex sets (POCS) algorithm [37, 36, 53, 55, 60, 68]. POCS provides a very flexible approach to SRMRI by allowing the incorporation of many *a priori* conditions into the reconstruction[10].

Another approach to MRI resolution enhancement that has been extensively studied implements the auto-regressive moving average (ARMA) technique (see for example [4, 35, 68, 92]). A special case of the ARMA technique is the linear prediction model, discussed in Section 4.1.3, and Chapters 8 and 9.

Iterated Fourier Transform Systems

New approaches to MRI spatial resolution enhancement using only a *single* acquisition are considered in Chapters 7, 8, and 9. Chapter 7 introduces a new iterative fractal-based operator that operates directly on Fourier domain data, and explores some of its mathematical properties. This operator naturally exhibits an extrapolation relationship, suggesting that it can be used for spatial resolution enhancement. Experiments on simulated data suggest the operator can be used to perform extrapolation. This operator however is not yet connected to to a physical model of the MRI data, and unlike the methods discussed in Chapters 8 and 9, there is currently no physical basis for this approach for SRMRI.

[10]Research has focused on the use of a very limited amount of the available *a priori* information, based on the constraints listed in Eq. 3.41. More research is needed to further the effectiveness and better understand this already successful approach to SRMRI.

Self-Similar and Linear Prediction Methods

Spatial domain images can possess a property known as *self-similarity*. Intuitively, for a two-dimensional discrete image $u(m,n)$ to possess self-similarity, we mean that a given subblock, $\mathbf{r}$, of u is well approximated by a linear combination of N_{DB} subblocks, $\mathbf{d}_n,\ n = 1,2,3,\ldots,N_{DB}$, found elsewhere in the data set. Affine mappings on the subblocks may be allowed, so that:

$$\mathbf{r} = \beta + \sum_{n=1}^{N_{DB}} \gamma_n \mathbf{d}_n, \tag{1.8}$$

where parameters β and γ are real numbers, that may be determined, for example, by using a least-squares procedure. Eq. 1.8 describes a *multi-parent* transform. If a sufficient approximation to $\mathbf{r}$ can be found using only a set of subblocks $\mathbf{d}_n$ taken only from a *local* neighbourhood of $\mathbf{r}$, then the image possess a certain degree of *local self-similarity*. A more thorough description of self-similarity is provided in Section 6.2.

In Chapter 8, we show that one-dimensional discrete functions of the form[11]

$$s(m) = \sum_{n=1}^{N_c} c_n e^{\nu_n m},\ m = 0,1,2,\ldots,N_s-1,\quad c_n,\nu_n \in \mathbb{C}, \tag{1.9}$$

are locally self-similar. This connection is made by making use of the fact that signals of the form 1.9, are linearly predictable, so that

$$s(m) = \sum_{n=1}^{N_c} a_n s(m-n),\ m = N_c, N_c+1, N_c+2,\ldots N_s-1,\quad a_n \in \mathbb{C}. \tag{1.10}$$

This connection establishes a physical basis for the use of self-similar techniques in MRI data extrapolation and resolution ehnahcement. These theoretical results are derived for the one-dimensional case, and a two-dimensional analysis is presented in Chapter 9.

Empirical results in Chapter 8 provide further evidence of the self-similarity of frequency domain data. Collage error histograms are used to study the self-similarity of measured data, derived from both phantom and human brain MRI data.

1.4 Thesis Organization

The organization of this thesis is as follows. Chapters 2 to 4 provide a background of MRI physics and review basic approaches found in the literature that perform SRMRI. Chapter 2 introduces the necessary background on one and two-dimensional Fourier theory required for the remaining chapters in this thesis. The treatment is brief, and the

[11] Eq. 1.9 has been used to model the measured MR data in order to improve the image reconstruction process and to reduce artifacts [34].

interested reader may wish to consult [8, 9], or [76] for a more thorough review of Fourier theory.

Chapter 3 provides a literature review of MR physics. Section 3.1 provides a derivation of the Bloch equations from quantum mechanical relationships. These equations are used to model the MRI signal. Section 3.2 provides a mathematical model of the image acquisition process, used in later sections of the thesis.

Chapter 4 is a literature review of SRMRI methodologies. Section 4.1 explores *single* acquisition image reconstruction techniques, and Section 4.2 explores *multiple* acquisition reconstruction techniques found in the MRI literature.

Chapter 5 first discusses the use of multiple acquisitions in SRMRI in the frequency encoding direction[12]. This chapter introduces new methods to ascertain the amount of new information in each acquisition after the first.

Section 6.1 provides a literature review of iterated function systems, necessary for Chapter 7, which introduces a fractal-based method over (complex-valued) Fourier data[13]. This chapter describes the method of iterated Fourier Transform systems, which has a natural mathematical connection with the fractal-based method of iterated function systems with greyscale maps (IFSM) in the spatial domain.

Chapter 8 presents a theoretical basis for the author's work on the self-similarity of MRI data by developing a model of the measured signal in one dimension. This chapter shows that discrete functions of this form are locally self-similar. Section 8.3 presents a statistical analysis of the local, two-dimensional self-similarity of human volunteer and phantom MRI frequency data[14]. Self-similarity of these MR images is described with collage error histograms [1]. Our results suggest that complex MRI data can be locally self-similar using multi-parent transforms.

Chapter 9 presents recent, unpublished work that expands on the ideas in Chapter 8. This chapter provides a basis for future work on the self-similarity of Fourier domain MRI data (i.e. - complex and discrete trigonometric polynomials). A partial list of potential research projects is provided in Section 10.3.2.

[12]Results presented in Chapter 5 are based on the author's published research [63, 64, 66].
[13]Results presented in Chapter 7 are based on the author's published research [65].
[14]Results presented in Chapter 8 is an overview of the author's published research [67].

Chapter 2

Fourier Theory

This chapter is a review of material present in the literature on one and two-dimensional Fourier Transforms. As such, the notation and concepts used in this chapter are also presented in well-known texts on Fourier theory (for example, see [8, 61, 59, 76]). The purpose of this chapter is to introduce concepts relevant to later chapters of this thesis.

2.1 The One-Dimensional Fourier Transform

Throughout this thesis we will be working with the L^p function spaces, i.e., $L^p(X)$, $p = 1, 2, \cdots$, where

$$L^p(X) = \left\{ u : X \to \mathbf{R} \mid \| u \|_p = \left[\int_X |u(x)|^p \right]^{1/p} dx \ < \ \infty \right\}. \tag{2.1}$$

We will also work with complex functions that operate on real numbers, $f(x) : \mathbb{R} \to \mathbb{C}$, and are in $L^2(\mathbb{R})$, the space of square integrable complex valued functions:

$$\int_{-\infty}^{\infty} \left| f(x) \right|^2 dx < +\infty. \tag{2.2}$$

Any function $f \in L^2(\mathbb{R})$ has a Fourier Transform, $F(k) : \mathbb{R} \to \mathbb{C}$, defined as

$$\mathcal{F}\{f\} = F(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x k} dx, \ x, k \in \mathbb{R}. \tag{2.3}$$

For the purposes of this thesis, variables x and k are spatial and spatial-frequency variables respectively.

Functions $f(x)$ and $F(k)$ constitute a *Fourier Transform pair*. Several Fourier Transform pairs used throughout this thesis are provided in Table 2.1. Proofs of the results in this table can be found in the literature [8, 38, 59]. Variables in this table are dimensionless; many signal processing textbooks, including Bracewell [8], consider dimensionless variables. Elsewhere in this thesis, however, variables may or may not have units.

$f(u) = \int_{-\infty}^{\infty} F(v)e^{+2\pi iuv}dv$	$F(v) = \int_{-\infty}^{\infty} f(u)e^{-2\pi iuv}du$
$\delta(u - u_0)$	$e^{-i2\pi u_0}$
$\mathrm{III}(u) = \sum_{n=-\infty}^{\infty} \delta(u - n)$	$\mathrm{III}(v) = \sum_{n=-\infty}^{\infty} \delta(v - n)$
$\Pi(u) = \begin{cases} 0 & \lvert u\rvert > \frac{1}{2} \\ \frac{1}{2} & \lvert u\rvert = \frac{1}{2} \\ 1 & \lvert u\rvert < \frac{1}{2} \end{cases}$	$\mathrm{sinc}(v) = \frac{\sin(\pi v)}{\pi v}$
$\Lambda(u) = \begin{cases} 1 + u, & 1 < u < 0 \\ 1 - u, & 0 \le u < 1 \\ 0, & \text{otherwise} \end{cases}$	$\big(\mathrm{sinc}(v)\big)^2 = \frac{\sin^2(\pi v)}{(\pi v)^2}$
$e^{-\pi u^2}$	$e^{-\pi v^2}$

Table 2.1: One-dimensional Fourier Transform pairs of univariate functions used throughout this thesis. Variables in this table are dimensionless.

Several well-known Fourier theorems that are used throughout this thesis are provided in Table 2.2. Proofs of these theorems may be found throughout the literature [8, 38, 59].

$f(u) = \int_{-\infty}^{\infty} F(v)e^{+2\pi iuv}dv$	$F(v) = \int_{-\infty}^{\infty} f(u)e^{-2\pi iuv}du$	Theorem
$\frac{df}{du}$	$+2\pi ivF(v)$	Derivative Theorem
$\int_{-\infty}^{\infty} f_1(u)f_2^*(u_0 - u)du$	$F_1(v)F_2(v)$	Convolution Theorem
$f(u - u_0)$	$e^{-2\pi ivu_0}F(v)$	Shifting Theorem

Table 2.2: One-dimensional Fourier Transform theorems used throughout this thesis. $f(u)$, $f_1(u)$, $f(u)$, and $F(v)$ are complex functions in $L^2(\mathbb{R})$. Variables in this table are dimensionless.

The inverse Fourier Transform, defined as follows, returns the original function $f(x)$ from $F(k)$

$$f(x) = \int_{-\infty}^{\infty} F(k)e^{+2\pi ixk}dk, \ x \in \mathbb{R}. \tag{2.4}$$

Thus, if $F(k)$ for all $k \in \mathbb{R}$ were known, $f(x)$ could be obtained for all x using Eq. 2.4. In MRI, it is unrealistic to obtain measurements over the entire real line, and $F(k)$ is only measured over a bounded subset of $\mathbb{R}$. One of the objectives of the research presented in this thesis is to estimate $F(k)$ over an interval larger than that which is measured in order to better estimate its inverse Fourier Transform, $f(x)$.

The definition used for the Fourier Transform (Eq. 2.3) follows from what is conventional in the MRI literature. Different definitions of the Fourier Transform and its inverse can be found in other areas of research. For example, one such alternative definition uses

angular frequency in place of frequency, with the substitution $\omega = 2\pi k$ in Eq. 2.3. This substitution would require a different definition for the inverse transform, alternative results for the transform pairs in Table 2.1, and so forth. This thesis, without exception, uses the definition stated in Eq. 2.3 to avoid unnecessary confusion.

It will be useful to consider sets of functions that are *band-limited*[1] and *spatially-limited*. A band-limited function $f(x)$ is one for which its Fourier Transform $F(k)$ has compact support over $k \in [a, b]$, for finite $a, b \in \mathbb{R}$. For convenience, the interval of support of $F(k)$ is often chosen to be the symmetric interval $\mathcal{K} = [-K/2, K/2]$, for an appropriate value of $K > 0$. Consequently, a band-limited function $f(x)$ satisfies, for some constant $K \in \mathbb{R}$,

$$f(x) = \int_{-\frac{K}{2}}^{+\frac{K}{2}} F(k) e^{+2\pi i x k} dk, \tag{2.5}$$

so that

$$F(k) = \Pi(\frac{k}{K}) F(k). \tag{2.6}$$

The function $\Pi(k)$ is defined in Table 2.1.

In MRI, we measure signals in the Fourier domain that come from objects that have finite size in the spatial domain. As such, throughout this thesis, we will be working with k-space functions whose inverse Fourier Transforms have compact support. It will be convenient therefore to introduce the set of all spatially-limited functions, $\mathcal{R}_R$, as the space that contains *frequency domain* functions whose inverse Fourier Transforms have compact support:

$$\mathcal{R}_R = \left\{ F(k) \in L^2(\mathbb{C}) \mid F(k) = \int_{-\frac{R}{2}}^{+\frac{R}{2}} f(x) e^{-2\pi i x k} dx, x \in \mathbb{R} \right\}, \tag{2.7}$$

for some appropriate value of $R \in \mathbb{R}^+$. This space is used in MR constrained reconstruction algorithms discussed in Chapter 4.1[2].

2.2 Mathematical Definition of the MRI Spatial Sample Spacing

The central point of this section is the derivation of the equation for the spatial sample spacing in MRI (Eq. 2.13, also Eq. 1.5) in terms of only two data acquisition parameters,

[1] Interestingly, the signal and image processing literature uses different spellings for this term; some authors [61, 91, 42] use "bandlimited", while others [42, 77, 88, 10, 94, 9] prefer to use "band-limited". The latter was arbitrarily chosen for this thesis.

[2] The primary example of the use of this space in MR reconstruction is the Papoulis-Gerchberg algorithm and its generalizations.

the number of measurements, N, and the frequency sample spacing, Δk. This fundamental equation provides the central mechanism for performing resolution enhancement: frequency extrapolation.

In MRI, *spatial* representations of the data are derived from measurements acquired in the *frequency* domain. The acquired data are continuous (along a prescribed trajectory), complex, and in the frequency domain. Furthermore, before the continuous data are transformed, they must be sampled. Thus, the conventional process of transforming the sampled *frequency* data into a *spatial* representation must be understood to define the spatial sample spacing.

2.2.1 The Sampling Theorem for MR Data

One of the strengths of MRI is its enormous flexibility over how the frequency data may be measured. There are however, particular conditions that must be observed when MR data are being sampled to reduce the degree to which inevitable artifacts distort the data. To describe these conditions, let us consider a function, $F(k) \in \mathcal{R}_R$, which represents a hypothetical MRI signal measured for all $k \in \mathbb{R}$. As such, its inverse Fourier Transform is spatially limited (or has compact support). Sampling of this function at regular intervals of length Δk may be modeled as a multiplication of $F(k)$ with the comb function, $\text{III}(\frac{k}{\Delta k})$, defined in Table 2.1. Thus, the sampled signal can be expressed as, for all $k \in \mathbb{R}$ (see Appendix A.2 for a complete derivation):

$$\hat{F}(k) = F(k)\text{III}(\frac{k}{\Delta k}) = \Delta k \sum_{n=-\infty}^{\infty} F(k)\delta(k - n\Delta k) \tag{2.8}$$

An inverse Fourier Transform may be employed to obtain $\hat{f}(x)$, which is the inverse Fourier Transform of $\hat{F}(k)$. By using the convolution theorem, we know that $\hat{f}(x)$ will be periodic with unit period.

Under sufficient conditions described by the well-known Sampling Theorem, it is possible to recover $F(k)$ for all k from knowledge of its samples. The Sampling Theorem is usually described in textbooks for either the spatial or time domain [8, 61, 59]. However, the same principles can be applied to the frequency space case [34, 76] which, for obvious reasons, is of more interest to us.

Let $F(k) \in \mathcal{R}_R$ be a spatially-limited signal with spatial extent R, that is sampled with a frequency sample spacing of Δk. If $\Delta k < R^{-1}$, then the signal $F(k)$ may be expressed with knowledge of its samples using the cardinal series [61]:

$$F(k) = \sum_{n=-\infty}^{\infty} \hat{F}(n\Delta k)\text{sinc}(Rk - n), \ \forall \ k \in \mathbb{R}. \tag{2.9}$$

If, however, the sample spacing is increased beyond R^{-1}, it may not be possible to reconstruct $F(k)$, and aliasing errors may be present in $f(x)$. In MRI, we are not interested in obtaining $F(k)$ for all $k \in \mathbb{R}$. However, the sampling theorem also provides us with information on the aliasing errors that we would like to minimize[3].

Many other sampling theorems have been derived that allow the sample spacing to be increased above R^{-1} if other criteria are met or if additional information is supplied. For example, a multi-channel sampling theorem has been derived that provides conditions under which $F(k)$ may be recovered from its samples when multiple data sets are obtained [61].

2.2.2 The Spatial Sample Spacing

Thus far we have assumed that an infinite number of samples have been measured, which is not obtainable in practice. We must also consider the effect of using a finite number of measurements to form the spatial domain signal.

Consider a frequency domain signal measured over a finite interval $[-k_{max}, k_{max}]$, and uniformly sampled $N = \frac{2k_{max}}{\Delta k}$ times with a frequency sample spacing Δk:

$$\tilde{F}(k) = \Pi\Big(\frac{k}{2k_{max}}\Big)\text{III}\Big(\frac{k}{\Delta k}\Big)F(k), \qquad \forall\ k \in \mathbb{R}. \tag{2.10}$$

Eq. 2.10 represents a finite set of uniformly spaced measurements approximating the continuous signal, $F(k)$, over the interval $[-k_{max}, k_{max}]$. Using the measurements $\tilde{F}(k)$, an approximation to $f(x)$ is desired.

To approximate $f(x)$, we perform the inverse Fourier Transform of Eq. 2.10 [34] (see Appendix A.3 for a complete derivation of Eq. 2.11) to obtain

$$\begin{aligned}
\tilde{f}(x) &= \int_{-\infty}^{\infty} \tilde{F}(k)e^{i2\pi kx}dk, \qquad \forall\ x \in \mathbb{R} \\
&= \int_{-\infty}^{\infty} \text{III}\Big(\frac{k}{\Delta k}\Big)\Pi\Big(\frac{k}{2k_{max}}\Big)F(k)e^{i2\pi kx}dk \\
&= \Delta k \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} F(n\Delta k)e^{i2\pi n\Delta kx}.
\end{aligned} \tag{2.11}$$

This inverse Fourier Transform converges because, by assumption, $F(k) \in L^2(\mathbb{R})$. We have not yet assumed that $\tilde{F}(k)$ is periodic, although it can be shown that $\hat{f}(x)$ is periodic with period $\frac{1}{\Delta k}$.

Due to the nature of the inverse Fourier Transform, Eq. 2.11 is valid for all x, but in practice, it is feasible to perform reconstructions for only a *finite* number of x values.

[3] Aliasing errors may be reduced with the application of an anti-aliasing filter. This process will be of importance in Chapter 5.

Hence, a discretization over spatial steps, of uniform length Δx must be introduced. To arrive at a sampled set of spatial coordinates, consider the effect of making $\tilde{F}(k)$ periodic, with period $2k_{max}$:

$$\tilde{F}_p(k) = \text{III}(\frac{k}{2k_{max}}) * \tilde{F}(k).$$

With this new periodic signal, we cannot implement the inverse Fourier Transform to obtain a spatial representation because $\hat{F}_p(k) \notin L^2(\mathbb{R})$. Reconstruction is instead carried out using an inverse discrete Fourier Transform (IDFT):

$$\tilde{f}_p(n\Delta x) = \sum_{p=-\frac{N}{2}}^{\frac{N}{2}-1} \tilde{F}_p(p\Delta k) e^{i2\pi(p\Delta k)(n\Delta x)}, \qquad n = 0, 1, 2, \cdots, N-1. \tag{2.12}$$

The reconstructed signal is periodic: $\tilde{f}_p(n\Delta x) = \tilde{f}_p(n\Delta x + \frac{1}{\Delta k})$. Since the spatial sample spacing is equal to the period length divided by the number of measurements, it follows that:

$$\Delta x = \frac{1}{N\Delta k}. \tag{2.13}$$

This fundamental relationship provides an important concept behind this project. Since, by Eq. 2.13, the sample spacing is inversely related to the spectral extent, we can reduce the sample spacing by incorporating higher frequencies in the partial sum in Eq. 2.12. SR methods attempt to estimate higher frequencies given a set of low frequency measurements thereby enhancing the spatial resolution. This project investigates methods of increasing the spectral extent by incorporating *a priori* knowledge of the object being imaged and the acquisition sequence employed.

2.3 The Two-Dimensional Fourier Transform

The two-dimensional Fourier Transform is a straightforward extension of its one-dimensional analogue[4], and is reviewed because of its extensive use throughout this thesis. The Fourier Transform of a two-dimensional function $f \in L^2(\mathbb{R}^2)$ is:

$$F(k_x, k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-2\pi i x k_x} e^{-2\pi i y k_y} dx dy.$$

The inverse of the above two-dimensional Fourier Transform is given by

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(k_x, k_y) e^{+2\pi i x k_x} e^{+2\pi i y k_y} dk_x dk_y.$$

This equation bears a fundamental importance for this thesis, as raw MR data is given in terms of a two-dimensional Fourier Transform. An approximation to its inverse is used to construct a spatial domain image.

[4]Higher dimensional generalizations of the Fourier Transform can easily be made, but will not be needed for this thesis.

The following sections list properties of the two-dimensional Fourier Transform, frequently used throughout this thesis. For other well known Fourier theorems (such as the convolution theorem, the Parseval theorem, and so forth), the interested reader will be referred to relevant areas of the image processing literature.

Shift Theorem

The shifted image

$$f(x - x_0, y - y_0),$$

has the Fourier Transform

$$e^{-2\pi i(x_0 k_x + y_0 k_y)} F(k_x, k_y).$$

The shift theorem bears a key importance in our discussion of super-resolution methods that use multiple images related by spatial translations. We use this shift theorem throughout Chapter 5.

Rotation Theorem

We use the rotation theorem in our discussion of the projection slice theorem. It will be convenient to first define the following rotation operator:

$$g(\hat{x}, \hat{y}) = \mathbf{R}_\theta\big\{f(x, y)\big\} = f(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$$
$$\text{where : } \begin{array}{l} \hat{x} = x\cos\theta - y\sin\theta \\ \hat{y} = x\sin\theta + y\cos\theta \end{array} .$$

The operator $\mathbf{R}_\theta\big\{f(x, y)\big\}$ rotates $f(x, y)$ by an angle θ. Here, we measure θ in radians, in a counter clockwise direction from the x-axis[5].

The rotation theorem for the two-dimensional Fourier Transform states that a rotation of $f(x, y)$ by an angle θ, rotates $F(k_x, k_y)$ by an angle θ. The two-dimensional Fourier Transform of $\mathbf{R}_\theta\big\{f(x, y)\big\}$ is simply:

$$F(k_x\cos\theta - k_y\sin\theta, k_x\sin\theta + k_y\cos\theta).$$

In other words, the counterclockwise rotation of $f(x, y)$ results in a counterclockwise rotation of $F(k_x, k_y)$. A proof of this theorem is provided in the image processing literature [98].

[5]Naturally, in k-space we measure θ in a counter clockwise direction from the k_x-axis.

The Projection Slice Theorem

The projection slice theorem (see [9], Chapter 14) is used in our discussion of phase encoding (Section 3.2.5). This section provides a description of this theorem, which first requires a definition of a particular projection operator.

We define the projection of an image onto the straight line, S, that passes through the origin at angle θ, with the following line integral:

$$g(\hat{x}) = \mathbf{P}_\theta\big\{f(x,y)\big\} = \int_{-\infty}^{\infty} \mathbf{R}_{-\theta}\big\{f(x,y)\big\}d\hat{y}.$$

The projection slice theorem states that $\mathbf{P}_\theta\big\{f(x,y)\big\}$ is the inverse Fourier Transform of $F(k_x,k_y)$ along the line passing through the k-space origin at angle θ. To see this, let us first consider the two-dimensional Fourier Transform:

$$F(k_x,k_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-2\pi i x k_x}e^{-2\pi i y k_y}dxdy.$$

We see that setting $k_y = 0$, yields:

$$\begin{aligned} F(k_x,0) &= \int_{-\infty}^{\infty}\Big\{\int_{-\infty}^{\infty} f(x,y)dy\Big\}e^{-2\pi i x k_x}dx \\ &= \int_{-\infty}^{\infty}\Big\{\mathbf{P}_{\theta=0}\big\{f(x,y)\big\}\Big\}e^{-2\pi i x k_x}dx. \end{aligned}$$

This demonstrates that $F(k_x,0)$ is the one-dimensional transform of $\mathbf{P}_0\big\{f(x,y)\big\}$ (the projection of $f(x,y)$ onto the x-axis).

A more general situation can be described for any straight line, S, passing through the k-space origin at an angle θ. A straight line path, S, through Fourier data $F(k_x,k_y)$, is equal to the one-dimensional Fourier Transform of the projection of $f(x,y)$ onto that same line:

$$G(\hat{k}_x) = F(k_x,k_y)\big|_S = \int_{-\infty}^{\infty}\mathbf{P}_\theta\big\{f(x,y)\big\}e^{-2\pi i(\hat{x}\hat{k}_x)}d\hat{x},\ \ \hat{k}_x \in \mathbb{R}. \tag{2.14}$$

As before, the line S travels through the k-space origin at angle θ. We may interpret the right hand side of Eq. 2.14 as a rotation of $f(x,y)$, a projection of the rotated function onto the $\hat{x}$-axis, and finally a transform of the projected function into the Fourier domain. Although this equation is valid for any $\theta \in \mathbb{R}$, in subsequent chapters, we only consider projections onto lines parallel to one of the coordinate axes.

Chapter 3

MR Physics and Data Acquisition

This chapter is a literature review of material needed for results presented in later chapters of this thesis. Section 3.1 provides a derivation of the Bloch equations, which are used to model the process by which the MR signal is generated and detected. A mathematical model of typical MR data acquisition and processing steps is provided in Section 3.2. This model will be used for example, in Chapter 5, to describe the efficacy of SRMRI methods in the literature.

3.1 The Bloch Equations

In this section, fundamental concepts behind nuclear magnetic resonance (NMR) are introduced by illustrating how manipulations of the proton spin orientations within a sample are induced by a collection of applied magnetic fields. The classic NMR experiment employs a main magnetic field $\vec{B} = B_0\hat{k}$ to produce a net magnetization vector $\vec{M} = [M_x, M_y, M_z]^T$ within an object of interest. Manipulations of this vector through applied radiofrequency (RF) fields produces an NMR signal containing information about the object.

This section also discusses the nature of the Bloch equations. These equations and their solutions are not used explicitly in other chapters of this thesis. However, the Bloch equations are a fundamental concept for the understanding of how data is acquired in MRI. The discussion in this chapter follows that presented in [34].

3.1.1 Quantization of Angular Momentum

Throughout this chapter[1] we describe the quantization of angular momentum for the hydrogen proton. Many quantum particles (electrons, protons, neutrons, etc) possess an

[1] A more detailed discussion of the material in this chapter may be found in Chapters 5 and 6 of [34].

intrinsic angular momentum or spin (analogous to how a spinning top precessing about the gravity axis possesses an angular momentum). We let $\vec{S}$ denote the spin angular momentum of the proton. As the proton has a non-zero spin, it also has an angular momentum, $\vec{\mu}$, proportional to its spin.

A measurement of the z-component of any atomic or nuclear angular momentum, $\vec{J}$, yields integer (or half-integer) multiples of[2] $\hbar$

$$J_z = m_j \hbar, \tag{3.1}$$

m_j is called the magnetic quantum number, and the $2j+1$ values of m_j are

$$m_j = -j, -j+1, -j+2, \ldots, j-1, j. \tag{3.2}$$

The number $j \in \mathbb{Z}$ or $j \in \mathbb{Z}/2$ holds a relation to the total angular momentum, $\vec{J}$, according to

$$J^2 = j(j+1)\hbar^2, \tag{3.3}$$

where $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$, and $J = ||\vec{J}||$. The total angular momentum for an atomic or nuclear system has a contribution from orbital motion $\vec{L}$ and a contribution from its intrinsic spin $\vec{S}$:

$$\vec{J} = \vec{L} + \vec{S}. \tag{3.4}$$

The intrinsic spin, $\vec{S}$, has a magnitude $S = ||\vec{S}||$ that satisfies

$$S^2 = s(s+1)\hbar^2, \qquad s = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots \tag{3.5}$$

The z-component S_z of $\vec{S}$ is given by $S_z = m_s \hbar$. For this thesis, we are mainly concerned with the hydrogen proton. We assume this nuclear species has no orbital motion, implying that $\vec{J} = \vec{S}, j = s$, and $m_j = m_s$.

The quantum mechanical spin angular momentum, $\vec{S}$, is an operator that can be represented by finite matrices. The matrix representation for the spin one-half vector operator is given by

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma}, \tag{3.6}$$

where the three linearly independent *Pauli spin matrices* are

$$\sigma_{\mathbf{x}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_{\mathbf{y}} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_{\mathbf{z}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.7}$$

It has been shown experimentally that the proton has spin one-half, so that $s = \frac{1}{2}$ and $m_s = \pm\frac{1}{2}$.

[2] $\hbar$ is Planck's constant divided by 2π. Planck's constant is approximately 6.626×10^{-34} J s.

3.1.2 Precession Frequency

In this section, the relationship between precession and quantum energy differences is made, which will be needed for the quantum analysis of proton spin. A hydrogen proton, in the presence of a static magnetic field, $\vec{B} = B_0\hat{k}$, has a potential energy[3] U, given by

$$U = -\vec{\mu} \cdot \vec{B} = -\mu_z B_0 = -\gamma J_z B_0 = -\gamma m_s \hbar B_0. \tag{3.8}$$

Here, $\vec{\mu}$ is the magnetic moment vector for an individual spin, and γ is the gyromagnetic (or magnetogyric) ratio[4]. The empirical relation $\vec{\mu} = \gamma\vec{J}$, and Eq. 3.1 has been used to derive Eq. 3.8. The magnitude of the energy absorbed or released by the proton spin system upon a transition between higher ($m_s = -1/2$) and lower energy ($m_s = 1/2$) states is found to be

$$\begin{aligned} \Delta E &= E_{m_s=-\frac{1}{2}} - E_{m_s=+\frac{1}{2}} \\ &= \frac{1}{2}\gamma\hbar B_0 - (-\frac{1}{2}\gamma\hbar B_0) \\ &= \hbar\omega_0. \end{aligned} \tag{3.9}$$

Here, ω_0 is the *Larmor precession frequency*, given by:

$$\omega_0 = \gamma B_0. \tag{3.10}$$

This is an example of the Larmor equation[5], a particularly important relation in MRI that is used throughout the remainder of this chapter.

3.1.3 Spin Solutions for Constant Fields

Consider the magnetic-moment interaction of a proton at rest and immersed in the constant external field, $\vec{B} = B_0\hat{k}$. We are assuming that no orbital motion is present. Let us show how the solution of the Schrödinger equation (for zero kinetic energy) leads to the quantum numbers and matrix elements introduced above.

The potential energy of this system, U, is then given by

$$U = -\mu \cdot \vec{B} = -\gamma J_z B_0 = \omega_0 J_z = -\omega_0 S_z = \frac{-\omega_0\hbar}{2}\sigma_z. \tag{3.11}$$

The matrix representation for the Hamiltonian, H, becomes

$$H = U = \frac{-\omega_0\hbar}{2}\sigma_z = \begin{pmatrix} -\frac{1}{2}\hbar\omega_0 & 0 \\ 0 & +\frac{1}{2}\hbar\omega_0 \end{pmatrix}. \tag{3.12}$$

[3]This is an example of the general Zeeman effect, where atomic moments in the presence of a magnetic field produce splittings in the atomic energy levels.

[4]The gyromagnetic ratio is different for each nuclear species. For hydrogen, $\gamma/2\pi$ is roughly 42.58 MHz/T.

[5]Named after Sir Joseph Larmor, who was first to demonstrate the precession of charged particles in 1897 [69].

Clearly, the Pauli representation for σ_z leads to a diagonal Hamiltonian, whose eigenvectors are (by inspection):

$$\psi_+ := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \psi_- := \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.13}$$

Indeed, the use of a 2-by-2 matrix representation for J_z in Eq. 3.11, implies that the wave function ψ on which it operates is a column matrix with two entries. The eigenvalues of $H = -\gamma B_0 S_z$, that satisfy $H\psi_\pm = E_\pm \psi_\pm$, are given by:

$$E_\pm = \mp \frac{1}{2}\hbar\omega_0. \tag{3.14}$$

Note also that ψ_+ and ψ_- are also eigenfunctions of S_z,

$$S_z\psi_\pm = \pm\frac{1}{2}\hbar\psi_\pm. \tag{3.15}$$

The eigenvalues are then $m_s = \pm\frac{1}{2}$. Using these eigenvalues and associated eigenfunctions the general solution to the time-dependent Schrödinger differential equation

$$H\Psi = i\hbar\frac{\partial\Psi}{\partial t}, \tag{3.16}$$

has the form

$$\Psi(t) = C_{\frac{1}{2}}\psi_{\frac{1}{2}}e^{\frac{1}{i\hbar}E_{\frac{1}{2}}t} + C_{-\frac{1}{2}}\psi_{-\frac{1}{2}}e^{\frac{1}{i\hbar}E_{-\frac{1}{2}}t}. \tag{3.17}$$

Here, and throughout this thesis, t refers to time. In Eq. 3.17, we assumed that the wave function has no spatial dependence since we are neglecting any motion by the proton.

3.1.4 Derivation of Quantum Precession

In this section, we use quantum mechanics to derive the dynamics of the magnetic moment for a proton at rest in a *static* field. This leads to a precession of the quantum expectation value of the magnetic moment that is an important concept of nuclear magnetic resonance.

It can be shown that the expectation values of the magnetic moment vectors are given by (see [34], page 75):

$$< \mu_x(t) > = \frac{\gamma\hbar}{2}\sin 2\theta\cos(\phi_0 - \omega_0 t) \tag{3.18}$$

$$< \mu_y(t) > = \frac{\gamma\hbar}{2}\sin 2\theta\sin(\phi_0 - \omega_0 t) \tag{3.19}$$

$$< \mu_z(t) > = \frac{\gamma\hbar}{2}\cos 2\theta \tag{3.20}$$

Definitions of constants ϕ_0 and θ, and a complete derivation of Eq. 3.19 can be found in Appendix A.4. The expectation values for the magnetic moment components describe the behaviour of a vector with magnitude $\gamma\hbar/2$. This vector rotates about the z-axis at a fixed polar angle 2θ (Fig. 3.1). Further manipulations of the magnetic moment are obtained with additional magnetic fields, described in the following section.

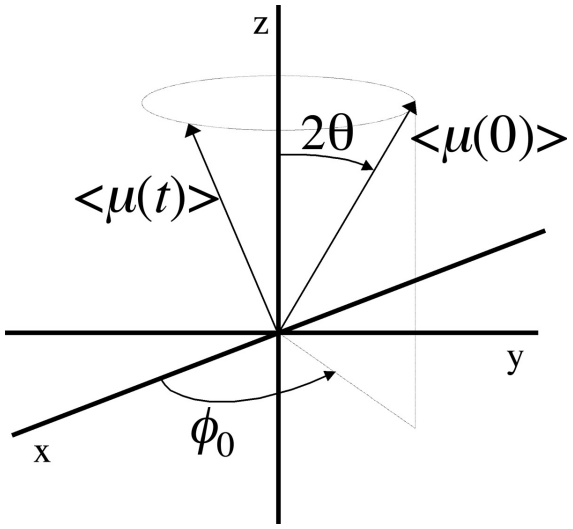


Figure 3.1: Precession of the quantum expectation value of the magnetic moment operator in the presence of a constant external field pointed in the z-direction.

3.1.5 RF Spin Tipping

This subsection introduces a time varying RF field, $\vec{B}_1(t)$, which plays an important role in NMR studies. Applying energy that the object absorbs, the RF field tips $\vec{M}(t)$ away from its equilibrium position by applying a torque to rotate the net magnetization by a desired angle[6]. The rotating magnetization vector (oscillating at the Larmor frequency of the sample) then produces an electromotive force in an RF coil that can detect changes in the magnetization in the transverse plane. This produces the signal that is measured for NMR studies.

Specifically, we are interested in obtaining the nuclear magnetic moment expectation value, in the presence of a combined static and circularly polarized field:

$$\vec{B}(t) = B_0\hat{k} + B_1[\hat{i}\cos(\omega_{rf}t) - \hat{j}\sin(\omega_{rf}t)]. \tag{3.21}$$

B_1 is the (constant) magnitude of the time-varying component of the RF field, rotating at frequency ω_{rf}. This particular set of applied fields is interesting, because of its relevance to our discussion of signal excitation and slice encoding, which we discuss in Section 3.2.3. In this section, we discuss the magnetic moment operator expectations under the presence of $\vec{B}(t)$.

[6]The RF pulse can rotate $\vec{M}(t)$ by a desired amount that depends upon the strength and duration of the pulse. For example, if the excitation is set to rotate $\vec{M}(t)$ by 90 degrees, $\vec{M}(t)$ will be in the transverse plane after the RF pulse is turned off.

It is convenient to replace the conventional Cartesian basis with a rotating basis:

$$\begin{aligned}\hat{i}_{rot} &= \hat{i}\cos\omega_{rf}t + \hat{j}\sin\omega_{rf}t \\ \hat{j}_{rot} &= \hat{j}\sin\omega_{rf}t - \hat{i}\cos\omega_{rf}t \\ \hat{k}_{rot} &= \hat{k},\end{aligned} \tag{3.22}$$

so that the applied magnetic field becomes:

$$\vec{B}(t) = B_0\hat{k} + B_1\hat{i}_{rot}. \tag{3.23}$$

With this rotating coordinate system, it can be shown that the expectation values are given by ([34], page 80):

$$< \mu_x(t) > \;=\; \frac{\gamma\hbar}{2}\sin\Theta\cos\Phi \tag{3.24}$$

$$< \mu_y(t) > \;=\; \frac{\gamma\hbar}{2}\Big[\cos\Theta\sin(\omega_1 t) + \sin\Theta\sin\Phi\cos(\omega_1 t)\Big] \tag{3.25}$$

$$< \mu_z(t) > \;=\; \frac{\gamma\hbar}{2}\Big[\cos\Theta\cos(\omega_1 t) - \sin\Theta\cos\Phi\sin(\omega_1 t)\Big] \tag{3.26}$$

A complete derivation of $< \mu_z(t) >$ and definitions of Θ and Φ can be found in Appendix A.5. In the rotating coordinate system, the expectation at $t = 0$ has a polar angle Θ and an azimuthal angle Φ about the axis of rotation, $\hat{k}$. We also see that in the rotating frame, $< \mu_x(t) >$ is time-independent. Specifically, these solutions describe the rotation of the angular momentum vector, with fixed magnitude, rotating about the $\hat{i}_{rot}$ axis, in the rotating reference frame.

3.1.6 Net Magnetization and the Bloch Equations

In this section, we introduce the Bloch equations, which characterize the trajectory of a net magnetization vector, $\vec{M}(t)$, important to the modeling of the MRI signal. In MRI studies, the measurable signal is created by manipulating $\vec{M}(t)$ with a carefully selected set of applied magnetic fields. For the time being, we assume $\vec{M}$ is spatially invariant, but we will consider its spatial dependence in subsequent sections of Chapter 3.

Consider a tiny volume V in a given object. Assume that the object is placed in a uniform magnetic field, $\vec{B} = B_0\hat{k}$, and that in V there are to be found N_p protons. Furthermore, let the n^{th} proton have magnetic moment $\vec{\mu}_n$. The net magnetization produced within V is defined as

$$\vec{M}(x,y) = \frac{1}{V}\sum_{i=1}^{N_p}\vec{\mu}_n \ , \tag{3.27}$$

where N_p is the number of protons in V (the sum is performed over all hydrogen protons in the volume). Classically, a moment vector $\vec{\mu}$ in a uniform magnetic field $\vec{B}$ would

satisfy (see [34], Chapter 2)

$$\frac{d}{dt}\vec{\mu} = \frac{d}{dt}\gamma\vec{J} = \gamma\frac{d}{dt}\vec{J} = \gamma\vec{N} = \gamma\vec{\mu}\times\vec{B}\ . \tag{3.28}$$

$\vec{J}$ is the total angular momentum vector (as in Eq. 3.4), and $\vec{N}$ is the torque exerted on the magnetic moment by the external field. Thus, combining Eqs. 3.27 and 3.28 we obtain

$$\frac{d}{dt}\vec{M} = \gamma\vec{M}\times\vec{B}\ . \tag{3.29}$$

This equation is a *simplified form of the Bloch equation*[7]. If $\vec{B} = B_0\hat{k}$, then Eq. 3.29 can be expressed as

$$\begin{pmatrix} \frac{dM_x}{dt} \\ \frac{dM_y}{dt} \\ \frac{dM_z}{dt} \end{pmatrix} = \begin{pmatrix} 0 & \gamma B_0 & 0 \\ -\gamma B_0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} M_x(t) \\ M_y(t) \\ M_z(t) \end{pmatrix}. \tag{3.30}$$

The solution to this system of differential equations is

$$\begin{pmatrix} M_x(t) \\ M_y(t) \\ M_z(t) \end{pmatrix} = \begin{pmatrix} \cos\omega_0 t & \sin\omega_0 t & 0 \\ -\sin\omega_0 t & \cos\omega_0 t & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} M_x(0) \\ M_y(0) \\ M_z(0) \end{pmatrix}. \tag{3.31}$$

The solutions to the simplified form of the Bloch equations are *similar* to those derived earlier using quantum mechanics (see Eqs. 3.18 to 3.20). Note that our quantum mechanical description contains an angle θ between the axis of rotation and the expectation of the momentum operator, that is not present in Eq. 3.31. In taking an average (Eq. 3.27) over N_p magnetic moment vectors in the given volume V, the transverse component of the net magnetization vector has vanished.

3.1.7 Boltzmann Equilibrium Values

In magnetic resonance studies, an equilibrium magnetization is generated in a magnetic field. To describe the mechanics of this process, the equilibrium magnetization M_0 for a system of particles with spin at temperature T in the presence of a static magnetic field $\vec{B} = B_0\hat{k}$ will be derived.

Consider a system consisting of N_{spins} spins that are in thermal contact with each other at a temperature T (N_{spins} is taken to be very large). The z-component of the thermal equilibrium value of M_z for a magnetic number m_s is given by ([34], Chapter 6)

$$M_0 = \rho\sum_{m_s=-s}^{s} P(\epsilon(m_s))\mu_z(m_s), \tag{3.32}$$

[7]The full Bloch equation also includes relaxation terms related to T_1 and T_2 processes, that we have ignored in Eq. 3.29.

where ρ is the density of spins per unit volume in the homogeneous isochromat volume V. $P(\epsilon)$ is the probability of finding a system with energy ϵ, while in contact with a much larger system at temperature T, and is given by

$$P(\epsilon) = \frac{e^{-\epsilon/kT}}{Z}, \quad Z = \sum_{\epsilon} e^{-\epsilon/kT}. \tag{3.33}$$

In Section 3.1 the constant k is the Boltzmann constant[8], $k = 1.381 \times 10^{-23}$ J K^{-1}.

It can be shown that the equilibrium magnetization is (see Appendix A.6 for derivation)

$$M_0 \approx \rho \frac{s(s+1)\gamma^2\hbar^2}{3kT} B_0. \tag{3.34}$$

For a proton, with spin $s = 1/2$, this expression is simply

$$M_0 \approx \rho \frac{\gamma^2\hbar^2}{4kT} B_0. \tag{3.35}$$

This expression provides a relationship between the equilibrium magnetization, M_0, and the proton density. We shall use this relationship in our discussion of the mathematical model of the MRI signal, in Section 3.2.1 (see Eq. 3.41). As we shall see, intensities in an MR *image* are proportional to the proton density of the object being imaged[9], via Eq. 3.35.

3.1.8 Proton Spin Excess

The tendency for protons in a magnetic field to align themselves with the field and thus drop to the lowest energy level is opposed by thermal interactions. These thermal motions distribute the populations in the two energy levels, and the resultant equilibrium distribution can be predicted.

Consider an $s = \frac{1}{2}$ (proton) system whose total number of spins per unit volume, N_{spins} is given by $N_{spins} = n_+ + n_-$, where n_+ and n_- represent the respective number of spins parallel and anti-parallel to the magnetic field. From the normalized Boltzmann probability (Eq. 3.33), the probability for each state is given by:

$$P_\pm = \frac{n_\pm}{N_{spins}} = \frac{e^{\pm u/2}}{e^{+u/2} + e^{-u/2}}, \qquad u = \frac{\hbar\omega_0}{kT}. \tag{3.36}$$

It follows that the ratio of the two populations is given by:

$$\frac{n_-}{n_+} = \frac{e^{-u/2}}{e^{+u/2}} = e^{-u} = e^{-\frac{\hbar\omega_0}{kT}} \tag{3.37}$$

[8]Elsewhere in this thesis, k denotes a spatial frequency variable.

[9]It is also interesting to note that intensities in the MR image are also proportional to the field strength through Eq. 3.35. Increasing B_0 by improving machine hardware increases the SNR of the data it can produce.

For $T = 300$ K and $B_0 = 1.5$ T, $u \approx 1.0214 \times 10^{-5}$. It can be easily shown that the number of parallel spins, n_+, in a volume containing one million spins (so that $N_{spins} = 10^6$), is given by:

$$n_+ = \frac{10^6}{1 + \frac{n_-}{n_+}} \approx 500002.5 \tag{3.38}$$

Thus, for the above conditions, a difference of roughly 5 spins is found between the parallel and the anti-parallel spin populations[10]. Since the spin excess of 5 spins is a million times smaller than the total number of protons, it might be guessed that no significant signal could be detected at room temperature and 1.5 T! However, in MR experiments there are often large enough proton densities in the sample being studied to permit a measurable signal.

3.2 MRI Data Acquisiton

Implementation of image reconstruction algorithms in MRI requires knowledge of the data acquisition process. Many of the processing steps and physical constraints on the MRI hardware not only affect the performance of reconstruction algorithms, they provide information that has been used to enhance image quality when incorporated into reconstruction algorithms [53, 55]. A description of the most pertinent concepts of MRI data acquisition is therefore a necessity for later discussion.

The equations that model the measured MRI signal and the basic operations that are used to process it are important relations for Chapter 5. The MR signal equation is carefully derived in many references, including Liang and Lauterbur's book [55], a well-known article by Hinshaw and Lent [40], and the book by Haacke, *et al.* [34], which devotes several chapters to its origins. The signal equation provides a means of modeling the measured data using physical parameters, many of which are known *a priori*, or have various constraints imposed upon them. Later chapters of this thesis investigate new means of using these constraints to improve the reconstructed image.

3.2.1 The Signal Equation

Let the measured signal produced by the set of applied fields be represented by $s_m(t) \in L^2(\mathbb{R})$, $t \geq 0$. This signal may be modeled as[11] (see [34], page 101):

$$s_m(t) \propto \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} M_\perp(x, y, z, 0) \sin(\phi(x, y, z, t)) dx dy dz. \tag{3.39}$$

[10] This simple example is based on an example exercise on page 87 of [34].

[11] Here we have ignored T_2 relaxation effects, noise, and inhomogeneities in the receive field.

The function $M_\perp(x,y,z,t)$ is the magnitude of the transverse component of the net magnetization vector, $\vec{M}(x,y,z,t) = [M_x, M_y, M_z]$, so that: $M_\perp(x,y,z,t) = \sqrt{M_x^2 + M_y^2}$. $\vec{M}(x,y,z,t)$ is produced by, and interacts with, the applied magnetic fields used to create and spatially encode an MR signal.

The phase $\phi(x,y,z,t)$ in Eq. 3.39 may be modeled as a sum of two terms:

$$\begin{aligned}\phi(x,y,z,t) &= \omega_0 t + \phi_G(x,y,z,t) \\ &= \gamma B_0 t + \phi_G(x,y,z,t).\end{aligned} \tag{3.40}$$

The constant γ is the gyromagnetic ratio. B_0 is the magnitude of the applied magnetic field. The function $\phi_G(x,y,z,t)$ is the local phase at time t of the proton spin that is determined by a set of applied magnetic gradient fields.

A derivation of the signal equation is well beyond the scope of this thesis. A thorough derivation of this equation from fundamental quantum and electromagnetic concepts can be found in the literature (for example, see [34], Chapter 7).

It is useful to draw a connection between $M_\perp(x,y,z,0)$ and the proton density $\rho(x,y,z) \in L^2(\mathbb{R})$, of the object being imaged. The function, $M_\perp(x,y,z,0)$, is approximately linearly proportional to $\rho(x,y,z)$ (see Eq. 3.34, or [34] page 86). The proton density, $\rho(x,y,z)$, satisfies the following constraints:

$$\begin{aligned}&1.\ \rho(x,y,z) \text{ is real } \forall\ (x,y,z), \\ &2.\ \rho(x,y,z) \geq 0\ \forall\ (x,y,z), \\ &3.\ \rho(x,y,z) = \Pi(\frac{x}{R_x})\Pi(\frac{y}{R_y})\Pi(\frac{z}{R_z})\rho(x,y,z) \text{ for some } R_x, R_y, \text{ and } R_z.\end{aligned} \tag{3.41}$$

In later chapters, we will use ρ in place of $M_\perp$. The above constrains are used, for example, in SRMRI algorithms based on projection techniques for frequency extrapolation, and will be discussed in Section 4.1. Namely, the first and second contraints are used in the phase constrained algorithms, while the third is used in the Papoulis-Gerchberg algorithm.

3.2.2 Demodulation

After the continuous signal is measured, a number of processing steps must be carried out to create a signal that can be stored on a computer. The precise set of processing steps used varies between MRI scanners, but the standard processing steps, as described in standard MRI texts [34, 40, 55], that are most relevant to subsequent chapters of this thesis will now be outlined. Namely, we discuss in this section mathematical models for *demodulation* (which includes *analog filtering*) and *sampling*.

The measured data must be *demodulated*[12]. The demodulation process is also referred to as phase-sensitive detection (PSD) [55], and is outlined in Fig. 3.2. Although the raw signal, $s_m(t)$ is real, this function is converted into a complex entity through the process of demodulation. Furthermore, high frequency oscillations present in the signal that stem from the $\omega_0 t$ term in Eq. 3.40, must be removed. Demodulation simultaneously creates a complex signal and removes this high frequency component.

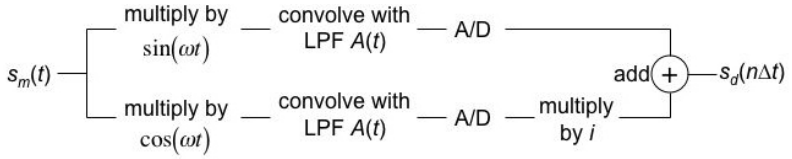


Figure 3.2: Elementary modeling of the MR signal. The data is split into two channels, multiplied by either a cosine or sine function, passed through a low pass filter, and then converted to a discrete signal.

PSD is carried out in a series of steps, demonstrated in Fig. 3.2. The signal is split into two channels, each channel is multiplied by either a sine or cosine function, and finally, a low-pass filter is applied to each channel. The channels are summed together, forming a complex demodulated signal, $s_d(t)$ (for derivation see Appendix A.7):

$$s_d(t) \approx \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} M_\perp(x, y, z, 0) e^{i\phi_G(x,y,z,t)} dx dy dz. \tag{3.42}$$

Comparing this equation with Eq. 3.39, we see that the measured signal, s_d, is now a complex function, and the high frequency term has been removed.

3.2.3 Slice Encoding

This section briefly discusses the mechanics of slice encoding (SE). How SE relates to SR methods will be discussed in Section 4.2.5. Research on SRMRI presented in this thesis focuses on resolution enhancement in the frequency and phase encoding directions, and does not extend to SE methods. Nonetheless, SRMRI in the SE direction (currently) remains a promising direction for research for MR resolution enhancement with multiple MR images. In fact, we argue in Chapter 5 that for research dealing with multiple acquisitions, spatial resolution enhancement in the SE direction currently appears to be *the most promising direction for future research in SRMRI*[13].

[12]The process of demodulation is actually carried out *after* the slice and phase encoding processes are carried out. We place this discussion of demodulation prior to our discussion of these encoding schemes to simplify our derivations. Namely, demodulation removes the $e^{i\omega_0 t}$ term from our model of the measured signal.

[13]Recent work [75] has introduced SRMRI with multiple RF coils. We briefly discuss this possible research direction in Chapter 10.

SE is fundamentally different from other encoding schemes. While frequency (Section 3.2.4) and phase (Section 3.2.5) encoding methods collect information in the Fourier domain, SE is a *spatial* encoding scheme. SE isolates information along the z-axis according to a prescribed distribution, determined by a chosen RF pulse and magnetic field gradient. Also, because the acquisition procedures related to SE are applied before those associated with frequency and phase encoding, our discussion of SE in this thesis is placed before our discussion of these other encoding methods.

For a typical MR experiment, the measured signal at a given (x, y, z) coordinate can be found through solving a system of differential equations derived from the Bloch equations. The process of SE is carried out with the application of two magnetic fields: an applied linear (in z) gradient field, and an RF field that produces signal excitation. The gradient field, $\vec{G}(z,t) = \gamma G_z(t) z\hat{k}$, exhibits a *linear* dependence between the gradient magnitude and z. This field enables the required spatial localization needed to isolate information from a desired region of the object being imaged. The RF field, $\vec{B}_1(t)$, excites the object, so that a signal can be measured. The shape of this function also determines the slice profile, or the distribution of spins along the z-axis that will contribute to the signal.

The Bloch Equation in the Rotating Frame

This subsection discusses the mathematical model for slice encoding. It is convenient to convert the coordinate system into one rotating about the z-axis with frequency ω_{rf}. To this end, we relate the *laboratory frame* coordinates to the *rotating frame* coordinates, by choosing rotating unit vectors, $\hat{i}_{rot}$, $\hat{j}_{rot}$, and $\hat{k}_{rot}$, such that:

$$\begin{aligned}
\hat{i}_{rot} &= \cos(\omega_{rf}t)\hat{i} - \sin(\omega_{rf}t)\hat{j} \\
\hat{j}_{rot} &= \sin(\omega_{rf}t)\hat{i} + \cos(\omega_{rf}t)\hat{j} \\
\hat{k}_{rot} &= \hat{k}.
\end{aligned}$$

Now let us define vectors $\vec{M}_{rot} = [M_{x,rot}, M_{y,rot}, M_{z,rot}]$ and $\vec{B}_{rot} = [B_{x,rot}, B_{y,rot}, B_{z,rot}]$, which represent vectors $\vec{M}$ and $\vec{B}$ in the rotating frame:

$$\begin{aligned}
\vec{M}(x,y,z,t) &= M_x\hat{i} + M_y\hat{j} + M_z\hat{k} \\
&= M_{x,rot}\hat{i}_{rot} + M_{y,rot}\hat{j}_{rot} + M_{z,rot}\hat{k}_{rot} \\
&= \vec{M}_{rot},
\end{aligned} \tag{3.43}$$

and

$$\begin{aligned}
\vec{B}(x,y,z,t) &= B_x\hat{i} + B_y\hat{j} + B_z\hat{k} \\
&= B_{x,rot}\hat{i}_{rot} + B_{y,rot}\hat{j}_{rot} + B_{z,rot}\hat{k}_{rot} \\
&= \vec{B}_{rot}.
\end{aligned} \tag{3.44}$$

In the rotating frame, for a given (x, y) coordinate, the magnetization vector $\vec{M}_{rot}$, for any time, t, and z position is found by solving the following system of equations (derived from the Bloch equations in Appendix A.8):

$$\frac{\partial \vec{M}_{rot}(z,t)}{\partial t} = \begin{pmatrix} 0 & \gamma G_z(t) z & 0 \\ -\gamma G_z(t) z & 0 & \omega_1(t) \\ 0 & -\omega_1(t) & 0 \end{pmatrix} \begin{pmatrix} M_{x,rot}(z,t) \\ M_{y,rot}(z,t) \\ M_{z,rot}(z,t) \end{pmatrix} \tag{3.45}$$

The function $\omega_1(t)$ represents the oscillation frequency of the RF field, and $\omega_1(t) = \gamma B_1(t)$. To arrive at Eq. 3.45, a number of assumptions were made in Appendix A.8:

1. The general form of the applied RF field is:

$$\vec{B}_1(t) = \frac{\omega_1(t)}{\gamma}[\cos(\omega_{rf} t)\hat{i} - \sin(\omega_{rf} t)\hat{j}] = B_1(t)\hat{i}_{rot}$$

   The acquisition parameter ω_{rf} is a constant and is set by the MR scanner operator. This parameter is the *carrier frequency* of the RF field.

2. The carrier frequency, ω_{rf}, is assumed to be set (or *tuned*) to the natural frequency of the object being excited, so that $\omega_{rf} = \omega_0$.

3. We are further assuming that the main magnetic field, $\vec{B}_0$ is spatially invariant. In practice, there are small spatial variations[14] in $\vec{B}_0$.

4. Relaxation terms (due to T_1 and T_2 relaxation) were ignored, meaning that the B_1 field is assumed to be a pulse of short duration compared to T_1 and T_2. This is an accurate assumption in most situations.

The solution to Eq. 3.45 yields the distribution of spins along the z-direction that will contribute to the measured signal. However, this system does not, in general, permit an analytical solution. Approximate solutions can be obtained, for example, with numerical techniques [56], perturbation analysis [41], or Fourier methods [78].

The Signal Equation in Two Spatial Dimensions

This subsection describes two simplifications to Eq. 3.42. Firstly, the desired effect of SE is to create a spatial localization in z. Suppose that we seek to have spatial information localized about a z coordinate, z_0, with slice width Δz. Assuming that information about $M_\perp(x, y, z, 0)$ is acquired uniformly across the slice, and that no information outside of

[14] See [34], Chapters 20 and 27, for discussion of magnetic field inhomogeneities and main magnet coil design.

the slice is recorded, we may simplify Eq. 3.42 so that it no longer is a function of z, by introducing the following function[15]:

$$M_{z_0}(x,y) = \int_{-\infty}^{\infty} \Pi(\frac{z-z_0}{\Delta z}) M_\perp(x,y,z,0)dz = \int_{z_0-\Delta z/2}^{z_0+\Delta z/2} M_\perp(x,y,z,0)dz. \tag{3.46}$$

Furthermore, as spatial information in the z direction has now been encoded, it is no longer necessary to use gradients along this direction to encode spatial information. As such, ϕ_G is no longer a function of z, and $\phi_G = \phi_G(x,y,t)$. We will maintain a dependence in ϕ_G on x and y, as we still must implement spatial encoding in the x and y directions. The following sections of this chapter discuss spatial encoding in these other directions.

Using these simplifications, Eq. 3.42 becomes

$$\begin{aligned} s_d(t) &= \iiint M_\perp(x,y,z,0)\mathrm{e}^{i\phi_G(x,y,z,t)}dxdydz \\ &= \iint M_{z_0}(x,y)\mathrm{e}^{i\phi_G(x,y,t)}dxdy. \end{aligned} \tag{3.47}$$

This equation will be modified to represent a two-dimensional Fourier Transform of $M_{z_0}(x,y)$ in the following sections of this chapter.

3.2.4 Frequency Encoding

Discussions in previous sections of this chapter described basic physics required to acquire an MR signal and how to isolate information from a slice. After a slice has been selected, the next step in the acquisition process is to create a linear relationship between frequency and position in the object being scanned. This is the process of frequency encoding (FE).

Eq. 3.47 provides an expression for the demodulated signal, $s_d(t)$, for two spatial coordinates. We now require an explicit expression for the phase, $\phi(x,y,t)$, in terms of the magnetic field gradients. The function $\phi_G(x,y,t)$ represents the local phase due to applied magnetic gradient fields. Linear gradients are often used, so that the time-varying magnitude of the fields changes linearly in space

$$B_G(x,y,t) = G_x(t)x + G_y(t)y = \vec{r}\cdot\vec{G}(t). \tag{3.48}$$

The spatially invariant magnetic field vector $\vec{G}(t)$ is the magnetic field gradient

$$\vec{G}(t) = [\frac{\partial B_x(t)}{\partial x}, \frac{\partial B_y(t)}{\partial y}]. \tag{3.49}$$

[15]Some information from outside the slice is always measured, and information across the slice is not uniformly weighted. More thorough discussions of slice encoding can be found in the literature (see, for example, [34], Chapter 16).

With this magnetic field gradient, the signal Eq. 3.47 becomes (see Appendix A.10 for derivation)

$$s_d(t) = s_d(k_x, k_y) = \iint M_{z_0}(x,y)\exp[-i2\pi\vec{k}(t)\cdot\vec{r}]dxdy, \tag{3.50}$$

where

$$\begin{aligned}\vec{k}(t) &= \big[k_x(t), k_y(t)\big],\\ k_x(t) &= \frac{\gamma}{2\pi}\int_0^t G_x(t')dt',\\ k_y(t) &= \frac{\gamma}{2\pi}\int_0^t G_y(t')dt'.\end{aligned} \tag{3.51}$$

The temporal functions in Eq. 3.51 represent a change of variables that introduces the concept of k-space. Eq. 3.50 demonstrates that the demodulated signal is the two-dimensional Fourier Transform of $M_{z_0}(x,y)$. As noted in Chapter 2.1, k_x and k_y are both expressed in units of 1/cm.

The MRI raw k-space data is often thought of as a 2D Fourier Transform of a proton density function. A combination of Eqs. 3.50 and 3.35 yields[16]

$$s_d(k_x, k_y) = \iint \rho(x,y)e^{-i2\pi(k_x x + k_y y)}dxdy. \tag{3.52}$$

This relation is used in later sections of this thesis (see for example Eqs. 7.1, 8.2).

Eq. 3.51 suggests that the data is only measured along *trajectories* in frequency space. Indeed, careful selection of these trajectories is an integral part of MRI data acquisition, and the gradient amplitudes (which are functions of time) determine the paths in frequency space along which the data is measured. These trajectories always begin at the origin of k-space, which can be verified by setting $t = 0$ in Eq. 3.50. The direction along which these paths run is referred to as the *frequency encode*, or *readout* direction.

In theory, these trajectories may take any desired path that starts at the k-space origin. From Eq. 3.52, the G_x and G_y gradients will determine these trajectories, and are often modeled as piecewise constant, or piecewise linear functions. Consider the sequence diagram in Fig. 3.3a. As the G_x and G_y gradients are both turned on at time t_1 for a time T, then reversed in polarity and left on for a time $2T$, the gradients produce the linear trajectory, shown in Fig. 3.3b. If this sequence were repeated many times, each time with different G_x and G_y magnitudes, the k-space trajectories would describe a set of radial lines emanating from the origin of k-space, one from each acquisition. Each line starts at the origin and moves out at an angle θ where $\theta = \tan^{-1}\left(G_y/G_x\right)$.

[16]To arrive at this equation, we ignored constants found in Eq. 3.35, and assumed that the object being imaged has two spatial dimensions.

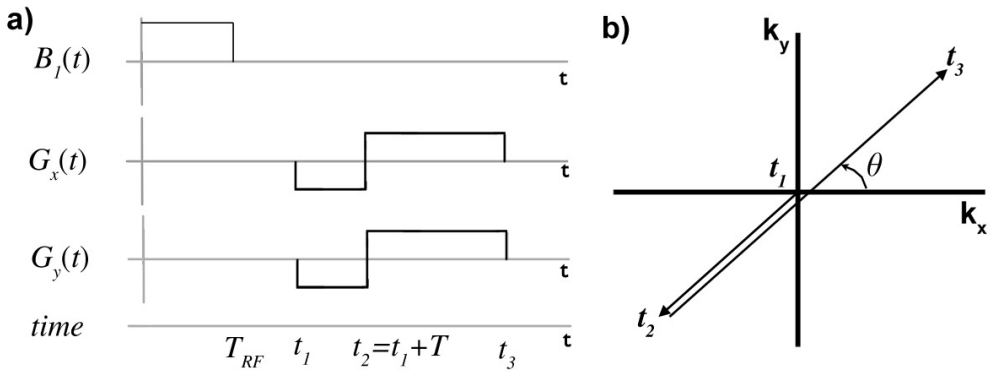


Figure 3.3: Fig. a) displays a timing diagram, showing an elementary pulse sequence to model the MRI data acquisition process. The $G_x(t)$ and $G_y(t)$ gradients are used to navigate through k-space, and the corresponding path is shown in b). The particular choice of gradients leads to a linear trajectory making an angle θ from the k_x-axis.

3.2.5 Phase Encoding

Frequency space is often measured inside a rectangular domain along a set of lines that run parallel to one of the coordinate axes[17]. This data acquisition model is described in the timing diagram in Fig. 3.4a. A gradient $G_y(t)$ is applied in the y direction for a prescribed time, T_{PE}, before recording the signal. During the time that the y gradient is being applied, the spins experience a field strength that increases linearly along y. This creates a linearly varying phase offset that is a function of y, thereby encoding information in this coordinate direction.

While the G_x gradient is turned on (from times t_2 to t_3), the MR signal may be "read out", or measured, and as such, $G_x(t)$ is often referred to as the *readout gradient*. The continuous signal must be sampled while $G_x(t)$ is being recorded (see Section 3.2.6). Successive repetitions of the scheme in Fig. 3.4 may be made by varying the magnitude of G_y to collect samples on a Cartesian lattice centered about the origin of k-space. The end result is a 2D discrete approximation to the Fourier transform of $\rho(x, y)$.

To be more precise, in the laboratory frame, spins oscillating with frequency

$$\omega(y) = \omega_0 + \gamma y G_y \ ,$$

after time interval T_{PE}, have accumulated a phase, $\phi(y)$ of:

$$\phi(y) = \omega_0 T_{PE} + \gamma G_y y T_{PE}. \tag{3.53}$$

[17]Other well-known approaches include spiral and radial k-space trajectories, which may be suited for particular applications. Different acquisition schemes are often preferred over one another to reduce relaxation artifacts, optimize the signal-to-noise ratio (SNR), scan time, or resolution for a given application.

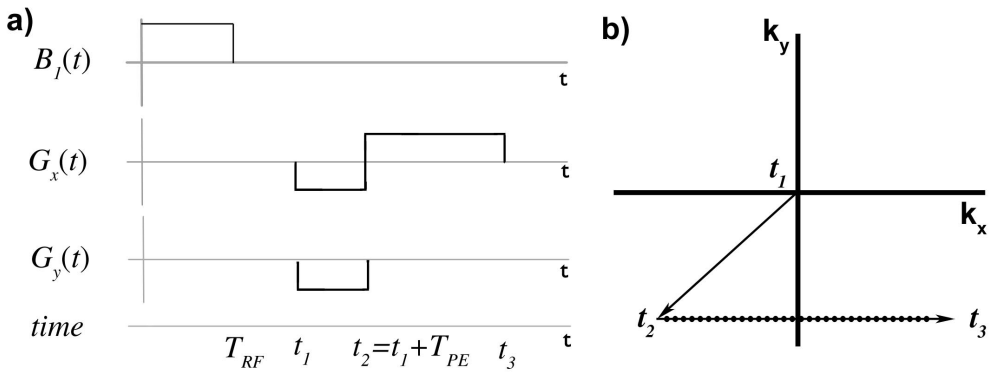


Figure 3.4: Fig. a) displays a timing diagram, showing an elementary pulse sequence to model the MRI data acquisition process. The $G_x(t)$ and $G_y(t)$ gradients are used to navigate through k-space, and the corresponding path is shown in b). The gradients in this example create a trajectory that travels along a line parallel to the k_x axis between times t_2 and t_3.

In the rotating frame, we may ignore the ω_0 term, so that the phase is simply $\gamma G_y y T_{PE}$. Essentially, spins are spatially *phase-encoded* before the signal is recorded. The amount of phase-encoding is dependent on the amplitude G_y and on the time interval T_{PE}.

An additional perspective of phase encoding provides further insight into how spatial information is encoded into the measured signal. We may rewrite the signal Eq. 3.47 as follows:

$$\begin{aligned} s_d(t) &= \iint M_{z_0}(x,y)\exp[-i2\pi(k_x(t)x + k_y(t)y)]dxdy \\ &= \int \left\{ \int M_{z_0}(x,y)e^{-i2\pi k_y y}dy \right\} e^{-i2\pi k_x x}dx \\ &= \int \left\{ \int M_{z_0}(x,y)e^{-i2\pi y G_y T_{PE}}dy \right\} e^{-i2\pi k_x x}dx \\ &= \int \left[\mathbf{P}_0\{ M_{z_0}(x,y)e^{-i2\pi y G_y T_{PE}} \} \right] e^{-i2\pi k_x x}dx. \end{aligned} \tag{3.54}$$

The function in the square brackets represents a projection[18] $\mathbf{P}$ of $M_{z_0}(x,y)$ multiplied by a complex weighting factor. Because this phase weighting factor warps the phase of the spins prior to the signal readout, this method of acquiring the data is sometimes referred to as *spin-warp imaging*.

Each k_y line measurement is based on a different yG_yT_{PE} product, which produces the phase encoding[19] of the data. Thus, because of the role that the y gradient plays, it is commonly referred to as the *phase-encoding* gradient. The process of phase encoding is essentially equivalent to that of frequency encoding. The only minor difference is that

[18]The specific projection operator we are using was introduced in Section 2.3.

[19]Typically, T_{PE} is fixed and G_y is varied.

the signal is often being recorded while the G_x gradient is on, whereas the G_y gradient is only briefly applied between successive applications of the readout gradient.

It should be emphasized, however, that although spin-warp imaging is the most prevalent strategy for acquiring data in MRI, it is only one of many different methods that may be considered. Indeed, many other acquisition strategies generated via particular combinations of $G_x(t)$ and $G_y(t)$ gradients may be implemented. However, throughout the remainder of this thesis, unless otherwise indicated, we assume that a spin-warp imaging sequence is employed.

3.2.6 Sampling the MR Signal

The demodulated signal is however still continuous and must be sampled. From Section 3.2.4, the measured, demodulated signal in raw analog form is (Eq. 3.50)

$$s_d(t) = \iint M_{z_0}(x,y)\mathrm{e}^{-i2\pi[k_x(t)x+k_y(t)y]}dxdy.$$

Recall that although data is being acquired in the time domain, the measured data has a specific two-dimensional k-space interpretation, and is acquired along trajectories that lie parallel to the k_x-axis. Along these lines, k_y is equal to the constant $G_{0,y}T_{PE}$

$$\begin{aligned} s_d(t) &= s_d(k_x(t)) \\ &= \iint M_{z_0}(x,y)\mathrm{e}^{-i2\pi[k_x(t)x+G_{0,y}T_{PE}y]}dxdy. \end{aligned} \tag{3.55}$$

Prior to sampling this data, a second low-pass filter, $A(t)$, is applied to reduce aliasing artifacts. The anti-aliasing filter is modeled by a convolution in the *frequency* (or time) domain:

$$s_d(k_x(t)) * A(k_x(t)) = \int_{-\delta k}^{\delta k} s_d(k_x-\kappa)A(BW_A k_x)d\kappa. \tag{3.56}$$

From the convolution theorem [8], it is possible to model this operation as a *multiplication* in the spatial domain. Conceptually, the filter attempts to null out the signal in the spatial domain that comes from outside the field-of-view. The width of the field-of-view will be determined by the chosen sample spacing in the frequency (or time) domain.

Following the application of the anti-aliasing filter, the data are discretized. Sampling the filtered signal yields $\hat{s}_d(k_{x,m})$:

$$\begin{aligned} \hat{s}_d(k_{x,m}) &= \mathrm{III}(\frac{k_x}{\Delta k_x})\big[s_d(k_x(t)) * A(k_x(t))\big] \\ &= \mathrm{III}(\frac{k_x}{\Delta k_x})\big[\int_{-\delta k}^{\delta k} s_d(k_x-\kappa)A(BW_A k_x)d\kappa\big] \\ &= \sum_{m=-M/2}^{M/2}\int_{-\delta k}^{\delta k} s_d(k_{x,m}-\kappa)A(BW_A k_{x,m})d\kappa, \end{aligned} \tag{3.57}$$

where $k_{x,m} = m\Delta k_x$. Sampling is often modeled with the use of the comb function in this manner. In practice, however, the measurements $\hat{s}_d(t)$ represent local *averages* of $s_d(t)$. For some MRI applications, the duration over which the analog signal is averaged (typically on the order of a few microseconds) is usually regarded as negligible, and Eq. 3.57 is used for mathematical models of the MR signal.

Data is naturally discretized in the z and y directions. Unlike the FE direction, *it is not possible to apply an anti-aliasing filter in the PE and SE directions.* In Chapter 5, the nature of the data acquisition process, and in particular, the anti-aliasing filter, will play an important role in the efficacy of SRMRI algorithms.

Chapter 4

MR Image Reconstruction

There are many methods presented in the MR literature that obtain the spatial domain image from raw k-space data. Although the most frequently used method is the 2D inverse discrete Fourier Transform (IDFT), many other methods have been proposed, many of which attempt to implement SR. This chapter presents a review of *single* acquisition methods for SRMRI (Section 4.1). Section 4.2 describes material related to resolution enhancement with *multiple* acquisitions. The chapter as a whole is a literature review of material needed for results presented in later chapters of this thesis.

4.1 Constrained Reconstruction MRI with One Image

Surpassing the limitations posed by the SNR equation (Eq. 1.2) was a highly active area of research in the 1980's and early 1990's, summarized by a comprehensive review article [53], and in Chapter 10 of Liang and Lauterbur's book [55]. Limitations posed by Eq. 1.2 are due in part to the reconstruction algorithm most commonly used: the IDFT. Constrained reconstruction algorithms offer new alternatives by incorporating additional *a priori* knowledge in the reconstruction process. It has been found that these algorithms can improve the quality of the image over the IDFT alone.

Constrained reconstruction algorithms often attempt to directly extrapolate the discrete measured data. If a sufficiently stable and accurate extrapolation technique can be employed, higher frequencies than those that are measured can be estimated, thereby improving the spatial resolution. Extrapolation algorithms that utilize *a priori* knowledge of the object being imaged are the focus of this section.

4.1.1 Extrapolation via Prolate Spheroidal Wave Functions

Extrapolation using prolate spheroidal wave functions (PSWF's) [21] was originally reported in a classic signal processing paper by Slepian and Pollak [91]. This technique

provides the unexpected result that under sufficient conditions, a *unique* solution to the spatially limited continuous extrapolation problem can be found.

To perform extrapolation with PSWF's, one evaluates

$$s(k) = \sum_{n=0}^{\infty} a_n \psi_n(k), \ \ \forall \ k \in \mathbb{R}, \tag{4.1}$$

where the complex coefficients, a_n, are found with the following equation:

$$a_n = \frac{1}{\lambda_n} \int_{-k_{max}}^{+k_{max}} \psi_m(k) s(k) dk. \tag{4.2}$$

The known functions $\psi_n(k)$ are the PSWF's, and $\lambda_n \in \mathbb{R}$ are their eigenvalues. For the interested reader, further details on PSWF's, and how they may be used to perform extrapolation is provided in Appendix B.1.

Although PSWF's provide some insight into the extrapolation problem, and are useful from an analysis perspective, they are *too sensitive to noise to be directly used for extrapolation.*

4.1.2 The Papoulis-Gerchberg Algorithm

The Papoulis-Gerchberg (PG) algorithm[1] is another technique that has been used for extrapolation problems [27, 77]. This technique achieves extrapolation via an iterative procedure which *converges to the unique solution* of the continuous extrapolation problem[2]!

Several authors later adapted the algorithm to address discrete extrapolation problems [10, 42, 87, 88, 89, 93]. The discrete PGA provides further insight into the extrapolation problem, but like extrapolation with PSWF's, this approach is *also too sensitive to noise to be directly used for extrapolation.* Our discussion here is included to help provide a basis for discussion of current iterative projection algorithms.

The Continuous PGA

Assume that a continuous frequency domain signal with finite spatial extent[3], $s(k) \in \mathcal{R}_R$. The signal $s(k)$ is only measured over the interval $[-k_{max}, +k_{max}]$ to produce $s_m(k)$:

$$s_m(k) = \Pi(\frac{k}{2k_{max}}) s(k). \tag{4.3}$$

[1]The PG algorithm was first discovered by Papoulis, but first published, independently, by Gerchberg (see [61], page 260). Some authors refer to this technique as the Gerchberg-Papoulis algorithm.

[2]See Section 1.2 for a description of the continuous extrapolation problem, and reference [77] for a proof of this convergence.

[3]Recall that $\mathcal{R}_R$, defined in Eq. 2.7, is the space of *frequency domain* functions that have a finite spatial extent of R.

Given s_m and R (which yields the continuous extrapolation problem, Section 1.2), the continuous PG algorithm may be used to calculate $s(k)$ for all k, using the following iteration, for $n \geq 0$:

$$s_{e,n+1}(k) = s_m(k) + \left[1 + \Pi(\frac{k}{2k_{max}})\right] R\text{sinc}(Rk) * s_{e,n}(k). \tag{4.4}$$

The function $s_{e,n}(k)$ is the extrapolation of $s_m(k)$ at iteration n. Typically, the initial condition $s_{e,0} = s_m(k)$ is used to calculate $s_{e,1}$. The first two iterations in this sequence are shown schematically in Fig. 4.1.

Using PSWF's it has been shown that [77]:

$$\int_{-\infty}^{+\infty} |s_{e,n}(k) - s(k)|^2 dk \to 0, \text{ as } n \to \infty. \tag{4.5}$$

The sequence therefore converges in L^2 to the unique solution, $s(k)$. Although this may seem impossible, an insightful diagram (Fig. 4.2), originally presented by Youla [99], illustrates the convergence of the PG algorithm. Like the extrapolation approach using PSWFs, the PG algorithm converges to $s(k)$ for all k with knowledge of only $s(k), k \in [-k_{max}, +k_{max}]$, and the spatial extent, R, of $s(k)$.

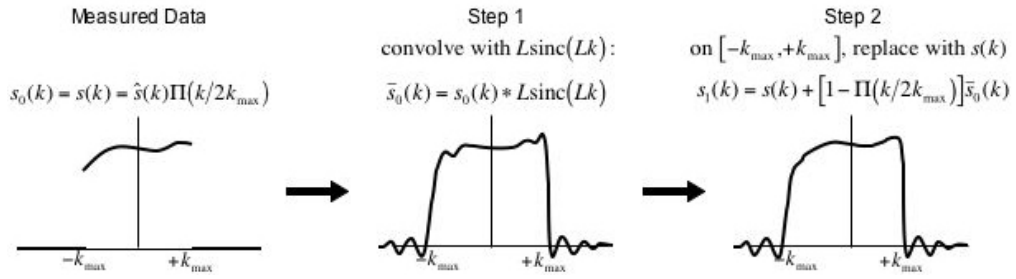


Figure 4.1: The continuous PG algorithm. Initially, the measured data, $s(k)$, is set equal to $s_0(k)$. To carry out step 1, s_0 is convolved with a sinc function with a spatial extent of L, to produce $\bar{s}_0(k)$. Step 2 consists of replacing $\bar{s}_0(k)$ with $s(k)$ over $[-k_{max}, +k_{max}]$. Steps 1 and 2 are repeated iteratively until a chosen convergence criterion is met.

The Discrete PGA

Practical implementation of the PG algorithm necessitates the use of discrete sequences with a finite number of samples. It has been shown that the discrete PG algorithm (DPGA) may not converge to a unique solution [42].

The literature has many slightly different formulas that are based on Eq. 4.4 for performing the DPGA. Approximating the convolution integral in Eq. 4.4 with a discrete sum with a finite number of terms can be problematic, as truncation errors must be

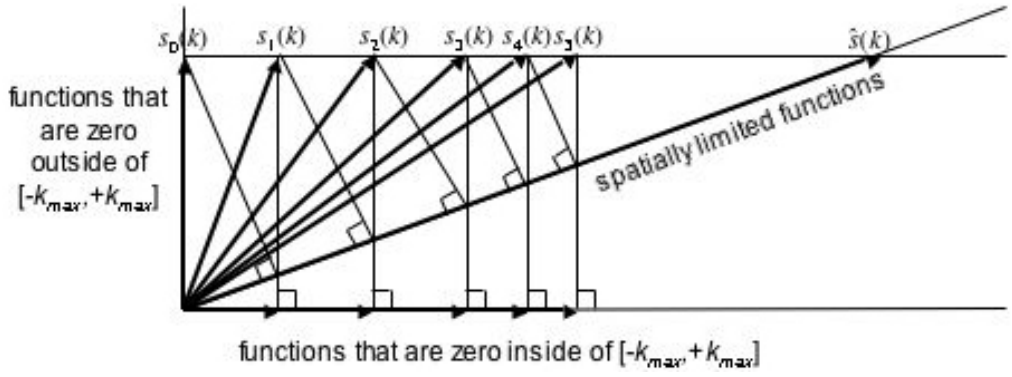


Figure 4.2: Youla's illustration [99] of the continuous PG algorithm's convergence. The algorithm consists of iteratively summing projections onto two spaces: the space of spatially limited functions, and the space of functions that are zero over $[-k_{max}, +k_{max}]$.

accepted in order to handle vectors with a *finite* number of elements. Alternatively, several improved iterative approaches have been proposed that circumvent this problem [10, 42, 87, 88, 89, 93].

A description of several experiments carried out using the DPGA on simulated data are provided in Appendix B.2 for the interested reader.

The original Papoulis-Gerchberg algorithm [27, 77] suffers from a slow convergence rate. Motivated by this problem, a non-iterative technique was explored for discrete signal extrapolation [87], and was later considered by several other authors [10, 42, 88, 89]. These approaches are often based on transforming the extrapolation problem into solving an integral equation that may be very difficult to solve. In fact, these non-iterative techniques are not generally recommended [53], as they often involve inverting a matrix, the extrapolation matrix [87], which is usually ill-conditioned. Typically, regularization methods are employed to provide a stable algorithm.

However, one interesting observation was obtained from this non-iterative approach. Paradoxically, in the discrete case, as the sample spacing Δk is decreased, the more ill-conditioned the extrapolation matrix becomes [10, 42, 53, 88, 89]. While the stability of many other extrapolation methods as a function of Δk has not been reported, this subject may be of interest when investigating other extrapolation techniques, especially when considering *multiple*-acquisition super-resolution.

4.1.3 Linear Prediction and Autoregressive Moving Average Modeling

The linear predictive (LP) extrapolation is based on the idea that each sample of a noiseless discrete complex trigonometric polynomial in $\mathcal{R}_R$ can be expressed as a linear

combination of its past samples. In order to use this property for data extrapolation in MRI, we must first introduce a model of the measured data. Several approaches to modeling MR data based on a priori knowledge of the acquisition process have been proposed to this end. In the *boxcar model* [35, 54], $\rho(x)$ is represented as a sum of N_C indicator functions:

$$\rho_{N_C}(x) = \sum_{m=1}^{N_C} c_m \Pi\Big(\frac{x - x_m}{\Delta x}\Big). \tag{4.6}$$

For simplicity, each of the indicator functions has width Δx (we may also consider indicator functions whose widths vary with m). Performing a Fourier Transform of this signal yields

$$\mathsf{P}_{N_C}(k) = \frac{1}{\pi k} \sum_{m=1}^{N_C} c_m \sin(\pi \Delta x k) e^{-2\pi i k x_m}. \tag{4.7}$$

Multiplication by $-2\pi ik$ and sampling converts $\mathsf{P}_{N_C}(k)$ into a complex trigonometric polynomial[4]:

$$\hat{s}_{N_C}(n) = -2\pi ik \mathsf{P}_{N_C}(n) = \sum_{m=1}^{2N_C} d_m e^{-2\pi i n \xi_m}, \ n = 0, 1, 2, \ldots, N_S - 1. \tag{4.8}$$

where the number of samples is N_S, and

$$\xi_m = \begin{cases} x_m - \Delta x/2, & m = 1, 2, \ldots, N_C \\ x_m + \Delta x/2, & m = N_C + 1, N_C + 2, \ldots, 2N_C \end{cases}$$

$$d_m = \begin{cases} c_m, & m = 1, 2, \ldots, N_C \\ c_{m-N_C}, & m = N_C + 1, N_C + 2, \ldots, 2N_C \end{cases}$$

It has been shown that discrete signals in the form of Eq. 4.8 obey [83, 97]:

$$\hat{s}_{N_C}(n) = -\sum_{m=1}^{2N_C} a_m \hat{s}_{N_C}(n - m), \quad n = 2N_C, 2N_C + 1, \ldots, N_s - 1. \tag{4.9}$$

This result is a consequence of Prony's method [97], and the coefficients $a_m \in \mathbb{C}$ are known as the *autoregressive* (AR) coefficients [58]. This model of the MR data provides us with relations for self-similar parameters discussed in Chapter 8.

Eq. 4.9 provides the basis for an LP approach for 1D extrapolation of k-space MR data, and is a special case of the autoregressive moving average (ARMA) technique. ARMA is a more general approach that has been extensively studied in MRI for resolution enhancement and artifact reduction (see [53, 55, 68, 92]).

[4] We see that this signal is now in the form of Eq. 1.9, and will show in Chapter 8 that signals of this form possess a type of self-similarity.

4.1.4 Conclusions

Constrained reconstruction in MRI has received much attention since the first papers started appearing in 1986. Many other reconstruction techniques that were not mentioned here have been proposed in the MRI literature. Namely, the generalized series technique and maximum entropy reconstructions have been found to yield promising results. One of the techniques that has been extensively studied by many authors is the phase constrained projection onto convex sets (POCS) technique. Descriptions of phase constrained techniques are provided in appendicies B.3 and B.4. Comparisons between these techniques and others can be found in the 1992 constrained reconstruction review article [53], as well as Liang and Lauterbur's book [55].

4.2 Super-Resolution MRI with Multiple Acquisitions

This section provides a review of SRMRI techniques published in the literature. Research on SRMRI with multiple acquisitions for medical imaging was recently reviewed in *The Computer Journal* [30].

4.2.1 Mathematical Modeling of MASRMRI

To describe the basic multiple acquisition (MA) SRMRI algorithm, consider the acquisition of two low-resolution, one-dimensional data sets, $l_1(2n_l\Delta x)$ and $l_2(2n_l\Delta x + \Delta x)$, $n_l = 0, 1, 2, \cdots, N_l - 1$. These two signals, or channels, which have a uniform sample spacing of $2\Delta x$, are vectors in $\mathbb{R}^{N_l}$. The channels l_1 and l_2 represent *spatial* domain data.

The discrete merged signal, $g(n\Delta x) \in \mathbb{R}^{N_h}$, $n = 0, 1, 2, \cdots, N_h - 1$, where $N_h = 2N_l$, is derived from the low-resolution channels as follows,

$$g(n\Delta x) = \begin{cases} l_1(n\Delta x) & n \text{ even} \\ l_2(n\Delta x) & n \text{ odd} \end{cases} . \tag{4.10}$$

The samples of the low resolution data sets are simply interleaved to form the merged data set.

A convolution model is then used to describe the relationship between $g(n\Delta x)$ and the desired high resolution data set, $h(n\Delta x) \in \mathbb{R}^{N_h}$:

$$g(n\Delta x) = \sum_{n'=0}^{N_h-1} h(n'\Delta x)\iota((n-n')\Delta x), \qquad n = 0, 1, 2, \cdots, N_h - 1. \tag{4.11}$$

The vector $\iota(n\Delta x)$ is a point spread function[5] estimated using information about the acquisition sequence and/or the object being imaged. After $\iota(n\Delta x)$ has been estimated and $g(n\Delta x)$ has been formed, $h(n\Delta x)$ is found using a chosen deconvolution algorithm.

[5] ι is the greek letter *iota*.

The MASRMRI algorithm may be described in the Fourier domain using the DFT's of the discrete signals, g, h, and ι. Using the DFT and the Fourier convolution theorem for discrete sequences, Eq. 4.11 becomes

$$G(n\Delta k) = H(n\Delta k)\mathsf{I}(n\Delta k), \qquad n = 0, 1, 2, \cdots N_h - 1. \tag{4.12}$$

Vectors $G(n\Delta k)$, $H(n\Delta k)$, and $\mathsf{I}(n\Delta k)$ are the respective DFT's of g, h, and ι, and the frequency sample spacing is $\Delta k = \frac{1}{N_h \Delta x}$. We may calculate the high resolution data set with

$$H(n\Delta k) = \frac{G(n\Delta k)}{\mathsf{I}(n\Delta k)}, \qquad \mathsf{I}(n\Delta k) \neq 0 \text{ for } 0 \leq n \leq N_h - 1. \tag{4.13}$$

Since $\mathsf{I}(n\Delta k)$ may be zero for some values of n, the calculation of $H(n\Delta k)$ may not be straightforward. However, before we concern ourselves with the deconvolution problem, there are a number of other more relevant problems at hand. As mentioned in Chapter 1, the nature of the additional information in each acquisition after the first is an issue that should be addressed.

4.2.2 Controversy on MASRMRI Research

One of the first SRMRI papers [81] that incorporated multiple acquisitions related by spatial translations was received with controversy [82, 90]. When applying SRMRI in the frequency encoding (FE) and phase encoding (PE) directions, it was claimed that there could be no new information present beyond the first image acquired. Later scientific contributions agreed with this proposal, and attention was then turned to applying SRMRI to the slice encoding direction [31, 33, 46, 47, 72, 79]. This new course appeared to be more encouraging; the physical basis for this approach was unchallenged and initial results seemed to be more promising. As a result, SRMRI in the FE direction did not appear to be - for a brief period - an effective technique for resolution enhancement in MRI.

However, a recent paper [11] published in the journal *Magnetic Resonance Imaging*, proposed that new information *can* be present in each new acquisition in the FE direction. Similar ideas were presented earlier [62, 70], using an alternative approach formulated in the frequency domain. The order of the processing steps carried out on the measured data, including the spatial shift, is significant. If the object is shifted prior to imaging, new information can in fact be present for FE data [11, 62, 63, 70]. This observation suggested that FESRMRI may still be a viable method for resolution and image quality enhancement.

The recent article [11] in *Magnetic Resonance Imaging*, presented encouraging results. Although the algorithm described in the article was not implemented, another algorithm - algebraic least-squares - yielded promising results when applied to noisy simulated data.

However, the authors found that this approach was not as successful in enhancing the spatial resolution of phantom data [11]:

> The results of the algorithm on the MRI data were nondecisive. The algorithm clearly shows resolution enhancement. Nevertheless, no clear advantage to the algorithm reconstruction is observed over the zero-padding interpolation as observed on the model results.

Several possible explanations for these results were provided, but it remains unclear as to why these results were nondecisive, despite their claim that new information is present in each image. Whilst we agree that new information about the object being imaged can be added in each acquisition, our contention is that the question of *how much new information each image contains must be, and to this point has not been, addressed.*

One of the conclusions we reach in Chapter 5 is that there is indeed new information in each acquisition when the object is shifted in the FE direction prior to imaging, as proposed earlier [11, 62, 63, 70]. *The caveat is that the amount of new information present in each acquisition after the first is relatively small* and may be difficult to detect in the presence of noise. Therefore, there may be no reasonable justification for acquiring more than one data set. Ultimately, any attempt to enhance the spatial resolution with multiple images related by a small, but known translation in the FE direction may not yield a significant advantage over more standard reconstruction approaches [53, 55, 68] that employ only a single image.

4.2.3 MASRMRI in the Frequency Encoding Direction

The discussion in this subsection appears in the author's M.Sc. thesis [62], and is summarized here for convenience. The series of processing steps that are carried out on the measured data and how the spatial shift is created are significant for MASRMRI. Following the discussion in Section 3.2.4, the expression for the unshifted, sampled signal before it is quantized in k-space is:

$$L_1(k) = \text{III}(\frac{k}{\Delta k}) \int_{-k_{max}}^{+k_{max}} s_m(k')A(k-k')dk', \qquad k \in [-k_{max}, +k_{max}]. \tag{4.14}$$

Note that the time dependence has been dropped for notational convience. The function $A(k) \in L^2(\mathbb{R})$ is an analogue, low pass filter (LPF) with inverse Fourier Transform $a(x)$. The LPF operates on the measured data $s_m(k) \in L^2(\mathcal{K})$, $\mathcal{K} = \{k \in [-k_{max}, +k_{max}]\}$, as a convolution in the frequency domain.

The second signal, $L_2(k)$, needs to be shifted in the spatial domain in the positive x direction by half a pixel, $\frac{1}{2\Delta x}$. By the Fourier shift theorem, we may obtain $L_2(k)$ from $L_1(k)$ by multiplying $L_1(k)$ with a phase ramp[6], $\exp(-i2\pi k\frac{1}{2\Delta x})$. As we will see, whether the phase ramp is applied before or after $s_m(k)$ is convolved with $A(k)$, and the nature of $A(k)$, can have different effects.

Consider an object with proton density $\rho(x)$ with spatial extent R:

$$\rho(x) = \Pi(\frac{x}{R})\rho(x),$$

so that its Fourier Transform is a function in $\mathcal{R}_R$. The object is placed in an MRI scanner, and two low-resolution data sets are produced, $L_1(k)$ and $L_2(k)$. A half pixel spatial shift is applied to the second data set by applying a phase ramp in the frequency domain. However, applying the phase ramp before or after the LPF will have different effects on $L_2(k)$. To explore how the LPF plays a role in MA SRMRI, we will consider two possibilities for how the phase ramp may be applied in the FE direction.

1. **Case I: Phase Ramp Applied After LPF**

   If the phase ramp is applied after the LPF, we have:

$$\begin{aligned} L_1(k) &= \text{III}(\frac{k}{\Delta k})\int_{-k_{max}}^{+k_{max}} s_m(k')A(k-k')dk' \\ L_2(k) &= e^{-i\pi k\Delta x}\text{III}(\frac{k}{\Delta k})\int_{-k_{max}}^{+k_{max}} s_m(k')A(k-k')dk'. \end{aligned} \tag{4.15}$$

   We can obtain $L_2(k)$ from $L_1(k)$ by multiplying $L_1(k)$ with another phase ramp *after* $\text{III}(\frac{k}{\Delta k})$ is applied! Therefore, there is no new information obtained through acquiring $L_2(k)$. Using this approach, it is not apparent that acquiring multiple data sets adds anything new.

2. **Case II: Phase Ramp Applied Before LPF**

   Contrary to what has been claimed in the literature [82, 90], if the phase ramp is applied before the LPF is applied, new information can be introduced into the second acquisition. For $k \in [-k_{max}, +k_{max}]$:

$$\begin{aligned} L_1(k) &= \text{III}(\frac{k}{\Delta k})\int_{-k_{max}}^{+k_{max}} s_m(k')A(k-k')dk' \\ L_2(k) &= \text{III}(\frac{k}{\Delta k})\int_{-k_{max}}^{+k_{max}} s_m(k')e^{-i\pi k'\Delta x}A(k-k')dk' \end{aligned} \tag{4.16}$$

   Unlike Case I, we cannot, in general, obtain $L_2(k)$ from $L_1(k)$ by multiplying $L_1(k)$ with a phase ramp. It seems that *some* new information may be present in $L_2(k)$.

[6]It is also possible to perform the spatial shift by physically moving the object. However, the advantage to this approach, over applying a phase ramp, is not clear. The accuracy of the spatial shift would also become an immediate concern.

We may also interpret Eq. 4.16 in the spatial domain. Because $s_m(k)$, which is the Fourier Transform of $\rho(x)$, was only acquired over $[-k_{max}, +k_{max}]$, $\rho(x)$ is convolved with a sinc function. After $\rho(x)$ is convolved with the sinc function it may no longer be spatially limited, and therefore the LPF may attenuate non-zero signal intensities in the spatial domain.

A priori knowledge about the object being imaged can improve the quality of the reconstruction, but how *much* information the second acquisition adds is not yet clear from this equation. The exploration of this subject is a major contribution of my thesis, and is explored in the next chapter.

4.2.4 MASRMRI in the Phase Encoding Direction

Initial work on MASRMRI presented in the literature attempted to enhance the spatial resolution in the phase encoding direction [33, 71, 80]. This strategy to enhance resolution has been found to be futile, because the data in this encoding direction is inherently discrete. Consider our model for MRI data (Eq. 3.54):

$$s_d(t) = s_d(k_x(t), k_y) = \iint M_{z_0}(x, y) e^{-i2\pi(k_x(t)x + yG_yT_{PE})} dxdy.$$

Along the PE, or k_y direction, k_x is constant. The measured data, s_d, is discrete along k_y, because $k_y = G_yT_{PE}$ is discrete[7]. Therefore, a phase ramp (applied before or after the data is sampled along the k_x direction), adds no new information to the data because the phase ramp can be easily undone with a second phase ramp of opposite slope. In other words: *the use of multiple acquisitions related by spatial translations in the PE direction to enhance the spatial resolution is futile* (see for example [90]).

4.2.5 MASRMRI in the Slice Encoding Direction

While approaches to SRMRI in the FE and PE directions have received little success, approaches to SR applied in the SE direction have encountered more promising results (and no controversy thus far).

In many applications of MRI, high resolution, three-dimensional isotropic resolution is desired. Acquisition strategies that do not employ slice encoding methods that can obtain isotropic resolution are often impractical (by increasing the acquisition time). On the other hand, acquisition techniques that employ SE methods reduce the acquisition time, but the resolution in the slice (or z) direction is often lower than the in-plane ($x-y$) resolution. It is common for example, to find reconstructed MR images with resolution $1 \times 1 \times 3$ mm^3, so that although the in-plane resolution is isotropic, the resolution in the SE direction is relatively poor.

[7] In practice, a discrete set of values are chosen for G_y.

SR methods have been investigated as a means to improve the spatial resolution in the SE direction. Essentially, the imaging volume may be acquired two or more times, with spatial shifts in the z direction between acquisitions so that the slices overlap. The slices are combined to create a merged volume, and a super-resolution algorithm is used to improve the spatial resolution in the slice-select direction.

Recent experiments in the literature (obtained from MR phantom, fruit (papaya and apple), human brain, and fMRI data) show promising results. A variety of SR algorithms have been employed, including the iterative back-projection algorithm [31, 32, 33, 47, 71, 80], regularized deconvolution [46, 47, 79], and POCS [18]. Quantitative comparisons to other approaches to resolution enhancement are often made by performing a least-squares fit of an edge within the image to a sigmoid function of the form [31, 79]

$$y(x) = \frac{1}{1 + \exp(-a(x - c))}. \tag{4.17}$$

Results obtained with SR algorithms are often compared to zero padded data sets. When SRMRI is applied to functional MRI (fMRI) data, comparisons between the activation volume in different data sets have also been used to validate results. In several studies, the SNR of the data sets improves through the use of SR techniques.

Although this area of research was not explored in the author's work on MASRMRI, this topic presents a current and viable area for research in SRMRI. Possible avenues for further investigation are discussed in Section 10.1.

Chapter 5

Frequency Encoded Super-Resolution MRI

5.1 Characterizing Information Gain

One of the main contributions of this thesis is the conclusion that SRMRI methods integrating multiple data sets related by spatial translations in the frequency encoding (FE) direction *are generally not practical*. These methods were initially proposed for resolution enhancement in MRI [81], but were initially received with controversy [90]. It was suggested that when spatial shifts are applied in the FE direction, no new information was acquired after the first image. Recent developments suggest however that shifting the object prior to imaging can add new information to each data set. For this reason, it was believed that SRMRI may be possible in the FE direction.

In this section, we point out that *the presence of new information in each acquisition is not sufficient for an SR algorithm to be practical*! Indeed, there are situations where the amount of new information is relatively small and may not be detectable in the presence of noise. We explore the question of *how much* new information can be present in each acquisition for FE SRMRI. In particular, this section investigates the extent to which the effect of the spatial shift - applied before the object is imaged - can be undone using simple image processing techniques.

Several measures are introduced in this chapter to quantify the amount of new information acquired in each data set. It is shown that when the amount of new information is relatively small, each image can be approximated by applying simple signal processing techniques to a *single* data set. Ultimately, our research suggests that little progress may be possible by MASRMRI to perform resolution enhancement in the FE direction.

The contents of this section have been published in the LNCS series [63], the European Society for Magnetic Resonance in Medicine and Biology 2006 Conference proceedings

[64], and in the journal *Magnetic Resonance Imaging* [66].

5.1.1 Modeling SRMRI

In order to describe how SRMRI could be performed in the FE direction and to permit an analysis of this technique, we use a model of the data acquisition process that assumes a one-dimensional FE acquisition. The scanner measures a k-space signal, $s(k)$, $k \in [-k_{max}, +k_{max}]$, that is convolved with a low-pass (anti-aliasing) filter $A(BW_{read}k)$. Here, BW_{read} denotes the bandwidth of $A(k)$.

In a classical SR imaging technique, multiple images, or channels, are acquired. Each image is related by a known translation, $r_n \in \mathbb{R}$, in the spatial domain. If N_{img} channels are acquired, the expression for the noiseless, demodulated, sampled MRI data $L(k_m, r_n)$ in an SR experiment is (Fig. 5.1):

$$L(k_m, r_n) = \int_{-\delta k}^{\delta k} s(k_m - \kappa) R(k_m - \kappa, r_n) A(BW_{read}\kappa) d\kappa, \tag{5.1}$$

where Δk is the frequency sample spacing, $k_{-N_l/2} = -k_{max}$, and

$$\begin{aligned} k_m &= m\Delta k, \\ m &= -N_l/2, -N_l/2+1, \cdots, N_l/2-1, \ N_l \text{ even} \\ n &= 0, 1, 2, \cdots, N_{img}-1, \\ R(k_m, r_n) &= \exp(-2\pi i r_n k_m), \\ |r_n| &< \Delta x. \end{aligned}$$

The parameter δk determines the length of the filter, $A(k)$, used to attenuate the signal. This low pass (anti-aliasing) filter attenuates the signal outside the chosen bandwidth BW_{read}.

From Eq. 5.1, the application of the filter $A(k)$ is modeled as a convolution in k-space. Mathematically, from the Fourier convolution theorem, this is equivalent to the multiplcation of the inverse Fourier Transforms of $A(k)$ and $s(k)$ [14, 29, 34, 40]. In order to reduce the amount of aliasing in the measured signal, the bandwidth of the anti-aliasing filter should be equal to or less than the bandwidth of the analogue-to-digital converter (ADC), which is $1/\Delta k$. Therefore, $BW_{read} \leq 1/\Delta k$. In this chapter, we will assume that BW_{read} is constant, and set it to its maximum value, $1/\Delta k$.

As a brief aside, the fact that the sampled signal values $L(k_m, r_n)$ are continuous functions of the spatial shift parameter r_n does not seem to have been mentioned elsewhere in the SRMRI literature. In particular,

$$L(k_m, r_n) \to L(k_m, 0) = L_1(k_m) \quad \text{as} \ \ r_n \to 0, \tag{5.2}$$

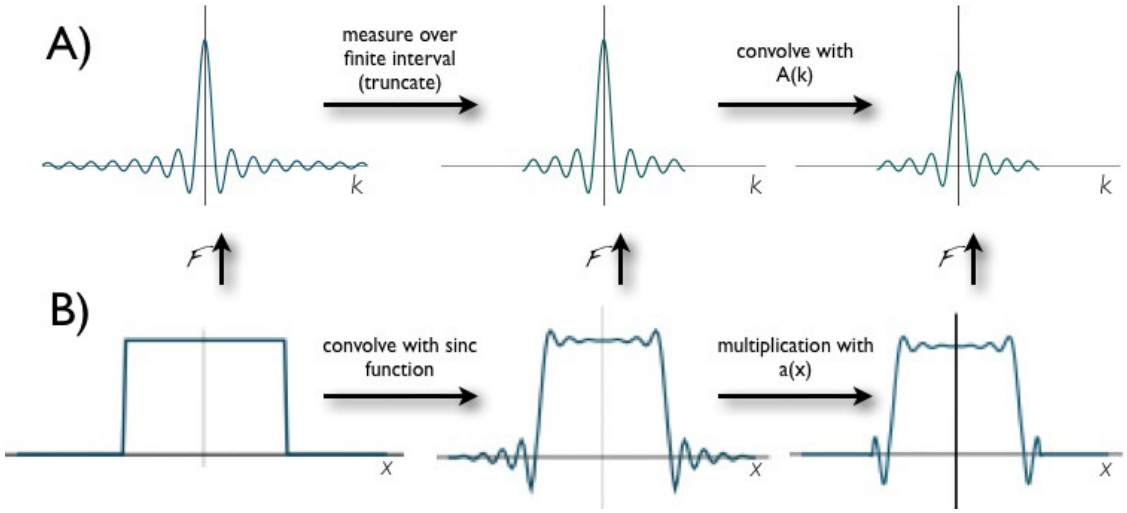


Figure 5.1: Simple schematic describing two steps in the processing of FE signals in the frequency and spatial domains. $\mathcal{F}$ denotes Fourier Transform. A) The frequency domain signal $s(k)$ is measured over only a finite range of frequencies. Before it is sampled, the signal is filtered by convolving $s(k)$ with an anti-aliasing filter, $A(k)$. B) The true object for this figure is a simple *box car* function. Measuring $s(k)$ over a finite interval spreads the spatial domain signal over the whole real line. This effect can be modeled as a convolution of the object with a sinc function. Multiplication of the result with the inverse Fourier Transform of $A(k)$ attenuates intensities in the object outside of BW_{read}.

where

$$L_1(k_m) = \int_{-\delta k}^{\delta k} s(k_m - \kappa) A(BW_{read}\kappa) d\kappa. \tag{5.3}$$

In fact, the function $L(k_m, r_n)$, and hence each component of the sampled signal, admits a Taylor series expansion in r_n which can be generated by formal expansion of the exponential in Eq. 5.1 followed by termwise integration[1]. The resulting complex-valued series will be written as

$$L(k_m, r_n) = \sum_{p=0}^{\infty} c_p r_n^p \ , \tag{5.4}$$

where (using the expression for the Taylor expansion of e^x)

$$c_p = \frac{(-2i\pi)^p}{p!} \int_{-\delta k}^{\delta k} s(k_m - \kappa)(k_m - \kappa)^p A(BW_{read}\kappa) d\kappa. \tag{5.5}$$

This result would then imply that one could construct the phase ramped signal $L(k_m, r_n)$ for nonzero values of r_n from the Taylor series to arbitrary accuracy. But – and this is the important point – one would have to know the coefficients c_m, which

[1]The boundedness of the integration interval $k \in [-k_{max}, k_{max}]$ and the uniform convergence of the Taylor expansion of the exponential over this interval guarantees the convergence of the Taylor expansion for $L(k_m, r_n)$.

means having access to the continuous frequency signal $s(k)$. The recorded signal is always discrete, but given $L_2(k_m, 0)$, $L_2(k_m, r_n)$ could be approximated using Eqs. 5.4 and 5.5 for small $|r_n|$.

5.1.2 Application of the Phase Ramp

The application of the phase ramp, $R(k_m, r_n)$, to the k-space data translates the position of the object by a distance $|r_n|$ [8, 48]. There are different methods of applying the phase ramp, which will affect the nature of the information in the measured data. We consider two cases: the phase ramp may be applied before, or after, the application of the filter $A(k)$. In each case, we will consider the acquisition of only two images, each related by a half-pixel shift relative to one another.

- **Case I: Phase ramp applied after application of $A(k)$**

  The process is illustrated schematically in Fig. 5.2. Using Eq. 5.1, we have the following two series,

  $$L_1(k_m) = L(k_m, 0) \tag{5.6a}$$
  $$L_2(k_m) = R(k_m, \Delta x/2)L(k_m, 0) = e^{-i\pi k_m \Delta x} L(k_m, 0). \tag{5.6b}$$

  Clearly, one can obtain $L_2(k)$ from $L_1(k)$ by multiplying $L_1(k)$ with a simple phase ramp after both data sets are measured. Although this method of shifting the spatial domain image has been used in the MRI literature [71, 80, 81], no new information is obtained by constructing $L_2(k_m)$, that is not already in $L_1(k_m)$. This point has been recognized by several authors in the SRMRI literature [31, 62, 82, 90].

- **Case II: Phase ramp applied before application of $A(k)$**

  The process is illustrated schematically in Fig. 5.3. In this case, we have the following series,

  $$L_1(k_m) = L(k_m, 0), \tag{5.7a}$$
  $$L_2(k_m, \Delta x/2) = L(k_m, \Delta x/2). \tag{5.7b}$$

  Note that the $L_2(k_m)$ series corresponds to the sampled and filtered data that would be obtained if the object being imaged were physically shifted by $\Delta x/2$ [62]. Unlike Case I, we cannot, in general, obtain $L_2(k_m)$ from $L_1(k)$ because of the noninvertibility of the operations behind the processing of the signal $s(k)$, namely, convolution, as well as the "degradation" that is produced by the limited range of integration. It would then seem that at least *some* new information is present in $L_2(k_m)$, which is contrary to what has been claimed in the literature [90]. The question of *how much* information is added by the second acquisition is explored in the following section of this chapter.

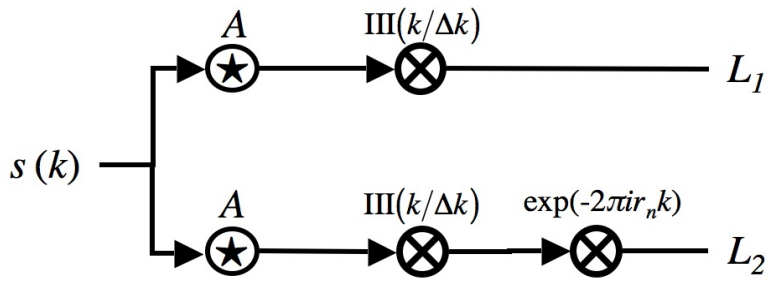


Figure 5.2: Phase ramp applied after filtering of raw frequency signal $s(k)$.

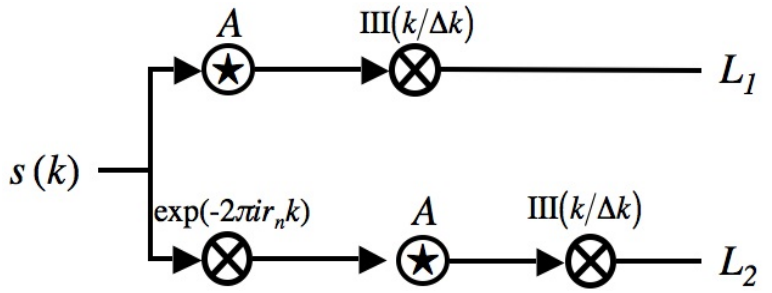


Figure 5.3: Phase ramp applied before filtering of raw frequency signal $s(k)$.

Indeed, the ramp must be applied prior to the application of the anti-aliasing filter for the acquisitions to yield independent information [11, 62, 63, 70]. To be more precise, we cannot undo the effect of the phase ramp because multiplication and integration do not commute in general:

$$\begin{aligned} R(k_m,-r_n)L(k_m,r_n) &= R(k_m,-r_n)\int_{-\delta k}^{\delta k} s(k_m-\kappa)R(k_m-\kappa,r_n)A(\kappa)d\kappa \\ &\neq \int_{-\delta k}^{\delta k} s(k_m-\kappa)R(k_m-\kappa,-r_n)R(k_m-\kappa,r_n)A(\kappa)d\kappa \\ &= L(k_m,0). \end{aligned} \tag{5.8}$$

Although it is clear that $L(k_m,0)$ and $R(k_m,-r_n)L(k_m,r_n)$ are not equal for general r_n, the extent to which they are different has not been studied extensively in the SRMRI literature. In the following section, we consider methods of quantifying their relative information, thereby determining how much new information is present in one signal relative to the other.

5.1.3 Measures of Information for Analysis of SR Data

We consider two measures that characterize the information contained in the k-space data $R(k_m,-r_n)L(k_m,+r_n) \in \mathbb{C}^{M\times N_{img}}$ for a given r_n in relation to the information contained

in $L(k_m, 0)$. An additional method is used that measures the difference between their spatial domain counterparts.

1. First, we calculate the cosine of the *angle* between $L(k_m, 0)$ and $R(k_m, -r_n)L(k_m, +r_n)$, which we denote by $C(r_n)$ [63]:

$$C(r_n) = \frac{\sum_m L^*(k_m, 0)\big[R(k_m, -r_n)L(k_m, r_n)\big]}{\sum_m |L(k_m, 0)|^2 \sum_m |L(k_m, r_n)|^2}, \tag{5.9}$$

   where * denotes complex conjugation. When $L(k_m, 0)$ and $R(k_m, -r_n)L(k_m, r_n)$ are approximately equal, $C(r_n) \approx 1$. In this case, the two signals are "highly correlated".

2. The difference, dL, between $L(k_m, 0)$ and $R(k_m, -r_n)L(k_m, r_n)$,

$$dL(k_m, r_n) = L(k_m, 0) - R(k_m, -r_n)L_2(k_m, r_n), \tag{5.10}$$

   provides an estimate of the new information contained in channel n relative to the $r_n = 0$ channel. The "energy" of dL,

$$E_{dL}(r_n) = \sum_m |dL(k_m, r_n)|^2, \tag{5.11}$$

   is a measure of the (squared) l^2 distance between the two signals. The magnitude of $E_{dL}(r_n)$ estimates the error in approximating $L(k_m, 0)$ with $R(k_m, -r_n)L(k_m, r_n)$.

3. The relative information of the two signals can also be characterized in the spatial domain. The inverse discrete Fourier Transforms (IDFT) of $L(k_m, r_n)$ and $R(k_m, -r_n)L(k_m, r_n)$ yield the respective vectors $l(j\Delta x, 0)$ and $l(j\Delta x - r_n, r_n)$, $j = 0, 1, 2, 3, \cdots, J/2 - 1$, $J = 2M$. Our analysis will use only the magnitudes of these signals. Although this ignores the phase information in the spatial domain, MRI applications almost always use only the magnitude data.

   A straightforward method of characterizing the information between them is to calculate their joint probability histogram, $p_{i,j}(r_n, N_{bin})$ [15]. The calculation of $p_{i,j}$ requires that a number of bins, N_{bin} be chosen. In our experiments, we have chosen and compared results obtained when using a range of values for N_{bin}. The joint probability histogram, $p_{i,j}(r_n, N_{bin})$, can tell us how similar $l(j\Delta x, 0)$ and $l(j\Delta x - r_n, r_n)$ are. Intuitively, if the two vectors are approximately equal for a given r_n, the intensities of $p_{i,j}$ will only occur on or near the main diagonal.

5.1.4 Analysis of the Merged Data

In the simplest FE SRMRI experiment, two data sets are acquired and $r_0 = 0$, $r_1 = \Delta x/2$. In this case, the data is "merged" in the spatial domain by interleaving pixels (as in Eq.

4.10)

$$g(j\Delta x/2) = \begin{cases} l(j\Delta x/2, 0) & j = 0, 2, 4, \cdots, J-2 \\ l((j-1)\Delta x/2, \Delta x/2) & j = 1, 3, 4, \cdots, J-1 \end{cases}, \tag{5.12}$$

where $J = 2M$. The discrete Fourier Transform (DFT) of g is given by (see Appendix A.11 for derivation)

$$G(k_m) = \begin{cases} dL(k_m + N_l\Delta k, \Delta x/2), & m \in [-N_l, -N_l/2 - 1] \\ 2L(k_m, 0) - dL(k_m, \Delta x/2), & m \in [-N_l/2, N_l/2 - 1] \\ dL(k_m - N_l\Delta k, \Delta x/2), & m \in [N_l/2, N_l - 1] \end{cases}. \tag{5.13}$$

This expression for the merged data in the frequency domain gives us another perspective on how much information is present in each acquisition. Since $dL(k_m, r_n)$ is in general nonzero for general r_n, some information may be present for $m \in [-N_l, -N_l/2 - 1]$ and $m \in [N_l/2, N_l - 1]$. However, if $dL(k_m, r_n)$ is sufficiently small relative to $2L(k_m, 0)$, the merged data may be approximated by zero padding $2L(k_m, 0)$. In the presence of noise, the difference between this zero padded data set and $G(k_m)$ may not be significant. Once again, there would be no need to acquire any more than one data set, $L(k_m, 0)$.

5.1.5 Methods

In this section, we present results from numerical calculations over a range of r_n values for two cases: (i) a simulated object, and (ii) phantom MRI data. In both cases, we ignore T_1 and T_2 relaxation.

Simulated Object

Our simulated object[2] is modeled as a sum of triangle functions[3] with unit height,

$$\rho(x) = \sum_{m=1}^{10} \Lambda\Big((x - x_m)/w\Big), \tag{5.14}$$

so that $s(k)$ was represented by a sum of 10 $\text{sinc}^2(k)$ functions modulated by phase ramps. The ten triangle functions are located at positions x_m, where $x_m = (-0.5 + m4/9)$ cm, $m = 0, 1, 2, \cdots, 9$. The parameter w controls the width of the triangle functions, and was set to 0.1 cm. The Fourier Transform of this function, shown in Fig. 5.4, is given by

$$s\big(k_m\big) = w\frac{\sin^2(w\pi k_m)}{(w\pi k_m)^2}\sum_{n=1}^{10}\exp(-2\pi i k_m x_n). \tag{5.15}$$

Eq. 5.8 was used to generate the low-resolution data sets, $L(k_m, r_n)$. For computational tractability, the filter $A(\frac{k}{\Delta k})$ is represented by a sinc function. Other filters [14, 29, 44]

[2]Simulations generated with MATLAB, version 7.0.1, R14; Mathworks, Natick, MA, USA.

[3]See Table 2.1 for the definition of Λ.

would require a significant amount of additional computational labour. The values of k that were used in our simulations were determined using typical parameters in an MR experiment. The frequency was set to $k_m = m\Delta k = \gamma G_x(m - N_l/2)\Delta t$, where $\gamma = 42.58$ MHz/T, $G_x = 0.01$ T/cm, $m = 0, 1, 2, \cdots, N_l - 1, N_l = 100000$, and $\Delta t = 10$ μs.

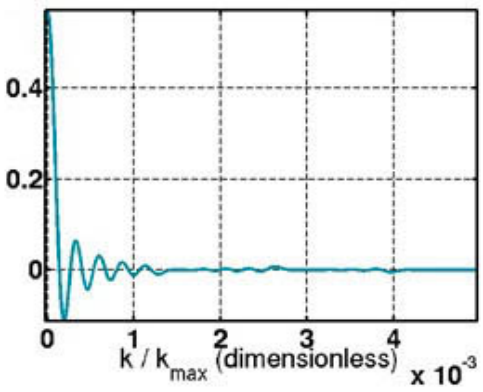


Figure 5.4: A close-up view of the real component of the k-space simulated data, $L(k_m, 0)$, near the origin. The signal is a sum of sinc^2 functions convolved with an anti-aliasing filter.

To approximate the convolution operation (Eq. 5.1), numerical integration was performed using a N_{GQ}-term Gaussian quadrature scheme [26, 57, 101]. The parameter δk in Eq. 5.1 was set to $100\Delta k$. To study the accuracy of the integration method, we compared results obtained using $N_{GQ} = 8, 16, 32$, and 64.

After the low resolution data sets were generated, a merged data set was calculated by performing the IDFTs of $L(k_m, 0)$ and $L(k_m, \Delta x/2)$. These two data sets were then interleaved according to Eq. 5.12. The DFT of the merged data set was calculated to obtain $G(k_m)$. Additionally, $C(r_n)$ and $E_{dL}(r_n)$ were calculated for $r_n \in [0, \Delta x]$.

The histogram $p(r_n, N_{bin})$ was calculated by first performing the IDFT on $L(k_m, 0)$ and $L(k_m, \Delta x/2)$. These two signals were normalized prior to calculating the histogram, so that their intensities ranged from zero to one.

Phantom Data

We also investigated the information contained in MR images of a resolution phantom (Fig. 5.5a), which consists of a series of plastic slats arranged on a 2 cm diameter teflon base. The phantom was submerged in water and imaged with a standard acquisition sequence (gradient echo, TR/TE=800/5 ms, FOV=3 cm). We acquired two k-space data sets and applied a phase ramp, $R(k_m, \Delta x/2)$ to the second data set prior to applying the anti-alaising filter. A third, unshifted image was acquired, for which the spectral extent in the FE direction was doubled. This third image corresponds to a "high-resolution image". The amplitude spectrum of the merged data set was calculated. We also computed the

difference between the merged data set and a data set obtained by simply zero padding $L(k_m, 0)$.

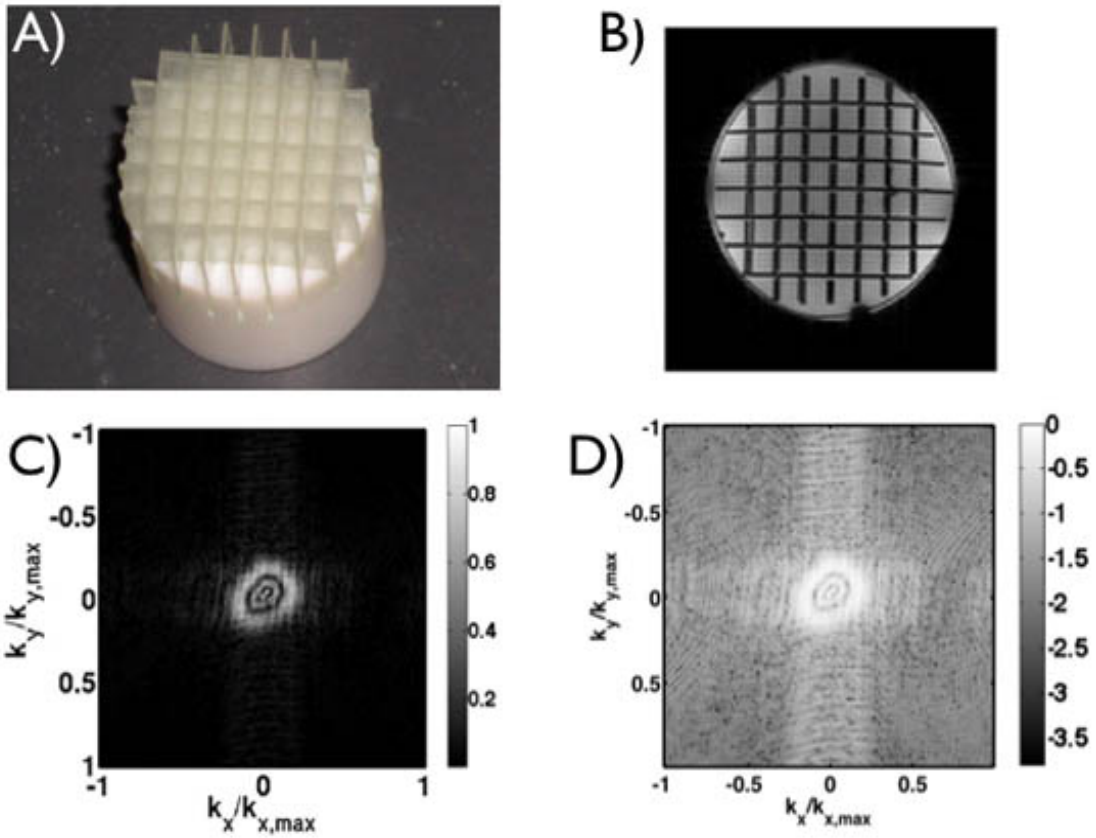


Figure 5.5: A) A photograph of the phantom consisting of plastic slats mounted on a 2 cm diameter base. B) MR image of the unshifted phantom immersed in water. C) The amplitude spectrum of the measured signal, and D) the amplitude spectrum on a logarithmic (base 10) scale.

5.1.6 Results

Simulated Object

Fig. 5.4 shows the simulated data, which is a frequency domain signal that was produced using Eq. 5.1. Simulations were carried out using a N_{GQ} term Gaussian quadrature scheme [26, 57], for $N_l = 100000$ discrete values of k_m, and $N_{img} = 11$ values of r_n. The maximum absolute value of the difference between $L(k_m, r_n)$ obtained for the cases *i*) $N_{GQ} = 8$ and $N_{GQ} = 64$, *ii*) $N_{GQ} = 16$ and $N_{GQ} = 64$, and *iii*) $N_{GQ} = 32$ and $N_{GQ} = 64$, was 0.2388, 5.7694×10^{-8}, and 1.4211×10^{-14} respectively. All of the simulation results presented in this chapter were obtained using a 64 term Gaussian quadrature scheme.

Fig. 5.6a shows the $\log_{10}$ of the amplitude spectrum of the merged data obtained

from the simulation experiment. The amplitude spectrum exhibits a sharp drop at $|k_{max}|$. Fig. 5.6b shows the ratio of E_{dL} to $E_{L(k_m,0)}$, which was monotonically increasing with increasing r_n. The maximum value was however less than 10^{-10} for the values of r_n that were used. $L(k_m, 0)$ and $R(k_m, -\Delta x/2)L(k_m, \Delta x/2)$ were found to be approximately equal. Also shown in Fig. 5.6b is a plot of $1 - C(r_n)$, which also increased with increasing r_n, but remained less than 10^{-10}. Since $C(r_n) \approx 1$, the vectors $L(k_m, 0)$ and $R(k_m, -\Delta x/2)L(k_m, \Delta x/2)$ were found to be highly correlated.

Fig. 5.7 shows the joint histograms that were calculated using $|l(j\Delta x, 0)|$ and $|l(j\Delta x - r_n, \Delta x/2)|$ for $N_{bin} = 50, 100$, and 1000. The signals were normalized so that their intensities ranged from zero to one prior to calculating the histograms, so that the bin sizes that were used were 0.05, 0.01, and 0.001. Most of the elements were found on the main diagonal in each case. The remaining mass was found on the sub-diagonals immediately adjacent to the main diagonal. More values can of course be found on off-diagonal elements when smaller bin sizes are chosen, but when the bin size was 0.05, 0.01, or 0.001, very little difference could be detected between the two signals using the joint probability histogram.

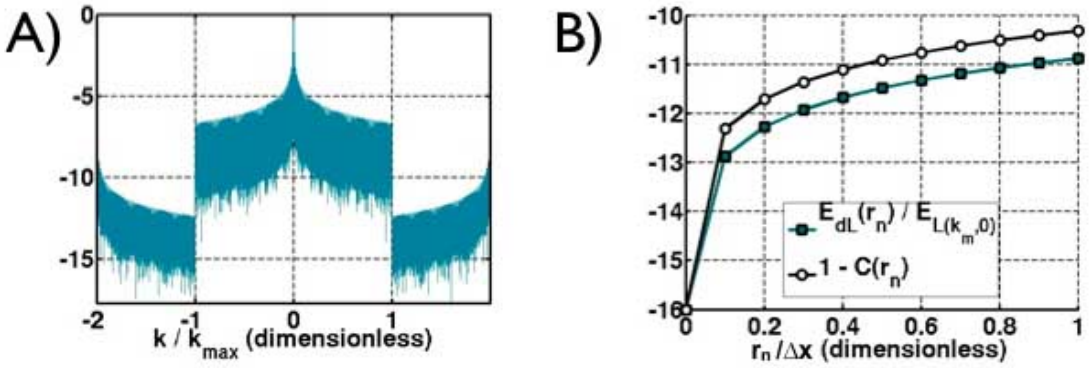


Figure 5.6: A) The $\log_{10}$ of the amplitude spectrum of the merged data from the simulation experiment. B) The $\log_{10}$ of $1 - C(r_n)$ and $E_{dL}(r_n)/E_{L(k,0)}$ for eleven values of r_n between 0 and Δx. Results in B) were also calculated from simulated data.

Phantom Data

Fig. 5.5b displays the magnitude of the measured k-space data. Fig. 5.5c is the MR image of the phantom immersed in water, obtained by using the IDFT.

Figs. 5.8a, 5.8b, and 5.8c all used the same $\log_{10}$ scale for comparison purposes. Fig. 5.8a shows the $\log_{10}$ of the magnitude of the high-resolution spectrum. This can be compared to the spectrum of the merged data, $G(k_x, k_y)$, shown in Fig. 5.8b. For $|k_y/k_{y,max}| > 1$, $G(k_x, k_y)$ appears to be mostly noise. Fig. 5.8c displays the difference between $G(k_x, k_y)$ and a zero padded version of $2L(k_x, k_y, 0)$. For $|k_y/k_{y,max}| < 1$,

$G(k_x, k_y) \approx 2L(k_x, k_y, 0)$. These observations agree with our mathematical model (Eq. 5.13) for the merged data.

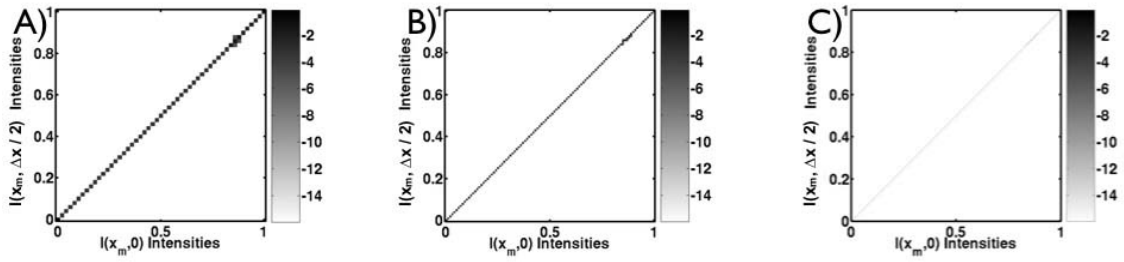


Figure 5.7: $\log_{10}$ of the joint probability histograms of spatial domain simulation data. Histograms were calculated using $l(x_m, 0)$ and $l(x_m - \Delta x/2)$. $\log_{10}0$ was defined to be 10^{-16}. No elements were found in these plots that were not on the main or sub-diagonals. The number of bins used were A) 50, B) 100, C) 1000.

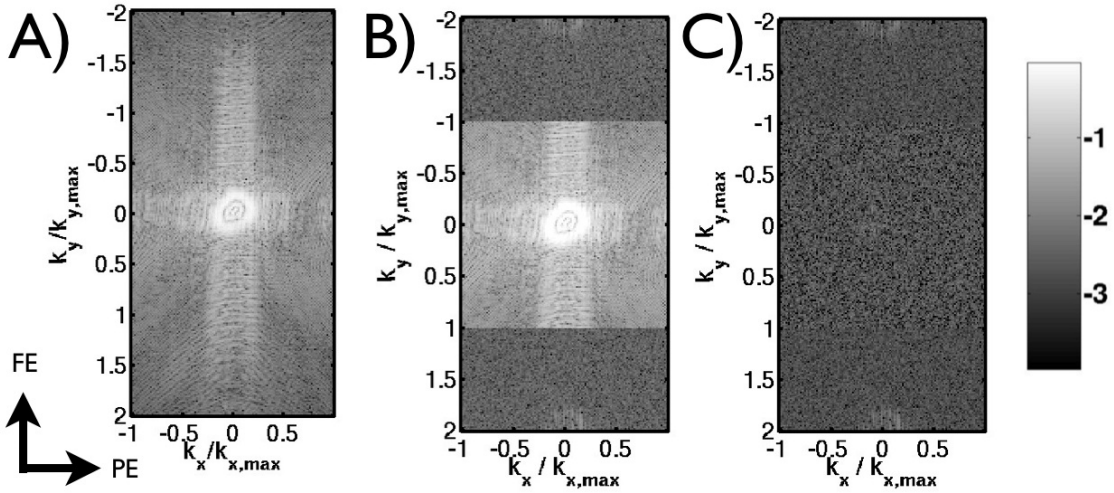


Figure 5.8: A) The $\log_{10}$ of the amplitude spectrum of the measured high resolution data set obtained by doubling the k_y extent. B) The $\log_{10}$ of the amplitude spectrum of the merged data. C) The $\log_{10}$ of the difference between a zero padded version of $L(k_x, k_y, 0)$ and the merged data.

5.1.7 Discussion

All results of our simulation experiments lead to the same conclusion for FE SRMRI: new information can be present in each acquisition. Although this result seems promising, the amount of new information is relatively small for $|r_n| < \Delta x$; $1 - C(r_n)$ and $E_{dL}/E_{L(k_m,0)}$ was less than 10^{-10} (5.6b). These quantities were nonzero - suggesting that some amount of new information may be present - but their magnitudes do not suggest that very much new information is present.

Similar results were obtained when investigating the joint probability histogram between $|l(j\Delta x, 0)|$ and $|l(j\Delta x - r_n, r_n)|$ (Fig. 5.7). Differences between these two normalized signals are at least as small as 0.002. The new information present in each channel is relatively small.

We offer two interpretations (one in the spatial domain, the other in k-space) for why $E_{dL(k_m,r_n)}$ is small relative to $E_{L(k_m,0)}$. The bandwidth of the ADC is typically larger than that of $s(k)$, because oversampling is often used in MRI to increase the SNR of the measured data [29, 44]. Since the bandwidth of $A(k)$ is set to that of the ADC, we also expect BW_{read} to be much greater than the bandwidth of s. Under these conditions, it is relatively easy to show (Appendix A.13) that $dL(k_m, r_n) \approx 0$, which is a result of the convolution integral in Eq. 5.1 having little effect on the measured signal.

There is an alternative interpretation for why $E_{dL(k_m,r_n)}$ is small relative to $E_{L(k_m,0)}$. Assume for a moment that k is a continuous parameter in Eq. 5.1 (we will later take into account the effects of sampling), and consider Eq. 5.11 in the limit as δk goes to infinity. If we again use $A(k) = \mathrm{sinc}\big(\frac{k}{\Delta k}\big)$ it can be shown that (Appendix A.14):

$$E_{dL}(r_n) = \Delta k \int_{-x_{max}}^{-x_{max}+r_n} |\hat{\rho}(x)|^2 dx - \Delta k \int_{x_{max}}^{x_{max}+r_n} |\hat{\rho}(x)|^2 dx, \tag{5.16a}$$

$$\hat{\rho}(x) = 2k_{max} \int_{-\infty}^{\infty} \rho(x')\mathrm{sinc}\big(2k_{max}(x - x')\big)dx'. \tag{5.16b}$$

The function $\rho(x)$ is the one-dimensional proton density of the object, and the variable x_{max} is half the field-of-view in the FE direction, which is typically much larger than the spatial displacement used to implement an SR procedure. For example, if there are 256 data points in the FE direction, r_n is at least 256 times smaller than $2x_{max}$.

The first integral in Eq. 5.16a is the energy of the signal that was measured in the $r_n = 0$ channel but filtered out when $r_n \neq 0$. Similarly, the second integral in Eq. 5.16a is the information that was not present in the first acquisition but present in the second acquisition. The difference between the two integrals *is the change in the information content* in the second acquisition. Essentially, the energy of $dL(k_m, r_n)$ is the amount of information from the first acquisition that was *lost* by shifting the field-of-view, minus the amount of information that was *added* by shifting the field-of-view.

We expect $E_{dL}(r_n)$ to be small relative to the energy of $L(k, 0)$. The energy of $L(k, 0)$ is calculated by integrating over the entire real line (in either the spatial domain, or k-space, using Parseval's theorem). However, the domain of integration in Eq. 5.16a is finite and small if $|r_n| \ll x_{max}$. A graphical interpretation of these concepts is presented in Fig. 5.9.

The following observations from our simple experiments with real MRI data support our model (Eq. 5.13) for the merged data set:

1. When comparing Figs. 5.8a and 5.8b, we see a very obvious drop in the amplitudes in $G(k_m)$ for $|k_y| \geq k_{y,max}$ that is not present in the high resolution data set (Fig. 5.8a). This suggests that very little information about the object is present in frequencies beyond $|k_{y,max}|$ in the merged data set, Fig.5.8b.

2. Fig. 5.8c is the difference between a zero padded version of $L(k_x, k_y, 0)$ and $G(k_x, k_y)$. The difference appears to be mostly noise. However, a small signal is detected above the noise near $|k_y/k_{y,max}| = 2$, agreeing with our simulation results (Fig. 5.6a).

These observations not only confirmed our mathematical model of the merged data set and agreed with our simulation experiments, but also verified our hypothesis that the relative amount of new information obtained in the second acquisition was difficult to detect in the presence of noise.

Thus far, research on SRMRI employing multiple acquisitions has only considered translations with magnitude $|r_n|$ less than the spatial sample spacing, Δx. Our simulation experiments suggest that the amount of new information can increase as $|r_n|$ increases. This is because the passband of any practical filter is not perfectly flat, and the amount of signal being pushed out of the filter bandwidth increases with increasing $|r_n|$. Clearly, the multiple data sets will contain a significant amount of information when all or part of the object is shifted out of the field-of-view. However, it still remains to be shown how this added knowledge will aid in the extrapolation of frequencies beyond k_{max}. While this approach could be used to extend the field-of-view, it does not necessarily aid in improving the spatial resolution.

Indeed, the set of N channels could still perhaps be used to improve the overall image quality of the data. The new information that is measured could be used to overcome small artifacts introduced by the anti-aliasing filter, or to increase the field-of-view, as suggested in Fig. 5.9. However, it is *not yet clear* how a FE SRMRI algorithm can achieve its intended purpose, which was to enhance the spatial resolution in a way that could compete with more established strategies that only require a single image [53, 55, 68].

Moreover, our work provides an explanation for some of the results obtained in the recent *Magnetic Resonance Imaging* article [11]. Results obtained when applying a FE SRMRI algorithm to real phantom MRI data [11] did not appear to surpass results obtained via zero padding, agreeing with the observations reported in this chapter (Figs. 5.6, 5.7, 5.8).

Ultimately, our research illuminates some of the difficulties in performing SRMRI in the FE direction and suggests why greater success has been encountered when applying MASRMRI in the SE direction.

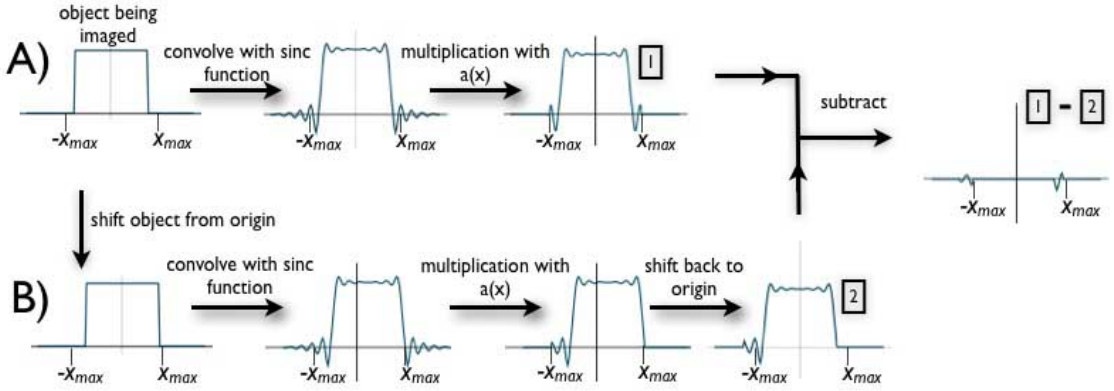


Figure 5.9: Spatial domain interpretation of $dL(k, r_n)$. A) The spatial domain data is distorted when its k-space data is scanned over a finite frequency range. The distortion can be modeled as a convolution with a sinc kernel (as shown in Fig. 5.1). After applying an anti-aliasing filter, the spatial domain data has a finite spatial extent. B) The same procedure is repeated except that the object has been shifted prior to imaging, and the object is translated again after the filter is applied. Finally, the difference between 1 and 2 is the inverse Fourier Transform of $dL(k, r_n)$.

Chapter 6

Self-Similar and Fractal Based Methods in Image Processing

6.1 Iterated Function Systems

This chapter reviews fractal-based techniques for image processing, needed for Chapters 7 and 8. One of the fundamental characteristics of fractals is their self-similarity, which arises when they are derived from an iterated function system (IFS). An IFS is a term that describes a multivariable discrete dynamical system. Under reasonable hypotheses, these systems have a unique compact invariant set. These sets exhibit certain self-similarity properties and have become known as *fractals*.

The following section describes the nature of IFS, which have applications in describing fractals and self-similar objects. This section will also introduce the inverse problem of approximation with IFS.

6.1.1 Contraction Mappings on a Metric Space

To facilitate a more precise understanding of IFS theory, several key definitions will be introduced. Much of IFS theory is based on contraction mappings and the Banach contraction mapping principle.

Definition 1 (Contraction Mappings) *Let (X, d) be a complete metric space. Then, a mapping $f : x \to x$ is said to be contractive if there exists a constant $c \in [0, 1)$ such that*

$$d(f(x), f(y)) \leqslant cd(x, y), \ \forall \ x, y \in X.$$

The constant

$$c_f = \inf_{0 \leq c < 1} c$$

is referred to as the contraction factor of f.

Another useful definition for this thesis is the set of all contraction maps on the complete metric space X.

Definition 2 (The Space of Contraction Maps) *The set of all contraction maps on a given complete metric space, X, is given by:*

$$Con(X) = \{f : x \to x \mid \exists \text{ a } c_f \in [0,1) \text{ s.t. } d(f(x), f(y)) \leqslant c_f d(x,y),\ \forall\ x, y \in X\}.$$

This leads us to the famous Banach contraction (or fixed point) theorem [5]:

Theorem 1 (Banach Fixed Point Theorem) *Let (X, d) be a complete metric space, and let $f \in Con(X)$ with contraction factor $c_f \in [0,1)$. Let the n^{th} composition of the contractive map f be denoted by $f_n(x)$. There exists a unique $\bar{x} \in X$ s.t. $f(\bar{x}) = \bar{x}$. Furthermore,*

$$d(f_n(x), \bar{x}) \to 0 \text{ as } n \to \infty,\ \forall\ x \in X.$$

The proof of this theorem may be found throughout the literature (see for example [17], page 338). The result of the Banach Contraction Theorem is that the iteration of a contraction mapping f on a complete metric space (X, d) converges to a fixed point. Given a contraction map, a metric space, and any point in that space, we can generate $\bar{x}$ via iteration.

6.1.2 Continuity of Contraction Maps

Suppose that we have two contraction maps, f_1 and $f_2 \in Con(X)$ that are very close to each other, so that they differ only by a slight perturbation, ϵ:

$$f_1(x) = \frac{1}{2}x, \qquad f_2(x) = \frac{1}{2}x + \epsilon.$$

We would guess that the respective fixed points of these maps would be close to each other. By inspection, the respective fixed points, $\bar{x}_1$ and $\bar{x}_2$, of f_1 and f_2 are:

$$\bar{x}_1 = 0, \qquad \bar{x}_2 = 2\epsilon.$$

Indeed, for small epsilon, the fixed points of the two maps are close to each other.

Fixed points of contractive maps vary continuously with respect to variations in the maps, as established by [12]. But because continuity requires the notion of a distance, we will first establish a distance function for contraction maps. To this end, for any $f, g \in Con(X)$, let us define the metric:

$$d_{Con(X)}(f(y), g(y)) = \sup_{x \in X} d_X(f(x), g(x)),$$

where d_X is the metric for space (X, d_X). The metric $d_{Con(X)}$ measures the distance between two maps $f, g \in Con(X)$ as being the farthest distance apart that a point x can be carried by maps f and g.

Theorem 2 *Let (X, d) be a metric space, and f, $g \in Con(X)$, with respective fixed points $\bar{x}_f$ and $\bar{x}_g$. If c_f is the contractivity factor of f, then [12]*

$$d_X(\bar{x}_f, \bar{x}_g) \leqslant \frac{1}{1 - c_f} d_{Con(X)}(f(x), g(x)), \ \forall \ x \in X.$$

A proof of this inequality is provided in Appendix A.15. Theorem 2 states that if f and g are close, then their respective fixed points are also close. This establishes that continuous variations in contraction maps produces continuous variations in their fixed points. This suggests that, under sufficient conditions that will be discussed shortly, a contraction map may be adjusted in a controlled manner about its fixed point.

6.1.3 The Collage Theorem

A related problem to that described above is as follows. Rather than searching for fixed points with a given map $f \in Con(X)$, we are given a point $x \in X$, and look for a contraction map f for which x is the fixed point. In general, it is not possible to construct a contraction map f whose fixed point is equal to a given x. Perhaps however, we can construct a map $f \in Con(X)$ with fixed point $\bar{x}$ that provides a close approximation to the given point, x. One question we need to ask is how close can the approximation $\bar{x}$ get to x. Mathematically, this can be expressed as the following inverse problem.

The Inverse Problem of Approximation by Fixed Points *Given an $x \in X$ and an $\epsilon > 0$, can we find an $f_\epsilon \in Con(X)$ whose fixed point $\bar{x}_\epsilon$ satisfies the inequality* [24]

$$d_X(x, \bar{x}_\epsilon) < \epsilon \ ?$$

This problem is the focus of Chapter 7, where we explore the possibility of constructing contraction maps with fixed points that approximate a measured signal, for the purpose of extrapolating frequency data.

Suppose that $x \in X$ is our target element, and that we have constructed a function, $f \in Con(X)$, with fixed point $\bar{x}$, to approximate x. The distance $d_X(x, \bar{x})$ is the error of approximating x with $\bar{x}$ (for the moment, we do not demand that $d_X(x, \bar{x})$ be small). By the triangle inequality,

$$\begin{aligned} d_X(x, \bar{x}) &\leqslant d_X(x, f(x)) + d_X(f(x), \bar{x}) \\ &= d_X(x, f(x)) + d_X(f(x), f(\bar{x})) \\ &\leqslant d_X(x, f(x)) + c_f d_X(x, \bar{x}). \end{aligned}$$

Then:

$$d_X(x, \bar{x}) \leqslant \frac{1}{1 - c_f} d_X(x, f(x)).$$

The above result is referred to as the *Collage Theorem* [5], which states that if $d_X(x, f(x))$ is small, the approximation error is small, provided that we can keep c_f away from 1.

The Collage Theorem presents an upper bound on the distance between a given fixed point, $y \in Y$ and another point $\bar{y}$ that is a fixed point to the mapping f, $f(\bar{y}) = \bar{y} \in Y$, where Y is a complete metric space. Although this suggests that f may be found that sends an element in Y close to the given element y, there is no indication from the above discussion on how to proceed to solve the inverse problem. It would seem that one way to solve the inverse problem would be to choose a set of N contraction maps $w_i \in Con(X)$, a contraction operator $\hat{w}$ from the w_i, a starting element s_0, and form an iteration sequence $s_{n+1} = \hat{w}(s_n)$. After a sufficient number of iterations, one could compute the distance between s_n and the attractor of $\hat{w}$. However, at this point, it becomes less obvious as to what direction to proceed in next; several questions arise. What if our approximation is not good enough? Do we choose another set of contraction maps, and if so, how? The Collage theorem by itself does not provide answers to these questions.

To be able to find a tractable solution to the inverse problem, one must use a more sophisticated approach to determining the contraction maps that define $\hat{w}(S)$. It has been found that by searching for a set of affine maps on the set S that are contractive, a solution to the inverse problem is possible. Namely, by looking at S as a union of smaller copies of itself that when mapped under a set of affine transforms, an approximation to the given point y can be found under appropriate conditions.

6.1.4 Iterated Function Systems with Greyscale Maps

In this section we outline the basic ideas of a special class of *iterated function systems with greyscale maps* (IFSM), a fractal-based method over functions. The treatment is brief[1] and quite specific to our applications discussed in Chapter 7.

In what follows, (X, d) denotes a compact metric space, the spatial support of our signal or image. This space is sometimes referred to as a *base space* for the given application. For this subsection, our applications will be one-dimensional real signals. We also let $\mathcal{B}(X)$ denote a suitable Banach space of functions $u : X \to \mathbb{R}$. In practice, these functions will be bounded, in which case we may also wish to specify their range[2] R_g, so that $u : X \to R_g$, where $R_g \subset \mathbb{R}$.

An N-map *affine IFSM* in x is defined by the following (for convenience, we consider the one-dimensional case, where $x \in \mathbb{R}$):

[1] For more generality and details, the interested reader is referred to [24].

[2] In the case of MRI, we may wish R_g to be $[0, 1]$ to not allow negative values to be present in the image.

1. **The spatial IFS component:** $\mathbf{w} = \{w_1, \cdots, w_N\}$, a set of affine contraction maps $w_m : X \to X$ having the form

$$w_m(x) = s_m x + a_m, \quad |s_m| < 1, \quad m = 1, 2, \cdots, N. \tag{6.1}$$

   We shall denote $X_m = w_m(X) = \{w_m(x) \mid x \in X\}$.

2. **The greyscale component:** $\mathbf{\Phi} = \{\phi_1, \cdots, \phi_N\}$, a set of associated affine greyscale maps $\phi_m : R_g \to R_g$ having the form

$$\phi_m(t) = \alpha_m t + \beta_m, \quad m = 1, 2, \cdots, N. \tag{6.2}$$

Associated with the above N-map affine IFSM is the *fractal transform* operator $\hat{T} : \mathcal{B}(X) \to \mathcal{B}(X)$. Given a function $u \in \mathcal{B}(X)$, its image $v = \hat{T}u$ is defined as

$$(\hat{T}u)(x) = \sum_{m=1}^{N} f_m(x) = \sum_{m=1}^{N} \left[\alpha_m u(w_m^{-1}(x)) + \beta_m\right] I_{X_m}(x). \tag{6.3}$$

Here $I_A(x)$ denotes the *indicator* or *characteristic function* of a set $A \subseteq X$: $I_A(x) = 1$ if $x \in A$ and $I_A(x) = 0$ otherwise. Each *fractal component* f_m is supported on the subset X_m.

Geometrically, the action of the fractal transform $\hat{T}$ may be viewed in terms of its action on the graph of u which is supported on X. Each term $u(w_m^{-1}(x))$ represents a spatially-contracted copy of the graph of u which is supported on the subset $X_m = w_m(X)$. The greyscale map ϕ_i then distorts this copy in the greyscale direction R_g to produce the fractal component $f_i(x)$. The fractal transform $\hat{T}$ then adds up these fractal components to produce $(\hat{T}u)(x)$.

We now consider the L^p function spaces, i.e., $\mathcal{B}(X) = L^p(X)$, $p = 1, 2, \cdots$, where

$$L^p(X) = \left\{ u : X \to \mathbf{R} \mid \| u \|_p = \left[\int_X |u(x)|^p \right]^{1/p} dx \; < \; \infty \right\}. \tag{6.4}$$

Consider a fixed integer $p \geq 1$. Then for any $u, v \in \mathcal{B}(X)$ [24],

$$\| Tu - Tv \|_p \;\; \leq \;\; C_p \, \| u - v \|_p, \quad C_p = \sum_{m=1}^{N} |s_m|^{1/p} |\alpha_m|. \tag{6.5}$$

A derivation of this inequality is provided in Appendix A.16. If $C_p < 1$, then $\hat{T}$ is a contraction over the space $L^p(X)$. From Banach's Fixed Point Theorem (Theorem 1), there exists a unique fixed point function $\bar{u} \in L^p(X)$ such that $\bar{u} = T\bar{u}$. Moreover, consider any "seed" function $u_0 \in L^p(X)$. If we construct the sequence of functions $u_n \in L^p(X)$ according to the iteration scheme

$$u_{n+1} = \hat{T}u_n, \quad n = 0, 1, 2, \cdots, \tag{6.6}$$

then the sequence u_n converges to $\bar{u}$, i.e., $\lim_{n\to\infty} \| u_n - \bar{u} \|_p = 0$.

From Eq. 6.3, the relation $\bar{u} = \hat{T}\bar{u}$ implies that the fixed point $\bar{u}$ exhibits *self-similarity*, as it may be expressed as a sum of spatially-contracted and greyscale-modified copies of itself.

6.2 Local IFS and Self-Similar Methods

There are certain limitations to the IFSM method described in Section 6.1.4. The fractal operator (Eq. 6.3), approximates a given target function with a linear combination of fractal components. Each of these components is a scaled and spatially contracted copy of the original target function over the base space X. However, most natural signals and images do not exhibit self-similarity in this sense.

A more practical method is based on ideas developed in the fractal coding literature [20]. Developing an approach which takes subblocks of the image and approximates them with linear combinations of subblocks of u, we construct a model of the given data that possesses a kind of self-similarity. To describe this model, let us consider normalized images, so that $u : X \to R_g$, where $R_g = [0, 1]$, and the support X of the image function u to be an $n_u \times n_u$ index matrix. The components of our model are as follows[3]:

1. A set $\mathcal{R}$ of N_R *range* blocks, $R_m \in \mathbb{R}^{n_R \times n_R}$. Each range block is supported on R_m, $m = 1, 2, 3, \ldots, N_R$, so that $R_m = u(\mathsf{R}_m)$. The range blocks are nonoverlapping and cover X, so that $\bigcup_n \mathsf{R}_n = X$.

2. A set $\mathcal{D}$ of N_D *domain blocks*, $D_m \in \mathbb{R}^{n_D \times n_D}$. Each domain block is supported on D_m, $m = 1, 2, 3, \ldots, N_D$, so that $D_m = u(\mathsf{D}_m)$. The set of domain blocks should cover X, so that $\bigcup_m \mathsf{D}_m = X$. This set may or may not be overlapping.

3. A set of N_D one-to-one *geometric transformations*, $w_{m_d,m_r,m_g}(t)$, that operate on D_m, $m = 1, 2, 3, \ldots, N_D$. For simplicity, we consider only affine transformations. There are eight possible mappings (flips and rotations) that may be accommodated. In the case that $n_R < n_D$, so that the domain blocks contain more pixels than the range blocks, it is also assumed that the contractive map w includes a pixel decimation operation. For notational convenience, let $\lambda = (m_d, m_r, m_g)$, and $\bar{D}_\lambda = \bar{D}_{m_d,m_r,m_g} = u(w_{m_d,m_r,m_g}(\mathsf{D}_m))$.

4. A set of *greyscale maps* of the form $\phi_n(t) = \alpha_n t + \beta_n$. The parameters α_n and β_n may or may not be constrained.

[3]The discussion in this section is based on [2].

Given an image function u, we examine how well the subimages $R_m = u(\mathsf{R}_m)$ are approximated by subimages $\bar{D}_\lambda$, written as

$$u(\mathsf{R}_{m_r}) \approx \phi_\lambda\big(\bar{D}_\lambda\big) = \alpha_\lambda \bar{D}_\lambda + \beta_\lambda, \tag{6.7}$$

where $\lambda = (m_d, m_r, m_g),\ 1 \le m_r \le N_R,\ 1 \le m_d \le N_D,\ m_g = 1, 2, 3, \ldots, 8$. The error associated with each domain and range pairing is given by:

$$\Delta_\lambda = \min_{\alpha_\lambda,\ \beta_\lambda} ||R_{m_r} - \phi_\lambda\big(\bar{D}_\lambda\big)||. \tag{6.8}$$

Often the L^2 (or squared L^2) norm is used to calculate Δ_λ. To minimize the collage distance for a given range block, we may perform an exhaustive search over all possible values m_d and m_g. This optimization may or may not be constrained, and in the unrestricted case, expressions for the optimal α and β yield standard regression formulas.

For the purposes of this thesis, we will focus on the following two cases of the above procedure:

1. Same-scale similarity: in this case, the domain and range blocks have the same size, so that $n_R = n_D$. In this case, the collage-error measures the similarity between domain and range blocks (or "approximability" of each range block with the set $\mathcal{D}$ of domain blocks). The statistics of same-scale approximations within an image have recently been investigated [2, 96].

2. Cross-scale similarity: in this case, the domain blocks are larger than the range blocks, so that $n_R < n_D$. This case has been explored extensively in fractal coding and compression methods [20].

The *same-scale* case has been explored recently for non-local means denoising. Here, the use of a linear combination of blocks can be used that may improve the collage error:

$$\Delta_\lambda = \min_{\alpha_\lambda,\ \beta_\lambda} ||R_{m_r} - \sum_{d \in \Theta} \phi_\lambda\big(\bar{D}_\lambda\big)||. \tag{6.9}$$

Some suitable choice of domain blocks in $\Theta \subset \mathcal{D}$ is made. This case is explored in Chapters 8 and 9, in both one and two dimensions, with MRI data.

The *cross-scale* case has been explored extensively in the fractal coding literature. Under appropriate conditions [24], the transform

$$\hat{T} = \phi_\lambda\big(\bar{D}_\lambda\big) = \alpha_\lambda \bar{D}_\lambda + \beta_\lambda \tag{6.10}$$

is a contractive operator in $L^2(X)$. By the Banach Contraction Mapping Theorem (page 64), a unique fixed point exists that may be generated via iteration. Starting from any *seed* image, u_0, the sequence $u_{n+1} = \hat{T} u_n$ converges to a fixed point, $\bar{u}$ as $n \to \infty$. This case is explored further in the following chapter, for one-dimensional k-space data.

Chapter 7

Iterated Fourier Transform Systems

In this chapter we introduce a fractal-based method over (complex-valued) Fourier Transforms of functions with compact support $X \subset \mathbb{R}$. This method of "iterated Fourier Transform systems" (IFTS) has a natural mathematical connection with the fractal-based method of "iterated function systems with greyscale maps" (IFSM) in the spatial domain, discussed in Section 6.1.4.

A major motivation for our formulation is the problem of resolution enhancement of magnetic resonance images, the main theme of this thesis. In an attempt to minimize sampling and transform artifacts, it is our desire to work directly with the raw frequency data provided by an MR imager as much as possible before image reconstruction. In this chapter, we show that our fractal-based IFTS method can be tailored to perform frequency extrapolation.

Work presented in this chapter is based on research the author has published in the LNCS series [65]. For simplicity of notation and presentation, we consider only one-dimensional MRI procedures in this chapter.

7.1 Frequency Domain Extrapolation

As we showed in Chapter 3, we assume that the object being imaged is spatially limited, and located within an interval $X = [-x_{max}, x_{max}]$ on the x-axis. It is the *proton density* of the object, to be denoted as $\rho(x)$ for $x \in X$. The (spatially) linearly varying magnetic field in the magnetic resonance spectrometer produces a complex-valued signal $s(k)$ of the real-valued frequency parameter k. The relation between $s(k)$ and the proton density

function $\rho(x)$ may be expressed as follows[1]:

$$s(k) = \int_{-\infty}^{+\infty} \rho(x)\exp(-i2\pi kx)dx, \qquad k \in \mathbb{R}, \tag{7.1}$$

so that $s(k)$ is related to $\rho(x)$ via the Fourier Transform (Chapter 2).

Because the object being imaged has a finite size (see Eq. 3.41), its Fourier spectrum necessarily has infinite support. But the data can only be collected over a finite interval, limiting the maximum obtainable spatial resolution. In MRI, the target function $U(k)$ is known only over a finite interval, $\Omega = [-k_{max}, +k_{max}]$, so that[2]

$$U_0(k) = \Pi(\frac{k}{2k_{max}})U(k). \tag{7.2}$$

The IFTS inverse problem (Section 7.3.1) can be solved to use $U_0(k)$ to estimate higher frequency components of $U(k)$, thereby achieving higher resolution in the spatial domain. In this chapter, we show that our fractal-based method of "iterated Fourier Transform systems" (IFTS) can be tailored to perform frequency extrapolation.

In order to understand the IFTS method, however, it is necessary to briefly review the method of fractal transforms in the spatial domain (see Section 6.1.4). As we will soon see, the IFTS method developed here is derived from the spatial domain fractal transform (Eq. 6.3).

7.2 From Fractal Transforms to IFTS

Our goal is now to formulate an IFS-type method over Fourier Transforms. Because our primary application (for this chapter) is signal processing, the following discussion is framed in the space $L^2(X)$ of square integrable functions on X. In the analysis that follows, we assume that our measured data comes from a spatially limited function $u \in L^2(X)$, such that

$$u(x) = u(x)\Pi\Big(\frac{x}{2x_{max}}\Big) = \begin{cases} u(x), & |x| \le x_{max} \\ 0, & |x| > x_{max} \end{cases} \tag{7.3}$$

Now consider its complex-valued Fourier Transform $U = \mathcal{F}[u]$, defined as follows,

$$\mathcal{F}[u] = U(k) = \int_{-\infty}^{\infty} e^{-i2\pi kx}u(x)\,dx, \tag{7.4}$$

[1]The raw signal is actually related to the transverse magnetization (Eq. 3.50), which in turn is related to the proton density (Eq. 3.35).

[2]In other words, we shall assume that the k-space data is acquired over a finite interval centered over the origin: $k \in \Omega = [-k_{max}, k_{max}]$.

where $k \in \mathbb{R}$. Now suppose that $v = \hat{T}u$, where $\hat{T}$ is the fractal transform in Eq. 6.3. Let the Fourier Transform of v be denoted by

$$\mathcal{F}[v] = V(k) = \int_{-\infty}^{\infty} e^{-i2\pi kx} v(x)\, dx. \tag{7.5}$$

A relationship between U and V may be derived by taking the Fourier Transform of the spatial domain fractal transform (Eq. 6.3). It can be shown that[3]:

$$\begin{aligned} V(k) &= \hat{M}U(k) \\ &= \hat{M}\mathcal{F}[u(x)] \\ &= \sum_{m=1}^{N_M} \mathrm{e}^{-2\pi i a_m k}\Big[\alpha_m s_m U(s_m k) + \beta_m \sigma_m \mathrm{sinc}(\sigma_m k)\Big], \end{aligned} \tag{7.6}$$

where

$$\sigma_m = 2x_{max} s_m.$$

Eq. 7.6 defines the action of the operator $\hat{M}$ that is induced by the fractal transform $\hat{T}$. The relation $v = \hat{T}u$ implies that $V = \hat{M}U$, where $U = \mathcal{F}[u]$ and $V = \mathcal{F}[v]$. Technically, $\hat{M}$ is a mapping from the space of functions that are Fourier Transforms of functions in $\mathcal{B}(X)$ to itself – let us call this space $\mathcal{B}^{FT}(X)$. We shall refer to $\hat{M}$ as an IFTS operator, for *iterated function systems on Fourier Transforms*.

Let us compare the frequency-space result in Eq. 7.6 with its spatial counterpart in Eq. 6.3. Here, $V(k)$ is a sum of distorted, *frequency-expanded* copies of $U(k)$ along with various modified copies of sinc functions. Note that the value of $U(k)$ is determined by values of U at *lower frequencies* $s_m k$. This can provide a mechanism for frequency extrapolation.

The following properties of the IFTS operator are derived (where necessary) in the appendices:

1. The IFTS operator, $\hat{M}$, maps functions $L^2(X)$ to itself (see Appendix A.18)
2. Let $U, V \in L^2(X)$. Then using the L^2 norm,

$$\| \hat{M}U - \hat{M}V \|_2 \;\leq\; C_{IFTS,2} \| U - V \|_2, \tag{7.7}$$

   where $C_{IFTS,2}$ is given by (see Appendix A.19)

$$C_{IFTS,2} = \sum_{m=1}^{N_M} |s_m|^{1/2} |\alpha_m|.$$

   The IFSM operator $\hat{T}$ and the IFTS operator $\hat{M}$ have the same Lipschitz constant in the L^2 norm[4].

[3] See Appendix A.17 for derivation.

[4] This is not the case for the L^p norm, as we show in Appendix A.19.

3. By the definition of contractivity, if the magnitude of the Lipschitz constant of the IFTS operator is less than one, then $\hat{M}$ is contractive over L^2 (and $\hat{T}$ will also be contractive).

4. As the IFSM and IFTS operators are one-to-one, their respective fixed points, $\bar{u} \in L^2$ and $\bar{U} \in L^2$, will be related as $\bar{U} = \mathcal{F}[\bar{u}]$ (Fig. 7.1).

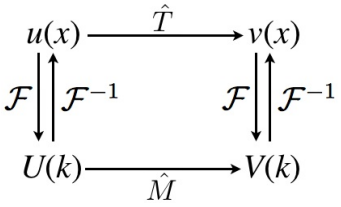


Figure 7.1: Relationship between the IFTS and IFSM operators.

Examples:

1. The following $N_M = 2$-map affine IFSM on $X = [-1, 1]$:

$$w_1(x) = \frac{1}{2}x - \frac{1}{2}, \;\; \phi_1(t) = \frac{1}{2}t + \frac{1}{2},$$
$$w_2(x) = \frac{1}{2}x + \frac{1}{2}, \;\; \phi_2(t) = \frac{1}{2}t + \frac{1}{2}.$$

Note that $X_1 = w_1(X) = [-1, 0]$ and $X_2 = w_2(X) = [0, 1]$: The two sets are nonoverlapping in the L^2 sense, even though they "touch" at $x = 0$. The fixed point for the associated IFSM operator $\hat{T}$, using Eq. 6.3, is $\bar{u}(x) = 1$ for $x \in X$. The fixed point for the corresponding IFTS operator $\hat{M}$, using Eq. 6.3 is the real-valued function $\bar{U}(k) = 2\frac{\sin \pi k}{\pi k} = \mathcal{F}(\bar{u})$.

2. A perturbation of the above system:

$$w_1(x) = \frac{3}{5}x - \frac{2}{5}, \;\; \phi_1(t) = \frac{3}{10}t + \frac{1}{2},$$
$$w_2(x) = \frac{3}{5}x + \frac{2}{5}, \;\; \phi_2(t) = \frac{1}{2}t + \frac{1}{2}.$$

Here $X_1 = w_1(X) = [-1, \frac{2}{5}]$ and $X_2 = w_2(X) = [-\frac{2}{5}, 1]$: The two sets overlap over the interval $[-\frac{2}{5}, \frac{2}{5}]$. The fixed points, $\bar{u}$ and $\bar{U}$ for, respectively, the IFSM operator $\hat{T}$ and the IFTS operator $\hat{U}$ are shown in Fig. 7.2 below. Note that the attractor $\bar{u}$ possesses a self-similarity – the "piling up" of function values over the overlap region is repeated over the entire interval X under the action of the IFS maps w_i. On the other hand, the Fourier Transform $\bar{U}$ is a mild perturbation of the transform of Example 1, most notably in the non-zero imaginary part. The attractors were computed by means of iteration of the operators, $\hat{T}$ and $\hat{U}$ in their respective spatial and frequency domains.

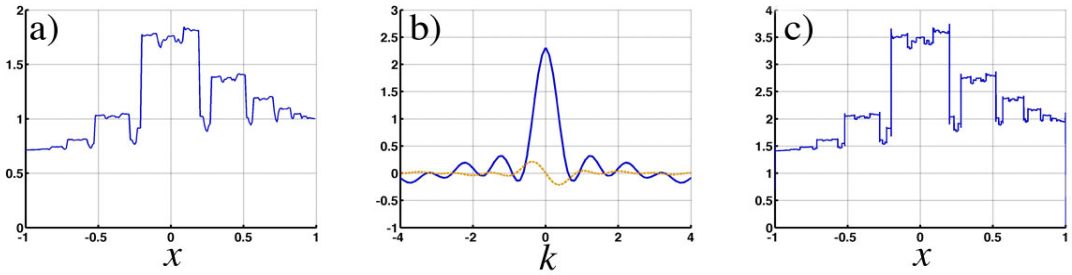


Figure 7.2: Fixed point attractors (a) $\bar{u}$ and (b) $\bar{U}$ (real and imaginary parts) for the IFSM/IFTS of Example 2 above. (c) is the IDFT of the fixed point $\bar{U}$ in (b).

7.3 IFTS and Frequency Extrapolation

We now have the ingredients to perform the following inverse problem of Fourier Transform approximation: Given a "target" Fourier Transform $U \in \mathcal{B}^{FT}(X)$, find an IFTS operator $\hat{M}$ with fixed point $\bar{U}$ that approximates U to an acceptable accuracy.

We adopt a *collage coding* strategy, and look for an operator $\hat{M}$ that minimizes the collage distance $\Delta = d_{FT}(U, \hat{M}U)$, where d_{FT} denotes a suitable metric on the space of functions $\mathcal{B}^{FT}(X)$.

7.3.1 Solving the IFTS Inverse Problem

Our procedure will mimic the usual procedure of IFSM fractal coding: We shall fix N_M, the number of components in the IFTS operator $\hat{M}$ in Eq. 6.3, as well as the affine spatial contraction maps w_i in Eq. 6.1, $1 \leq m \leq N_M$. The optimal values of the real greyscale coefficients α_i and β_i in Eq. 6.2 will then be given by

$$(\alpha^*, \beta^*) \;=\; \arg \min_{(\alpha,\beta)\in\Pi} d_{FT}(U, \hat{M}(\mathbf{s}, \mathbf{a}, \alpha, \beta)U). \tag{7.8}$$

Here, $\Pi \in \mathbb{R}^{2N_M}$ denotes a *feasible parameter space* that ensures contractivity of the IFTS operator $\hat{M}$.

As in traditional fractal image coding, it is convenient to use the L^2 metric. The squared L^2 collage distance between U and $\hat{M}U$ is

$$\begin{aligned}\Delta^2(\alpha, \beta) &= \| \, U - \hat{M}U \, \|_2^2 \\ &= \int_{\mathbb{R}} \left| U(k) - \sum_{m=1}^{N_M} \mathrm{e}^{-2\pi i a_m k} \Big[\alpha_m s_m U(s_m k) + \beta_m \sigma_m \mathrm{sinc}(\sigma_m k) \Big] \right|^2 dk. \end{aligned} \tag{7.9}$$

Note that Δ^2 may also be rewritten as a quadratic form,

$$\Delta^2(\mathbf{x}) = \| \, U \, \|^2 + \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}\mathbf{x}, \tag{7.10}$$

where $\mathbf{x}^T = (\alpha_1, \cdots, \alpha_{N_M}, \beta_1, \cdots, \beta_{N_M})$. Here, $\mathbf{A}$ may be expressed as a $2N_M \times 2N_M$ symmetric real matrix and $\mathbf{b}$ a real $2N_M$-vector. The minimization of Δ^2 may now be accomplished with quadratic programming algorithms, generally at a fraction of the time required for gradient descent methods.

Of course, the minimization must be performed with constraints. The contractivity constraint $C_{IFTS,2} < 1$ from Eq. 6.5 is a very strict condition, having been obtained by a simple application of the triangle inequality. From a knowledge of the maps w_i involved and the overlapping/nonoverlapping properties of the associated sets X_i, one may be able impose more relaxed conditions. Additional constraints may be imposed from the requirement that the associated spatial fixed point function $\bar{u}(x)$ be nonnegative over X, e.g., $\alpha_m, \beta_m \geq 0$ (see condition 2, page 28).

Another concern is to make the collage distance "suitably small." *Theoretically*, the collage distance can be made arbitrarily small by suitably increasing the refinement of the partition produced by the spatial IFS maps over the interval X. (The "proof" is basically the same as presented in [23].) A systematic and quite effective refinement is produced by using N_M-map truncations $\mathbf{w}_{N_M}$ of the infinite set $\mathbf{w} = \{w_{11}, w_{12}, w_{21}, \cdots\}$ of spatial IFS maps on $X = [-1, 1]$ given by

$$w_{nj} = \frac{1}{2^n}(x + 2j - 1 - 2^n), \quad n = 1, 2, \cdots, \quad 1 \leq j \leq 2^n. \tag{7.11}$$

This system of maps produces a decomposition of X into dyadic subintervals. The natural choices for N_M are 2, $2 + 4 = 6$, $2 + 4 + 8 = 14$, ..., $2 + \cdots + 2^p = 2^{p+1} - 2$, $p = 1, 2, \cdots$, which produce complete dyadic refinements of $[-1, 1]$ down to $\Delta x = 2^{-(p-1)}$. This system also permits "mixing" between overlapping intervals of varying size. Such a strategy was adopted successfully in [22].

Example 3: We consider a complex-valued function[5] $U(k) \in \mathcal{R}_R$, $R = 1$, the real and imaginary components of which are plotted in 7.3(a) and 7.3(b), respectively (solid line plot). Its inverse Fourier Transform, the spatial function $u(x)$, is plotted in 7.4(a) (apart from ringing artifacts outside the interval $X = [-1, 1]$ due to the finiteness of Ω).

With $U(k)$ as our target function, we applied L^2 collage coding, using the system of "dyadic" IFS maps given in Eq. 7.11. For $N_M = 2, 6$ and 14, the minimal collage errors Δ obtained were 2.13, 1.70 and 1.08, respectively. The Harwell HSL Archive FORTRAN subroutine VE17AD was used to solve the associated quadratic programming problems. In all cases, the solution was provided in a fraction of a second (in contrast, gradient search methods required on the order of an hour of computer time to solve the 14-map problem). Figs. 7.3(a) and 7.3(b) display the real and imaginary components of the fixed point $\bar{U}(k)$ of the 14-map IFTS. Note that the relatively well-behaved real part of the target $U(k)$ is very well approximated over the interval $[-k_{max}, +k_{max}]$.

[5]Recall that $\mathcal{R}_R$, defined in Eq. 2.7, is the space of *frequency domain* functions that have a finite spatial extent of R.

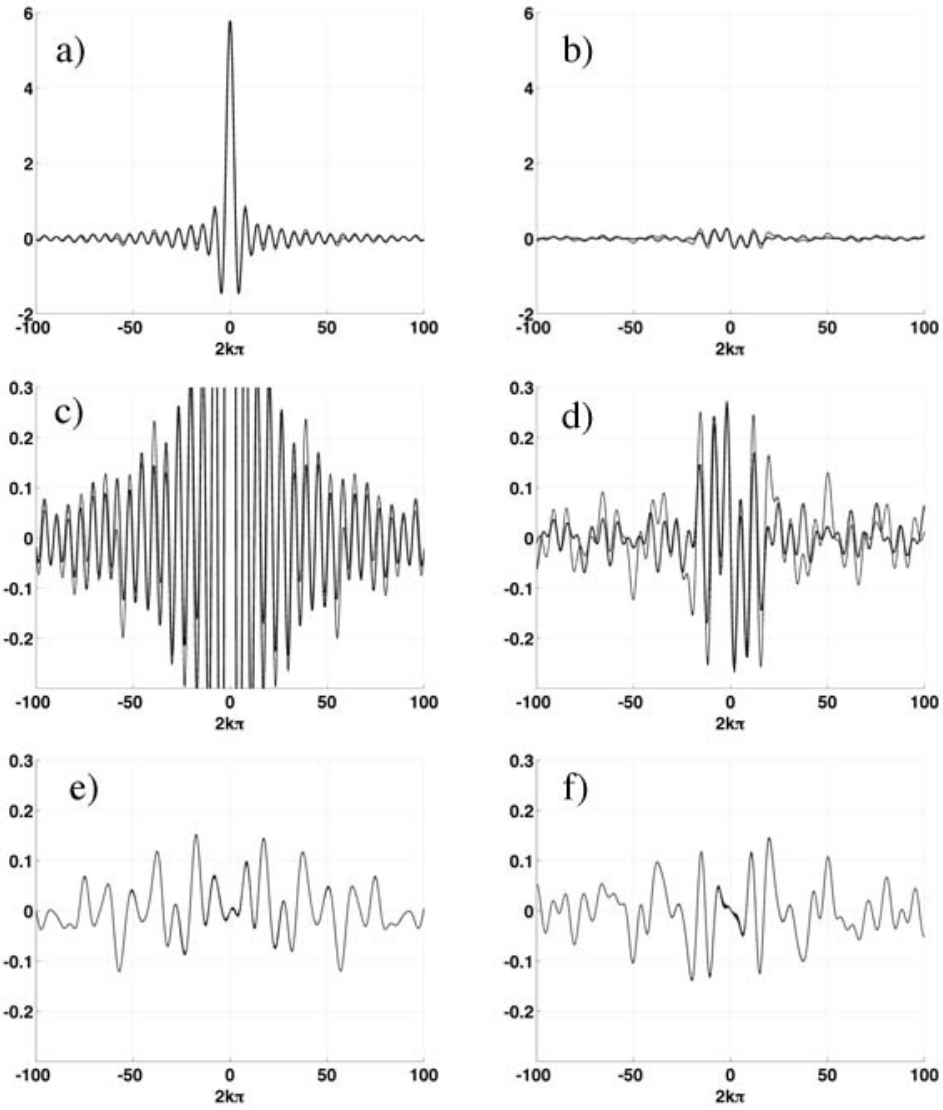


Figure 7.3: Figures (a) through (d) show the target $U(k)$ (solid plot) and fixed point approximation $\bar{U}(k)$ (dotted plot) yielded by a 14-IFS map IFTS described in Example 3. Shown in (a) and (c) are the real components. Shown in (b) and (d) are the imaginary components. Note that Figures (c) and (d) are vertically expanded versions of (a) and (b), respectively, in order to see better the details of the functions, especially the imaginary part. In (e) and (f) are shown the residuals of, respectively, the real and imaginary components, using the same scale as in (c) and (d).

7.3.2 Frequency Extrapolation

The use of IFTS to perform frequency extrapolation is motivated by the action of the IFTS operator $\hat{M}$ in Eq. 6.3 and noted earlier: The value of $V(k) = (\hat{M}U)(k)$ is determined by values of $U(k)$ at lower frequencies $s_m k$. Consider any U_0 with support $\Omega_0 = [-k_0, k_0]$. Now let $U_1(k) = \hat{M}U_0(k)$. It follows that the support of $U_1(k)$ is the interval $\Omega_1 = [-Ak_0, Ak_0]$, where $A = [\max_k |s_m|]^{-1}$. In general, the support of $U_n = \hat{M}^n U_0$ is $\Omega_n = [-A^n k_0, A^n k_n]$. In this way, lower frequency information makes its way outward, implying frequency extrapolation.

Working with the truncated transform U_0, however, introduces an additional approximation error into the *collage coding* procedure. The following result is a straightforward consequence of the Collage Theorem and the triangle inequality:

> Given a target transform $U \in L^2(X)$ and an approximation U_0 to U, with $d_{FT}(U, U_0) = \epsilon_1$. Now suppose that we perform collage coding on U_0, that is, we find an N_M-map affine IFTS $\hat{M}_0$ with contraction factor $c_0 \in [0, 1)$ such that $d_{FT}(U_0, \hat{M}_0 U_0) \leq \epsilon_2$, where $\epsilon_2 > 0$. Then
>
> $$d_{FT}(U, \bar{U}_0) \leq \frac{1}{1 - c_0}\epsilon_2 + \epsilon_1, \tag{7.12}$$
>
> where $\bar{U}_0$ is the fixed point of M_0.

In the limit $\delta \to 0$, we retrieve the Collage Theorem (Section 6.1.3).

No matter how small we may force the collage error ϵ by suitable refinement, the error $\delta > 0$ serves as a barrier toward the approximation of our (unknown) target U with the fixed point $\bar{U}$, as will be seen below.

Example 3 revisited: Let us return to the target transform $U(k)$ presented in Fig. 7.3, with support $\Omega = [-100, 100]$. When inverted, it yields the spatial profile $u(x)$ in Fig. 7.4(a). Now suppose that we are given only one-half of this data, i.e. $U_0 = U|_{[-50,50]}$, with support $\Omega_0 = [-50, 50]$. Inversion of U_0 yields the lower-resolution profile $u_0(x)$ pictured in Fig. 7.4(b). The L^2 error of approximation is $\delta = \| u - u_0 \|_2 = 0.066$. In this example, all spatial functions on $[-1, 1]$ have been normalized in the L^2 norm.

Collage coding of U_0 with a 14-map IFTS, using the w_{ij} maps in Eq. 7.11, produces a fixed point $\bar{U}_0$, the inversion of which is shown in Fig. 7.4(c). Even though $\bar{u}_0$ has much more spatial resolution, it is a poorer approximation to u than u_0, with an L^2 error $\delta = 0.100$. One possible reason for this observation: It is not guaranteed that the given low frequency component of U – namely, U_0 – is well approximated by $\bar{U}_0$. In order to bypass this problem – and in the spirit of the Papoulis-Gerchberg algorithm for frequency extrapolation – we simply replace the frequency components $\bar{U}_0|_{\Omega_0}$ with the original data U_0, so that

$$\bar{U}_0^{P_0}(k) = [1 - I_{\Omega_0}(k)]\bar{U}_0(k) + U_0(k), \qquad k \in \Omega. \tag{7.13}$$

Mathematically, $\bar{U}_0^{P_0}$ is the projection of $\bar{U}_0$ onto the subspace P_0 of transforms with frequency subspectrum U_0 on Ω_0.

Inversion of $\bar{U}_0^{P_0}$ over $[-100, 100]$ yields the spatial profile pictured in Fig. 7.4(d). Visually, it appears to be a much better higher-resolution approximation to u than $\bar{u}_0$. Quantitatively, its L^2 error of $\delta = 0.052$ is also smaller. This L^2 error is also lower than that of u_0, which achieves one of the main goals of this entire exercise.

Another strategy is to consider the fact that collage coding minimizes the distance between a target U and its image $\hat{M}U$. For this reason, we also consider the following construction,

$$U_1^{P_0}(k) = [1 - I_{\Omega_0}(k)](\hat{M}U_0)(k) + U_0(k). \tag{7.14}$$

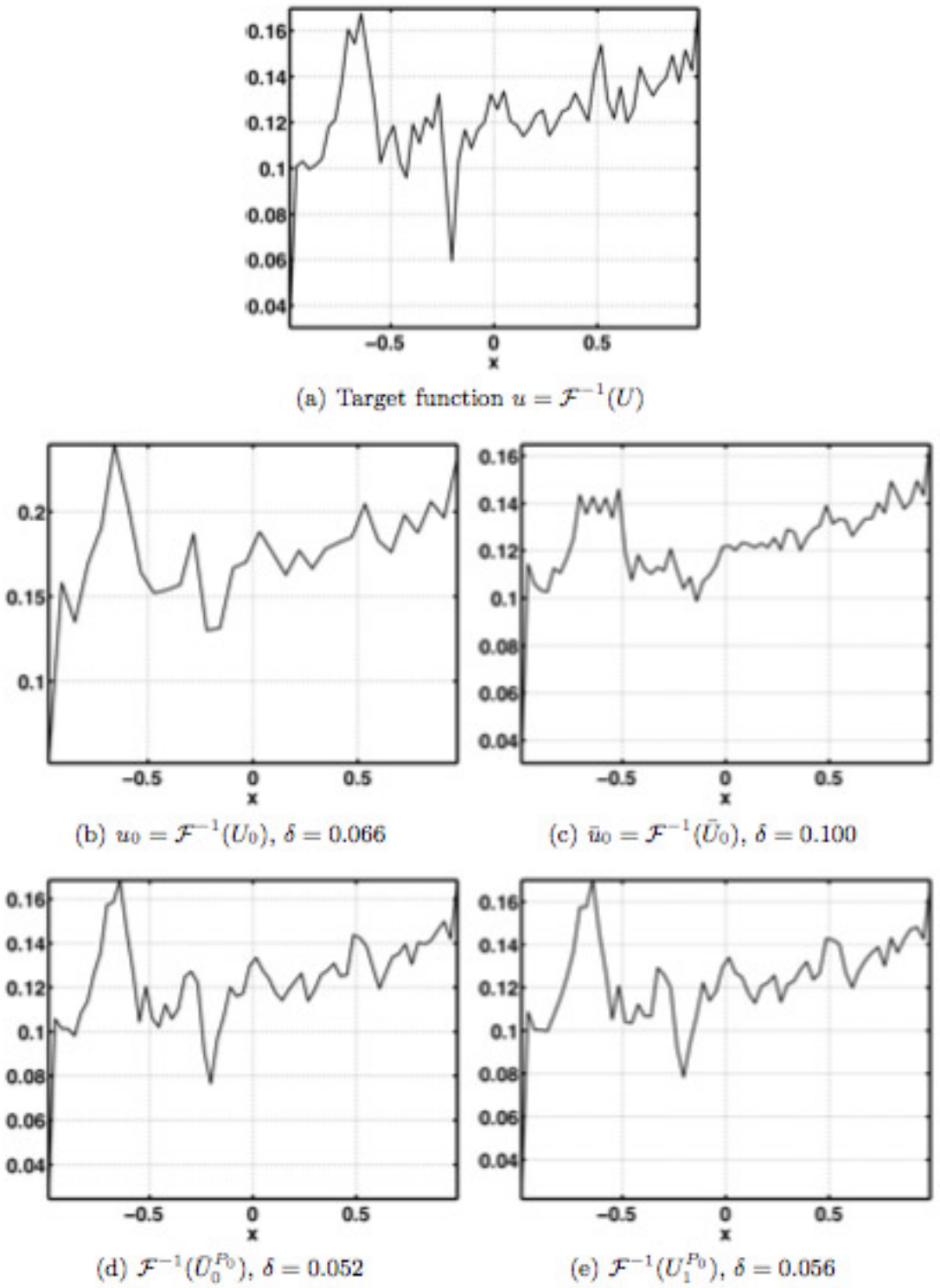
In other words, we collage code the original data U_0, construct the first iterate $U_1 = \hat{M}_0 U_0$ and once again project onto the subspace P_0. The resulting spatial profile is shown in Fig. 7.4(e). It also provides a very good higher-resolution approximation to the target spatial profile $u(x)$, both numerically (L^2 approximation error is $\delta = 0.056$) as well as visually.

7.4 Conclusions

Approaches in quantifying the relative quality of our reconstructions obtained using techniques presented in this chapter is consistent with what is commonly used in the MRI literature. Measures of error based on the L^2 metric, and "observer measures" are both commonly used strategies to compare results obtained from different post-processing algorithms (see for example [53, 68]).

Finally, the computations described in this chapter represent only a preliminary study of self-similar methods for frequency extrapolation. We see that because the IFTS operator is derived from the IFSM operator, it constructs fractal components from the entire function. However, as there is no connection between this approach to a physical model of the MR data, there is no reason why we would expect this approach for frequency extrapolation to be effective.

A more practical approach, used in fractal compression literature, is to first partition the data into subblocks and to then construct components based on these subblocks. In fact, we find in the following chapter that this procedure has several advantages over the methods presented in Chapter 7. The most important of which is the connection between linear combinations of these blocks in k-space, and a physical model for the MR signal.



(a) Target function $u = \mathcal{F}^{-1}(U)$

(b) $u_0 = \mathcal{F}^{-1}(U_0)$, $\delta = 0.066$

(c) $\bar{u}_0 = \mathcal{F}^{-1}(\bar{U}_0)$, $\delta = 0.100$

(d) $\mathcal{F}^{-1}(\bar{U}_0^{P_0})$, $\delta = 0.052$

(e) $\mathcal{F}^{-1}(U_1^{P_0})$, $\delta = 0.056$

Figure 7.4: (a): The target spatial function $u(x) = \mathcal{F}^{-1}(U)(k)$ corresponding to the Fourier Transform $U(k)$ plotted in Fig. 7.3. Figs. 7.4(b)-(e): Approximations yielded by the bandlimited data U_0 and various frequency extrapolation methods applied to U_0, as described in text. L^2 approximation errors are also given. All functions have been normalized in L^2 on $X = [-1, 1]$.

Chapter 8

Local Self-similarity of Fourier Domain MRI Data

8.1 Introduction

Results presented in this chapter represent an extension of the research presented in Chapter 7 on the use of self-similarity for the resolution enhancement of MRI data. In the previous chapter, we showed how resolution enhancement could be accomplished by means of a fractal-based method over complex-valued Fourier Transforms with compact support $X \subset \mathbb{R}$. Our method of "iterated Fourier Transform systems" (IFTS) is the Fourier domain counterpart of the fractal-based method of "iterated function systems with greyscale maps" (IFSM) in the spatial domain [24]. The action of an IFTS operator $\hat{M}$ on a Fourier Transform $U(k)$ is to produce frequency-expanded and phase-ramped, range-distorted copies of $U(k)$. With this method, higher frequency components of $U(k)$ outside the interval Ω_0 are estimated, thereby achieving higher spatial domain resolution.

The IFTS method described in the previous chapter, does not yet have a connection to physical parameters of the object being imaged. In this chapter, we explore the use of other self-similarity methods that may aid in frequency extrapolation, motivated by a connection we present here to ARMA methods and a mathematical model of the MRI data. A further major motivation is provided by recent work [2], showing that images are, in general, affinely self-similar locally: Given a "range block" $u(R_i)$ of an image, there are generally a good number of "domain blocks" $u(D_j)$ that can approximate it well under the action of affine greyscale transforms. This self-similarity in the spatial domain is dramatically demonstrated when we plot errors of approximation for all domain-range pairings.

Here we demonstrate that such self-similarity is also exhibited by subblocks of Fourier data. The underlying explanation for this block-based self-similarity is that a connection

can be made between the well-known result of autoregressive (AR) correlation coefficients and block-based model parameters. This justifies block-based coding in the complex Fourier domain, which we then employ for the purpose of frequency extrapolation.

To this end, we show that discrete functions of the form,

$$s(m) = \sum_{n=1}^{N} c(n) e^{\sigma_n m}, \qquad m \in \mathbb{Z}, \quad c_n, \sigma_n \in \mathbb{C} \tag{8.1}$$

are locally self-similar. Functions of this form are found in many applications, but here we consider their use in MRI. Eq. 1.9 has been used to model the measured MR data in order to improve the image reconstruction process and to reduce artifacts [35].

Block-based fractal coding in the wavelet domain is a rather standard procedure [95]. To the best of our knowledge, however, there has been little analysis to date on the use of block-based methods to analyze or process Fourier data. This research presented in this chapter (and the previous chapter, based on [65]) represents work in this direction, and is based on research presented in [67].

8.2 One-Dimensional MRI Data Models

Here, we outline general spatial and frequency domain models for MRI data, where the raw data is a Fourier spectrum of the desired image. For simplicity of notation and presentation, we first consider one-dimensional MRI procedures.

As discussed in Chapter 3, the MR scanner responds to the *proton density* of the object, $\rho(x)$, and produces a complex-valued signal $s(k)$ of the real-valued frequency parameter k. The relation between $s(k)$ and the proton density function $\rho(x)$ may be expressed as (Eq. 3.52):

$$s(k) = \int_{-\infty}^{+\infty} \rho(x) \exp(-i2\pi kx) dx, \qquad k \in \mathbb{R}. \tag{8.2}$$

In other words, $s(k)$ is related to $\rho(x)$ via the Fourier Transform (Chapter 2). If $s(k)$ is known $\forall\ k \in \mathbb{R}$, then $\rho(x)$ may be found by using the inverse Fourier Transform of $s(k)$. In practice, however, the true proton density, $\rho(x)$, cannot be obtained exactly, and must be estimated. One reason is that $s(k)$ is determined over a finite interval. Ultimately, reconstruction yields only approximations to $\rho(x)$ with finite spatial resolution.

8.2.1 Self-Similarity of the One-Dimensional MR Signal

Many of the self-similar based methods developed in the literature utilize local self-similarities between different regions of real *spatial domain* images. For these methods to

be effective on complex *Fourier domain* data we require that the data exhibit some degree of self-similarity. By self-similar, we mean that subblocks may be well approximated by other subblocks. This procedure has been examined for maps in the spatial domain [2]. The extent to which MRI Fourier data is locally self-similar has, to the best of the author's knowledge, only been explored in [67].

In order to initiate a discussion of the self-similarity of Fourier data, we first consider the complex, discrete signal, $s(n) \in l^2(\mathbb{R}), \quad n = 0, 1, 2, \ldots, N_S - 1$. Then define a partitioning of $s(n)$ into domain and range vectors (or *blocks*) and affine transforms that operate on these blocks. The last N_P values of $s(n)$ will constitute the range block, $\mathbf{r}$:

$$\mathbf{r} = [s(N_S - N_P), s(N_S - N_P + 1), \ldots, s(N_S - 1)]. \tag{8.3}$$

The domain vectors are constructed by partitioning $s(n)$ into $N_{DB} = N_S - N_P$ *overlapping* domain blocks:

$$\begin{aligned} \mathbf{d}_m = [s(p), s(p+1), s(p+2), \ldots, s(p+N_P-1)], \\ p = N_S - N_P - m, \\ m = 1, 2, \ldots, N_{DB}. \end{aligned}$$

The set of all domain blocks, $\mathcal{D} = \{\mathbf{d}_m \in \mathbb{C}^{N_P} |\, m = 1, 2, \ldots, N_{DB}\}$, comprises the *domain pool*.

Affine transforms can be used to search for similarities between the intensities of range and domain blocks. If for a given range block, we can find an affine transform T, and domain block d_m, such that $\mathbf{r} \approx T\mathbf{d}_m$, then the given data possesses some degree of self-similarity.

The simplest such affine transform assumes the form

$$\mathbf{r} \approx T_1(\mathbf{d}_m) = \alpha_m \mathbf{d}_m + \beta_m, \tag{8.4}$$

where $\alpha_m, \beta_m \in \mathbb{C}$. The subscript "1" of the operator T denotes that a *single* domain block is being used to approximate $\mathbf{r}$. The complex parameters α_m and β_m may be chosen by minimizing the l^2 (vector) norm, Δ_m of the difference between $\mathbf{r} \in \mathbb{C}^{N_P}$ and each affinely-transformed domain vector[1]:

$$\Delta_m = \sqrt{\frac{1}{N_P} \sum_{n=1}^{N_P} \left|\mathbf{r}(n) - [\alpha_m \mathbf{d}_m(n) + \beta_m]\right|^2}. \tag{8.5}$$

We may generalize this operator to one that uses a set of N_{PB} domain (or *parent*) blocks instead of a single $\mathbf{d}_m$ to approximate $\mathbf{r}$ [2], as given by

$$\mathbf{r} \approx T_{N_{PB}}(\mathcal{D}_Z) = \beta + \sum_{m \in Z} \gamma_m \mathbf{d}_m. \tag{8.6}$$

[1]The quantity Δ is simply the root mean squared error between $\mathbf{r}(n)$ and its approximation, and is referred to in the fractal literature as the *collage error*.

The vector set $\mathcal{D}_Z \subset \mathcal{D}$ denotes a set of N_{PB} parents[2] chosen from the domain pool, $\mathcal{D}$. With a multi-parent approach, the parameters γ_m may be chosen to combine the domain blocks in a number of ways [1, 6, 16].

Relationship to autoregressive based methods

In this section, we make a connection between self-similar and AR equations to demonstrate certain properties of one-dimensional MR data. Combining Eq. 8.3 for range block $\mathbf{r}$ with Eq. 4.9, we obtain (see Appendix A.20 for derivation):

$$\mathbf{r} = -\sum_{m=1}^{2N_C} a_m \mathbf{d}_m. \tag{8.7}$$

Comparing Eqs. 8.6 and 8.7, several key observations can be made. First, by allowing $\beta = 0$ and using an appropriate set for $\mathcal{D}_Z$, we obtain

$$\alpha_m = -a_m, \quad m = 1, 2, 3, \ldots, 2N_C,$$

so that the approximation in Eq. 8.6 becomes an equality. That is, multi-parent operators acting on signals of the form Eq. 1.9 can produce $\mathbf{r}$ *exactly* without the need for a β term[3]. Thus, using the boxcar model [35], *one-dimensional MR signals are self similar* upon multiplication of the signal by $-2\pi ik$.

Having established a connection between self-similar model parameters and AR modeling, new results could be explored with AR modeling that describe mathematical and physical properties of the parameters α and β. This connection enables a relationship between, for example, α and physical parameters (in this case, x_m and Δx of Eq. 4.6) through Prony or AR modeling [50]. Deriving such connections is, however, beyond the scope of this chapter.

Eq. 8.7 represents a theoretical derivation for *one*-dimensional complex signals. Although mathematical extensions of Prony and AR methods to two dimensions have been made (see for example [3, 85]), our research into the extension of the mathematical connection between self-similar and AR theory to two dimensions is presented in the next chapter. The next section explores the *empirical* evidence that two-dimensional complex MRI Fourier data can demonstrate self-similarity.

[2]For example, the N_{PB} parents that have the lowest collage errors, according to Eq. 8.5 may be chosen.

[3]Admittedly, we have considered the noiseless case, and would not expect that the collage error would remain zero in the presence of noise

8.3 Two-Dimensional Self Similarity of MRI Data

In this section, we present a preliminary analysis of the self-similar properties of two-dimensional MRI data. We explore the self-similarity of MR data, using two-dimensional blocks, in the Fourier domain, employing single and multiple parent transforms, with MRI data taken from two different MR imaging systems.

Photographs of MR "phantoms", constructed with plexiglass sheets mounted on 2 cm diameter cylindrical Teflon bases are shown in Figs. 8.1(a) and 8.1(b). Figs. 8.1(c) and 8.1(d) show MR images[4] of these phantoms immersed in water. Fig 8.1(c) has dimensions 256×256, Fig. 8.1(d) has dimensions 512×512. Figs. 8.1(e), (f) display the respective k-space real components between [-2.5e-4,+2.5e-4]. Figs. 8.1(g) and (h) display the respective k-space magnitudes and the relative amplitudes at the origin and at high frequencies.

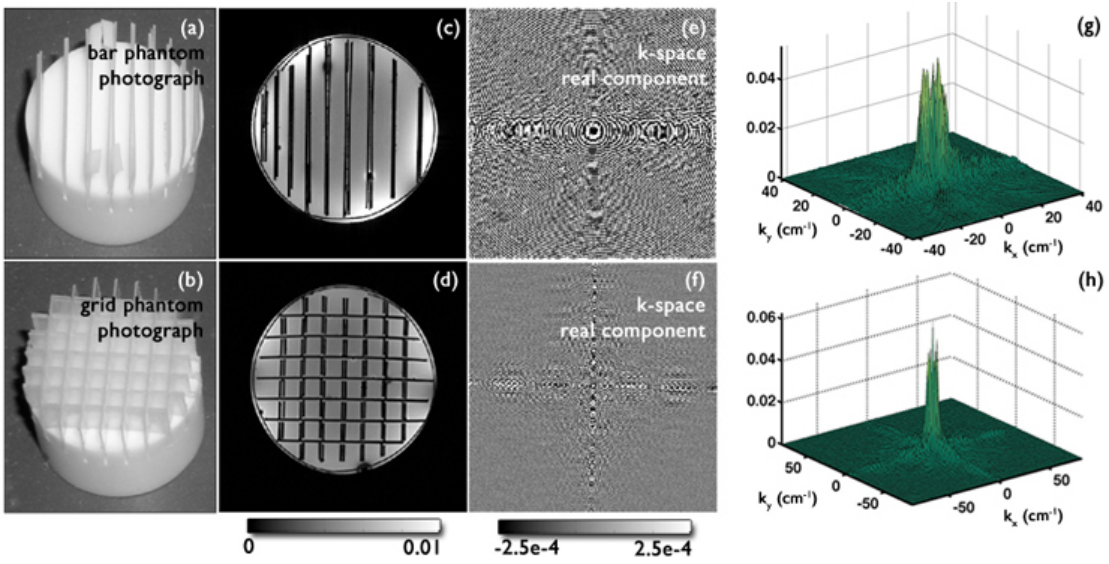


Figure 8.1: Photographs of the (a) "bar" and (b) "grid" phantoms. (c) and (d) display spatial domain magnitudes. (e) and (f) show the corresponding k-space real component between [-2.5e-4, 2.5e-4], and (g) and (h) show the k-space magnitude data. Data sets were normalized in l^2.

Fig. 8.2 shows data[5] from a 30 slice data set of a human volunteer. Each slice is a complex k-space data set with dimensions 128×128. The (a) spatial domain magnitude, (b) k-space real component, and (c) k-space magnitude are displayed. Each image was individually normalized by its respective l^2 norm.

Collage error probability histograms (CEPH) were constructed from the phantom

[4] 11.7 T MRI system (Bruker), using a gradient echo sequence, TR/TE 800/5 ms, FOV 3 cm.

[5] 3.0 T MRI system (General Electric Medical Systems; Waukesha, WI), using a multislice spoiled gradient-recalled echo sequence, FOV 24 cm, slick thickness 4 mm, TR/TE 200/3.1 ms, flip angle 18°.

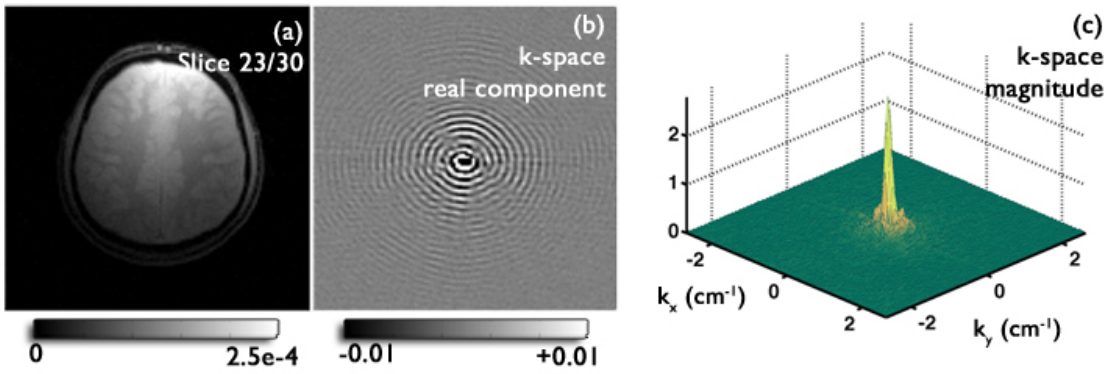


Figure 8.2: One slice from the 30 slice human volunteer data. (a) displays the magnitude of the spatial domain data, (b) the corresponding k-space real component intensities between -0.01 and +0.01, and (c) shows the k-space data magnitude.

data. The histograms were made by partitioning the k-space data sets into N_{DB} non-overlapping $N_P \times N_P$ *domain* and *range* blocks. The domain and range block sets were exactly the same. For each range block, all possible domain blocks were compared (with the exception of the domain block that was equal to the range block), using the 2D analog of Eq. 8.5, or *collage error*:

$$\Delta_{p,q} = \sqrt{\frac{1}{N_P^2} \sum_{m,n=1}^{N_P} |\mathbf{r}_p(m,n) - \alpha_q \mathbf{d}_q(m,n) - \beta_q|^2}, \;\; p,q = 1,2,\ldots,N_{DB}, \;\; p \neq q. \qquad (8.8)$$

A total of $N_{DB} \times (N_{DB} - 1)$ comparisons are made to construct a CEPH. The parameters α_q and β_q are determined using a least-squares fit to $\mathbf{r}_p$. CEPHs calculated using Eq. 8.8 from the *phantom* data are shown in Fig. 8.3 using $N_P = 8$.

To study the effect of noise on the distributions in our CEPHs, complex noise taken from a normal distribution with zero mean and SD 0.005 was added to the normalized phantom data. The corresponding CEPHs and SD histograms from the phantom data are displayed in Fig. 8.4

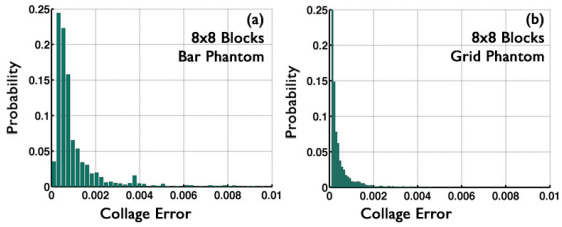


Figure 8.3: CEPHs from phantom data using 8×8 blocks.

CEPHs calculated using Eq. 8.8 from the *human volunteer* data are shown in Fig. 8.5 using $N_P = 4$. CEPHs with added noise and SD histograms are also presented in this figure.

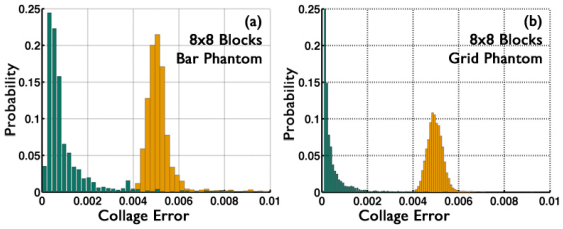


Figure 8.4: CEPHs from phantom data using 8×8 blocks (green), and the corresponding histograms with added complex zero mean noise with SD 0.005 (orange).

In all of the CEPHs (Figs. 8.4, 8.5, 8.6), the additive noise moved the distributions away from zero to a new distribution centered near 0.005. This effect is not surprising. Intuitively, it is generally more difficult to approximate noisy blocks, and we expect the collage errors to increase with increasing noise. We know that the CEPHs of *pure* noise is centered on its SD (see [28], page 5). We also see that the pure k-space data is more self-similar than the data obtained after adding noise, providing us with evidence that two-dimensional k-space data may be self-similar, or at least can be more self-similar with less noise.

The collage errors in Fig 8.5 were closer to the origin than the SDs. The SDs are simply the collage errors, using Eq. 8.8, with $\alpha_q = 0$. Including the α_q term reduces Δ, further suggesting that k-space data can be self-similar.

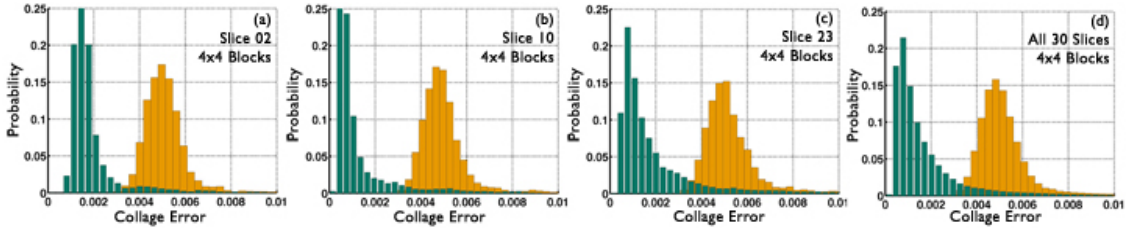


Figure 8.5: (a) to (c) human volunteer CEPHs (green) using 4×4 blocks and corresponding CEPHs after noise was added (orange). CEPHs in (d) were calculated from all 30 slices.

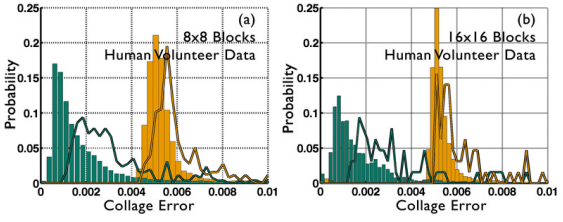


Figure 8.6: (a) and (b) human volunteer CEPHs (green) from all 30 slices using different block sizes, and the corresponding CEPHs after noise was added (orange). Green and orange lines correspond to range block SD histograms.

N_{PB}	Bar, $N_P = 8$	Grid, $N_P = 16$	Human Volunteer, $N_P = 8$
1	0.6179	0.8505	0.8300
30	0.2234	0.6221	0.3955
60	0.0435	0.5299	0.1028

Table 8.1: Mean normalized multi-parent collage errors using $N_{PB} = 1, 30$, and 60, from various data sets. Collage errors are normalized by the SD of $\mathbf{r}_p$. Only the N_{PB} blocks for each $\mathbf{r}_p$ with the lowest collage errors were used to calculate the least squares projection onto the N_{PB} blocks. Means are calculated across all range blocks.

CEPHs were also calculated using a *multi-parent* transform. The collage error between the range blocks and the least-squares projection of each range block onto the space spanned by the N_{PB} transformed domain blocks with the lowest collage errors for each range block was calculated. The CEPHs for $N_{PB} = 1, 30$, and 60 are shown in Fig. 8.7. The collage errors were normalized by the SDs of the range blocks, and the errors were plotted on a logarithmic scale. Table 1 shows the means of collage errors plotted in Fig. 8.7. As expected, the collage errors become smaller as more parents are used.

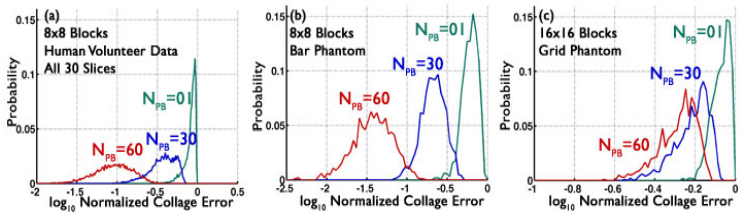


Figure 8.7: Normalized multi-parent CEPHs using $N_{PB} = 1, 30$, and 60, from various data sets. Errors are plotted on a $\log_{10}$ scale, and collage errors are normalized by the SD of $\mathbf{r}_p$. Only the N_{PB} blocks for each $\mathbf{r}_p$ with the lowest collage errors were used to calculate the least squares projection onto the N_{PB} blocks.

8.4 Conclusions

This chapter represents an important step in our research programme for frequency extrapolation of Fourier domain data using self-similar based methods. The author investigated the local self-similarity of two classes of data: 1D signals of the form Eq. 1.9, and 2D complex MRI k-space data.

Section 8.2 provided a mathematical investigation of the local self-similarity of signals of the form Eq. 1.9. Signals of this form were found to be self-similar, and partitioned blocks of the signal may be *predicted exactly with multi-parent transforms*; a consequence of Prony's method and AR modeling. One-dimensional noiseless MRI k-space signals can

be put into this form by using the boxcar model, thereby allowing the data to possess local self-similarity.

Section 8.3 provided an empirical study of the self-similarity of 2D complex Fourier data. Single and multi-parent transforms with complex model parameters were used to approximate Fourier domain blocks. Added noise decreased the self-similarity of the data, suggesting that self-similar methods can find and utilize self-similar structures present in raw Fourier data. Use of an α term and multi-parent transforms improved the collage error, suggesting that multi-parent based techniques are able to uncover k-space self-similarity. The use of domain block isometries (or geometric maps), overlapping domain blocks, and other standard fractal techniques [20] (modified for a same-scale affine transform) should further improve our results.

Establishing connections between self-similar and AR modeling theory provides new avenues for future investigation. Further relationships between self-similar, AR, and physical parameters can be made. Ultimately however, concepts established in this chapter are part of an overall research programme to use self-similarity to perform k-space extrapolation. Extension of mathematical results from Section 8.2 into two dimensions and use of multi-parent self-similar techniques explored in Section 8.3 may enable new approaches to the extrapolation of Fourier data, thereby providing a means of spatial domain resolution enhancement.

Chapter 9

Further Developments on the AR Modeling and Self-Similarity of MR Fourier Data

Results presented in the previous chapter represented original research the author recently published [67]. A number of recent experiments and extensions that have been made on the work presented in that chapter and are presented here. The work in this chapter *represents preliminary research.* It is hoped, however, that these developments will be useful for future researchers who may be interested in making further contributions to this area.

9.1 Prony's Method and the Separation of Exponentials Problem

This section provides a more detailed review of Prony's method than that provided in Section 4.1.3. The discussion here provides a basis for the material presented later in this chapter.

Let $s(n)$ represent a measured signal consisting of N_s evenly spaced samples, of the form

$$s(n) = \sum_{m=1}^{2\hat{N}_c} \hat{h}_m e^{-\hat{\sigma}_m n},\ \hat{h}_m, \hat{\sigma}_m \in \mathbb{C}, n = 0, 1, 2, \ldots, N_s - 1. \tag{9.1}$$

We assume the parameters $\hat{\sigma}_m$ are distinct, and that $\hat{h}_m \neq 0\ \forall\ m \in [1, 2\hat{N}_c]$. Prony's method [83] attempts to estimate the unknown parameters $\hat{h}_m, \hat{\sigma}_m$, and $\hat{N}_c$. The method

fits a discrete model[1]

$$s_{N_c}(n) = \sum_{m=1}^{2N_c} h_m e^{-\sigma_m n},\ h_m, \sigma_m \in \mathbb{C},\ n = 0, 1, 2, \ldots, N_s - 1, \tag{9.2}$$

to the observed $s(n)$ to find parameters N_c, h_m and σ_m that best describe the measured data, and so that $s(n) = s_{N_c}(n)$.

Prony's method first derives a linear prediction equation,

$$s(n) = -\sum_{l=1}^{2N_c} a_l s(n-l),\ n = 2N_c, 2N_c + 1, 2N_c + 2, \ldots, N_s - 1, \tag{9.3}$$

based on Eq. 9.2 (see Appendix A.21 for derivation). The variables $\{a_l\}$ are referred to as the *auto-regressive* (AR) parameters, and are estimated by solving

$$\begin{aligned}
\mathbf{r} &= \begin{bmatrix} s(2N_c) \\ s(2N_c+1) \\ \vdots \\ s(2N_c+N_P-1) \end{bmatrix} \\
&= \begin{bmatrix} s(2N_c-1) & s(2N_c-2) & \ldots & s(0) \\ s(2N_c) & s(2N_c-1) & \ldots & s(1) \\ \vdots & \vdots & \ddots & \vdots \\ s(2N_c+N_P-2) & s(2N_c+N_P-3) & \ldots & s(N_P-1) \end{bmatrix} \begin{bmatrix} -a_1 \\ -a_2 \\ \vdots \\ -a_{2N_c} \end{bmatrix} \\
&= \mathbf{D}\mathbf{a},
\end{aligned} \tag{9.4}$$

which follows directly from Eq. 9.3. The parameter N_P can be chosen so that $1 \leq N_P \leq N_s - 1$, and we see that $\mathbf{D} \in \mathbb{C}^{N_P \times 2N_c}$

It can be shown [49] that the row rank of $\mathbf{D}$ is $2N_c$, so long as $\mathbf{D}$ has at least $2N_c$ rows. Otherwise, $\mathbf{D}$ will have rank N_P. Therefore, choosing $N_P < 2N_c$, the above system does not have a unique solution. If we choose $N_P = 2N_c$, $\mathbf{D}$ will have full rank, and there will be a unique solution to Eq. 9.4. Choosing $2N_c < N_P \leq N_s - 1$, the system will be overdetermined, but a minimum norm least squares solution may be found.

Once the parameters a_m have been estimated, σ_m may be found by finding the zeros, $z_m \in \mathbb{C}$, of the polynomial (see Eq. A.69):

$$p(z) = \sum_{m=0}^{2N_c} a_m z^m,\ z \in \mathbb{C},\ a_0 = 1. \tag{9.5}$$

The zeros, $z_m \in \mathbb{C}$, provide the exponential parameters via $z_m = e^{-\sigma_m}$, so that

$$\sigma_m = -\ln z_m,\ m = 1, 2, 3, \ldots, 2N_c - 1. \tag{9.6}$$

[1]The negative signs in the exponents of Eq. 9.1 and 9.2 are simply included as a matter of convenience for our derivations in the appendices.

Finally, the parameters h_m are found by solving

$$\begin{bmatrix} s(N_S-1) \\ s(N_S-2) \\ \vdots \\ s(1) \\ s(0) \end{bmatrix} = \begin{bmatrix} z_1^{2N_S-1} & z_2^{2N_S-1} & \dots & z_{2N_c}^{2N_S-1} \\ z_1^{2N_S-2} & z_2^{2N_S-2} & \dots & z_{2N_c}^{2N_S-2} \\ \vdots & \vdots & \ddots & \vdots \\ z_1 & z_2 & \dots & z_{2N_c} \\ 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_{2N_c-1} \\ h_{2N_c} \end{bmatrix}. \tag{9.7}$$

This system follows from Eq. 9.2, and the change of variable $z_m = e^{-\sigma_m}$.

A one-dimensional frequency extrapolation method could use the linear prediction equation (Eq. 9.3). The remaining steps of the Prony method (Eqs. 9.5 to 9.7), which are useful for parameter identification problems, are (thus far) not needed for the analysis in this chapter. These steps were merely included to provide a more complete review of the Prony method.

Before developing connections between this method and MRI reconstruction, it should also be noted that the separation of exponentials problem remains an open ended research area, and has relevance in the NMR and MRI literature (see for example [4, 25]). The interested reader is also referred to a beautifully written thesis [52] recently completed at the University of Waterloo on the separation of exponential problems related to NMR data. Prony's method for solving separation of exponential problems is well described in [97], page 1404 of [45], and in a textbook by Lanczos [51].

9.2 Further Developments of Prony Modeling of MR Data

This section provides *recent* results on the Prony modeling of MR data. Section 4.1.3 developed a one-dimensional piecewise constant (boxcar) model for the proton density of the object being imaged. We first provide two extensions of this model: a one-dimensional piecewise linear model, and a two-dimensional piecewise constant model.

9.2.1 1D Piecewise Linear Model

In the piecewise *linear* model the proton density can be expressed as:

$$\rho_{N_c}(x) = \sum_{m=1}^{N_c} c_m \Lambda\left(\frac{x - x_m}{\Delta x}\right), \quad \Lambda(x) = \begin{cases} 1+x, & -1 \le x \le 0 \\ 1-x, & 0 \le x \le 1 \\ 0, & \text{else} \end{cases}, \tag{9.8}$$

and $x_m = (m - 1 - N_c/2)\Delta x$, $m = 1, 2, 3, \dots, N_c$. The proton density, $\rho_{N_c}(x)$, in this case is represented by a sum of overlapping triangle functions, each centered at x_m, with base width $2\Delta x$ (Fig. 9.1).

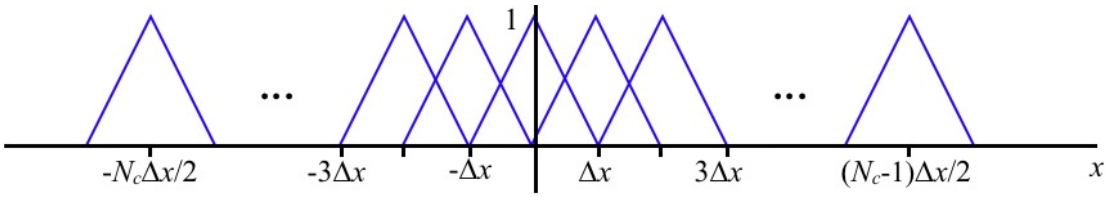


Figure 9.1: The proton density in Eq. 9.8 is expressed as a sum of N_c triangle functions, each centred at $x_m = (m - 1 - N_c/2)\Delta x$.

Fourier Transformation of the 1D piecewise linear model leads to[2]:

$$\begin{aligned}
\mathsf{P}_{N_c}(k_x) &= \int \rho_{N_c}(x) e^{-2\pi i k_x x} dx \\
&= \int \sum_{m=1}^{N_c} c_m \Lambda\left(\frac{x - x_m}{\Delta x}\right) e^{-2\pi i k_x x} dx \\
&= \sum_{m=1}^{N_c} c_m e^{-2\pi i x_m k_x} \int \Lambda\left(\frac{x}{\Delta x}\right) e^{-2\pi i k_x x} dx \\
&= \mathrm{sinc}^2(\Delta x k_x) \Delta x \sum_{m=1}^{N_c} c_m e^{-2\pi i x_m k_x}.
\end{aligned} \tag{9.9}$$

This expression may be put in the form (see Appendix A.22):

$$\mathsf{P}_{N_c}(k_x) = \frac{-1}{4\Delta x \pi^2 k_x^2} \sum_{m=1}^{3N_c} d_m e^{i\phi_m k_x}, \ d_m \in \mathbb{C}, \ \phi_m \in \mathbb{R}. \tag{9.10}$$

Multiplication by $-4\Delta x \pi^2 k_x^2$, and sampling this function N_S times with a regular sample spacing yields

$$\begin{aligned}
s^{(1)}_{N_c}(n) &= -4\Delta x \pi^2 n^2 \mathsf{P}_{N_c}(n) \\
&= \sum_{m=1}^{3N_c} d_m e^{i\phi_m n}, \ n = 0, 1, 2, \ldots, N_S - 1.
\end{aligned} \tag{9.11}$$

The superscript (1) is used to distinguish this signal from s_{N_c}, which was derived with a piecewise *constant* model in Section 4.1.3. Parameter sets $\{d_m\}$ and $\{\phi_m\}$ are defined in Appendix A.22.

It can be shown (with a similar derivation as that provided in Appendix A.21) that

$$s^{(1)}_{N_c}(n) = -\sum_{l=1}^{3N_c} a_l s^{(1)}_{N_c}(n - l), \ n = 2N_c, 2N_c + 1, 2N_c + 2, \ldots, N_s - 1. \tag{9.12}$$

Comparing this equation to Eq. 4.9, which was derived using a piecewise *constant* model, two immediate observations can be made. We see that an additional N_c prediction

[2]The Fourier Transform of $\Lambda(x) = \Pi(x) * \Pi(x)$, is $\mathrm{sinc}^2(k_x)$, by the convolution theorem.

coefficients are needed to construct the linear prediction relation. Additionally, to put this equation into a discrete, complex trigonometric polynomial form, we multiplied $\mathsf{P}_{N_c}(n)$ by a *quadratic* function, instead of a linear function. As the SNR of MRI Fourier data decreases at higher frequencies, we expect this piecewise linear model to be more sensitive to noise. As simple as this analysis is, it has, to the author's best knowledge, not been presented in the MRI literature.

9.2.2 2D Piecewise Constant Model

The two-dimensional piecewise constant model, expresses the proton density of the data as a sum of piecewise constant functions[3]:

$$\rho_{N_c}(x,y) = \sum_{m_1=1}^{N_c}\sum_{m_2=1}^{N_c} c_{m_1,m_2}\Pi\left(\frac{x-x_{m_1}}{\Delta x}\right)\Pi\left(\frac{y-y_{m_2}}{\Delta y}\right). \tag{9.13}$$

For simplicity, we assume $x_{m_1} = (m_1 - N_c/2)\Delta x$, and $y_{m_2} = (m_2 - N_c/2)\Delta y$. The 2D Fourier Transform of this model yields:

$$\begin{aligned}
\mathsf{P}_{N_c}(k_x,k_y) &= \int\int \sum_{m_1,m_2=1}^{N_c} c_{m_1,m_2}\Pi\left(\frac{x-x_{m_1}}{\Delta x}\right)\Pi\left(\frac{y-y_{m_2}}{\Delta y}\right)e^{-2\pi i(k_x x+k_y y)}dxdy \\
&= \sum_{m_1,m_2=1}^{N_c} c_{m_1,m_2}\int \Pi\left(\frac{x-x_{m_1}}{\Delta x}\right)e^{-2\pi i k_x x}dx \int \Pi\left(\frac{y-y_{m_2}}{\Delta y}\right)e^{-2\pi i k_y y}dy \\
&= \Delta x\Delta y \mathrm{sinc}(\Delta x k_x)\mathrm{sinc}(\Delta y k_y)\sum_{m_1=1}^{N_c}\sum_{m_2=1}^{N_c} c_{m_1,m_2}e^{-2\pi i(x_{m_1}k_x+y_{m_2}k_y)}.
\end{aligned}$$

Using the definition of the sinc function, this equation becomes

$$\mathsf{P}_{N_c}(k_x,k_y) = \frac{\sin(\Delta x\pi k_x)\sin(\Delta y\pi k_y)}{\pi^2 k_x k_y}\sum_{m_1=1}^{N_c}\sum_{m_2=1}^{N_c} c_{m_1,m_2}e^{-2\pi i(x_{m_1}k_x+y_{m_2}k_y)}. \tag{9.14}$$

Multiplication by $-4\pi^2 k_x k_y$ yields

$$\begin{aligned}
s_{N_c}(k_x,k_y) &= -4\pi^2 k_x k_y \mathsf{P}_{N_c}(k_x,k_y) \\
&= -4\sin(\Delta x\pi k_x)\sin(\Delta y\pi k_y)\sum_{m_1=1}^{N_c}\sum_{m_2=1}^{N_c} c_{m_1,m_2}e^{-2\pi i(x_{m_1}k_x+y_{m_2}k_y)}.
\end{aligned} \tag{9.15}$$

The factor of -4 is removed in the following step, where $s_{N_c}(k_x,k_y)$ is expressed as a sum of complex exponentials (see Appendix A.23 for derivation):

$$s_{N_c}(k_x,k_y) = \sum_{m_1=1}^{2N_c}\sum_{m_2=1}^{2N_c} d_{m_1,m_2}e^{i\zeta_{m_1}k_x}e^{i\xi_{m_2}k_y}. \tag{9.16}$$

[3]This is a two-dimensional version of Eq. 4.6.

Sampling this function N_S times along the k_x and N_S times along the k_y direction yields

$$\begin{aligned} s_{N_c}(m,n) &= -4\pi^2 mn \mathsf{P}_{N_c}(m,n) \\ &= \sum_{m_1=1}^{2N_c} \sum_{m_2=1}^{2N_c} d_{m_1,m_2} e^{i\zeta_{m_1} m} e^{i\xi_{m_2} n}, \end{aligned} \tag{9.17}$$

where $m, n = 0, 1, 2, \ldots, N_S - 1$. It has been shown [3] that signals of this form have a two-dimensional linear prediction equation[4]:

$$s_{N_c}(m,n) = -\sum_{p=1}^{p_1} \sum_{q=1}^{q_1} a_{p,q} s_{N_c}(m-p, n-q). \tag{9.18}$$

This equation is valid for $m = p_1, p_1 + 1, \ldots, N_S$, and $n = q_1, q_1 + 1, \ldots, N_S$.

Eq. 9.18 suggests that a two-dimensional linear prediction equation may be used to extrapolate MRI data. Previous works on data extrapolation and reconstruction in MRI that use ARMA models have only used one-dimensional models which are applied on each row (or column) of the frequency data matrix separately. We may find that a fully 2D approach would perform better than methods that have been previously implemented in the literature. Eq. 9.18 also suggests that, as we had for the one-dimensional case, the frequency data possess a natural self-similarity that may be used for image processing procedures. Indeed, a number of future research directions are suggested by Eq. 9.18, a partial list of which is provided in Section 10.3.2.

[4]The proof of this result presented in [3] essentially follows that presented in [97] and Appendix A.21 for the one-dimensional case.

Chapter 10

Conclusions

This thesis has investigated three broad research areas under the common underlying theme of spatial resolution enhancement in MRI in the Fourier domain. This chapter simply reviews the contributions made to each of these areas and offers possible directions for future research.

10.1 MASRMRI

Chapter 5 explored the use of SRMRI with multiple acquisitions related by spatial translations in the frequency encoding (FE) direction. Prior to the author's publication [66] of the research presented in Chapter 5, there existed some debate in the literature as to the efficacy of this approach. This thesis does not claim that SR with multiple acquisitions related by spatial translations in this specific encoding direction is *impossible*. However, conclusions reached in Chapter 5 state that the amount of new information after the first is relatively small, and can be difficult to detect in the presence of noise. Furthermore, most of the new information can be related to an increase in the field-of-view, and not to information at higher frequencies beyond what is measured.

10.1.1 Contributions

1. Conclusions reached in Chapter 5 help explain why SRMRI experiments in the FE direction have not met with success [11], and suggests that this particular approach may not be viable for resolution enhancement in MRI. It is hoped that these conclusions will point future researchers working on resolution enhancement in MRI towards more promising research directions that are more likely to make significant contributions to SRMRI.

2. The work presented in Chapter 5 may help emphasize to future researchers in MR reconstruction the importance of having an accurate physical model of the data that is being used. As pointed out throughout Chapter 5, much of the work carried out on MASRMRI in the FE and PE directions by various research groups ignored key signal processing steps of the acquisition process. The means by which measured MRI data is acquired and subsequently processed must be understood before any assumptions can be made as to its nature. Particular details relating to the acquisition process can determine the efficacy of a particular processing technique.

3. Results presented in Chapter 5 were published in the LNCS series [63], in the European Society for Magnetic Resonance in Medicine and Biology 2006 conference proceedings [64], and in MRI [66].

10.1.2 Future Work

There are at least two possible avenues for further research in SRMRI with multiple acquisitions.

1. Further research in MASRMRI could improve the efficacy of techniques implemented along the SE direction. One possible area of research for this would be to investigate the use of known information about the pulse sequence used. Incorporating the PSF, which can be determined from pulse sequence parameters (see discussion in Section 3.2.3), may help the image reconstruction in such an experiment. Published research in this area [18, 31, 32, 33, 46, 47, 71, 79, 80] has only made use of a PSF that was assumed to be given by $\Pi(x)$ or by a Gaussian function, with no physical basis for this choice.

2. Another MASRMRI technique that has been recently published [75] implements multiple translations using multiple radiofrequency coils. Although initial experimental results look promising, it is not immediately obvious why this approach would yield images with higher spatial resolution. Indeed, new information is introduced in each acquisition by extending the FOV. However, as with SR in the FE and PE directions, information is only recorded over the same frequency band in each image. An investigation, similar to what is presented in Chapter 5 (and published in MRI [66]) exploring the nature of the new information in each acquisition could be considered[1].

[1] Without access to MRI data acquired with multiple RF coils, perhaps a simple letter to the editor could be submitted, as was the case for the FE and PE SR models [82, 90].

10.2 The IFTS Operator

A fractal-based, conditionally-contractive IFTS operator over complex-valued functions was derived in Chapter 7. This operator was derived from the well-established, spatial domain IFSM operator. Properties of the IFTS operator were described, and the operator was used to extrapolate simulated data for the purpose of spatial resolution enhancement.

10.2.1 Contributions

1. The IFTS operator introduced in Chapter 7 was the first of its kind, in that most work carried out on fractal operators have worked in either the spatial or wavelet domains [20, 73]. The IFTS operator works directly on complex data, and is derived from the IFSM operator.

2. Several properties of the IFTS operator, listed in Section 7.2, were introduced. Conditions for this operator to be contractive were provided.

3. The work presented in Chapter 7 established several relationships between the Fourier domain IFTS and spatial domain IFSM operator from which it was derived:

   (a) The spatial domain IFSM operator expresses a function as a sum of scaled, shifted, spatially-*contracted* copies of itself. The IFTS operator expresses a Fourier domain function as a sum of scaled, modulated, frequency-*expanded* copies of itself[2].

   (b) The fixed point of the IFTS operator is the Fourier Transform of the fixed point of the corresponding IFSM operator.

   (c) The IFTS and IFSM operators posses the same Lipschitz constant in the L^2 metric.

4. Work presented in Chapter 7 was published in the LNCS series [65].

10.2.2 Future Work

The work carried out in Chapter 7 could be extended to two dimensions. However, comparing the work presented in Chapters 7 and 8, there is currently no connection to the physics of the MRI data acquisition process that would suggest that the IFTS operator could be used for frequency extrapolation. There is, however, a physical basis for linear prediction and self-similarity for frequency extrapolation of MR data that was described in Section 8.2.

[2]This is a consequence of the Fourier scaling theorem, Table 2.2.

Any future work on the IFTS operator should connect this operator to a physical model of measured data to provide an explanation for why it could be used for frequency extrapolation (or any other application in signal or image processing). Until then, there may be more value in pursuing further research on the material presented in Chapters 8 and 9, based, perhaps, on recommendations made in Section 10.3.2.

10.3 Self-Similarity and ARMA Modeling in SRMRI

Chapter 8 explored the self-similarity of a 1D physical model of MRI data, as well as the self-similarity of measured 2D Fourier data. Results presented in Chapter 8 were developed further in Chapter 9, where a 1D piecewise linear and a 2D piecewise constant model were considered. In each case, it was shown that these models lead to linear prediction relations in k-space.

10.3.1 Contributions

1. A connection was established between 1D ARMA methods and self-similar, multi-parent methods (Eq. 8.7). This connection suggests that a physical justification for the use of multi-parent methods in k-space exists.
2. The 1D and 2D linear prediction models suggest that frequency extrapolation could be obtained using self-similarity. Chapter 8 established a model in 2D that needs to be further explored to implement a fully 2D algorithm for extrapolation.
3. Methods for measuring and displaying the self-similarity of k-space MRI data were explored. It was found that some degree of self-similarity of this data can be found using multi-parent operators.
4. Work presented in Chapter 8 was published in the LNCS series [67]. Work presented in Chapter 9 remains unpublished, but may be helpful for future researchers working in SRMRI.

10.3.2 Future Work

Many directions for future research were discovered that are based upon the results presented in Chapter 8. As the research results therein do not explore the extrapolation of MR k-space data, the following items outline a set of research goals that may eventually lead towards a viable approach to spatial resolution enhancement.

1. The most immediate step in furthering the research results in Chapters 8 and 9, would be a development of a method to determine the parameters for the 2D AR model of Eq. 9.18.

2. A study of the stability of 2D linear prediction relations would be essential for developing practical extrapolation procedures. Any algorithm that determines AR coefficients would ideally do so in a way that could be used for a *stable* 2D extrapolation.

3. Once a method for determining parameters for the 2D AR model has been established (item 1 of this list), numerical simulations in two dimensions should be performed. These simulations would test methods of determining AR parameters and using these parameters for frequency extrapolation.

4. Provided a stable, and reasonably accurate implementation of a 2D algorithm can be obtained, this technique could be implemented on MRI phantom and human brain data.

There are also other research directions that can be taken, not considered in this thesis, that do not necessarily relate to frequency extrapolation.

1. Further experiments involving the CEPH, presented in Section 8.3 could be explored. The filter, $-4\pi^2 k_x k_y$, presented in Eq. 9.15, when applied to the 2D model of the proton density allowed the data to be linearly predictable. It may be easy to show, as in the 1D case, $s_{N_c}(k_x, k_y)$ is self-similar.

2. A more precise connection between self-similar and physical parameters could be made via Prony modeling. This would provide a connection between self-similar methods and parameter estimation, which may be helpful for proving theorems or uncovering new applications of self-similarity in image processing.

3. Research on the use of self-similarity for denoising MRI data could be explored. The use of multi-parent transforms for denoising spatial domain data has been studied [28]. The combination of linear prediction relations and self-similar methods may yield interesting results for the denoising of k-space data.

Appendix A

Derivations

A.1 Properties of the Discrete Extrapolation Problem

The purpose of this appendix is to provide a brief description of the discrete extrapolation problem for Section 1.2, and to show that this problem is ill-posed. Indeed, an important concept in the discrete spatially-limited (SL) extrapolation problem is that the solution is not unique[1]. To demonstrate this property, we require the use of two matrices that will model frequency limiting and spatial limiting operators on frequency domain discrete signals. Consider the matrix $\mathrm{R} = \{r_{i,j}\}$, where $i, j \in \mathbb{Z}$, and

$$r_{i,j} = \frac{\sin(\pi(i-j)R)}{\pi(i-j)}, \qquad R \in \mathbb{R}.$$

Matrix R is an infinite Toeplitz matrix and is a spatially limiting operator on discrete frequency domain signals.

Discrete frequency domain truncation is modeled with the matrix $\mathrm{K} = \{k_{i,j}\}$, with elements

$$k_{i,j} = \begin{cases} 1 & i = j, \quad -N \leq i \leq +N \\ 0 & \text{otherwise} \end{cases}, \qquad N \in \mathbb{Z}^+.$$

The matrix K has dimensions $(2N+1) \times \infty$.

Now let us consider a continuous, SL signal, $s_{true}(k)$, where $k \in \mathbb{R}$. This signal is sampled at regular intervals to produce $s_d(n)$, where $n \in \mathbb{Z}$. Application of matrix K to the infinite vector $s_d(n)$ yields the truncated signal, $s_{d,N}(n)$:

$$s_{d,N}(n) = \mathrm{K} s_d(n), \qquad n \in \mathbb{Z}. \tag{A.1}$$

That is, $s_{d,N}$ is a vector of infinite length, whose central $2N+1$ points are equal to $s(n)$, and elsewhere the elements of s_m are zero.

[1]The discussion in this appendix follows that presented in [42].

The measurment of a SL discrete signal $s_{d,N}(n)$, $n \in \mathbb{Z}$, is modelled with Eq. A.1. However, since s_d is spatially limited, we also have that $s_d = \mathrm{R}s_d$:

$$s_{d,N} = \mathrm{KR}s_d, \tag{A.2}$$

where KR is a rectangular matrix with dimensions $(2N+1) \times \infty$. Therefore, KR has rank of at most $2N + 1$. Eq. A.2 represents an underdetermined system of equations. The solution to this system, and hence the solution to the discrete extrapolation problem, is non-unique.

A.2 Sampling of The Measured Signal, $F(k)$

The purpose of this appendix is to prove Eq. 2.8. The sampled Fourier domain signal, $\hat{F}(k)$, can be expressed as, for all $k \in \mathbb{R}$:

$$\begin{aligned}
\hat{F}(k) &= F(k)\mathrm{III}(\frac{k}{\Delta k}) \\
&= F(k) \sum_{n=-\infty}^{+\infty} \delta\big(\frac{k}{\Delta k} - n\big) \\
&= F(k) \sum_{n=-\infty}^{+\infty} \delta\big((k - n\Delta k)\frac{1}{\Delta k}\big).
\end{aligned}$$

However, because of the scaling property of the Dirac delta function [38], $\delta(k/\Delta k) = \Delta k \delta(k)$, we may write this as:

$$\begin{aligned}
\hat{F}(k) &= F(k) \sum_{n=-\infty}^{+\infty} \Delta k \delta(k - n\Delta k) \\
&= \Delta k \sum_{n=-\infty}^{+\infty} F(n\Delta k)\delta(k - n\Delta k).
\end{aligned} \tag{A.3}$$

We have completed our proof of Eq. 2.8.

A.3 IFT of Infinitely Sampled Frequency Data

The purpose of this appendix is simply to prove Eq. 2.11, which is the inverse Fourier transform of a discretized function in frequency space, defined over the entire real line:

$$\begin{aligned}
\tilde{f}(x) &= \int_{-\infty}^{\infty} \tilde{F}(k) e^{i2\pi kx} dk, \qquad \forall\ x \in \mathbb{R},\ k \in \mathbb{R}, \\
&= \int_{-\infty}^{\infty} \mathrm{III}\Big(\frac{k}{\Delta k}\Big)\Pi\Big(\frac{k}{2k_{max}}\Big) F(k) e^{i2\pi kx} dk \\
&= \int_{-\infty}^{\infty} \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} F(k)\delta(\frac{k}{\Delta k} - n) e^{i2\pi kx} dk \\
&= \int_{-\infty}^{\infty} \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} F(k)\delta\big((k - n\Delta k)\frac{1}{\Delta k}\big) e^{i2\pi kx} dk \\
&= \Delta k \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} F(n\Delta k) e^{i2\pi n\Delta kx} \qquad \text{(sifting theorem).}
\end{aligned} \tag{A.4}$$

Eq. A.4 is Eq. 2.11, thereby completing our proof.

A.4 Derivation of Precession with Quantum Mechanical Principles

The purpose of this Appendix is to derive Eq. 3.19. Using the Schrödiner solution for a proton, (Eq. 3.17), the y-component of the expectation value of $\vec{\mu}$ is:

$$\begin{aligned}
< \mu_y(t) > &=< \Psi|\mu_y(t)|\Psi > \\
&= \int \Big(\sum_{m=\pm 1/2} C_m \psi_m \mathrm{e}^{-iE_m t/\hbar} \Big)^* \gamma J_y \Big(\sum_{m=\pm 1/2} C_m \psi_m \mathrm{e}^{-iE_m t/\hbar} \Big) dV \\
&= \frac{\gamma\hbar V}{2} \Big(\sum_{m=\pm 1/2} C_m \psi_m \mathrm{e}^{-iE_m t/\hbar} \Big)^* \sigma_y \Big(\sum_{m=\pm 1/2} C_m \psi_m \mathrm{e}^{-iE_m t/\hbar} \Big) \\
&= \frac{\gamma\hbar V}{2} \Big[+ iC_-^* \psi_m \mathrm{e}^{+iE_- t/\hbar}, -iC_+^* \psi_m \mathrm{e}^{+iE_+ t/\hbar} \Big] \sum_{m=\pm 1/2} C_m \psi_m \mathrm{e}^{-iE_m t/\hbar} \\
&= \frac{\gamma\hbar V}{2} \Big[- iC_+^* C_- e^{-i(E_- - E_+)t/\hbar} + iC_-^* C_+ e^{+i(E_- - E_+)t/\hbar} \Big],
\end{aligned} \tag{A.5}$$

where $C_+ = C_{+1/2}$, $C_- = C_{-1/2}$, $E_+ = E_{+1/2}$, $E_- = E_{-1/2}$. Using the relation $E_- - E_+ = \hbar\omega_0$ (Eq. 3.9), we find that

$$\begin{aligned}
< \mu_y(t) > &= \frac{\hbar\gamma V}{2} [-iC_+^* C_- e^{-i\omega_0 t} + iC_-^* C_+ e^{+i\omega_0 t}] \\
&= \hbar\gamma V \mathrm{Real}\big\{ iC_-^* C_+ e^{+i\omega_0 t} \big\}.
\end{aligned} \tag{A.6}$$

For convenience, we introduce the following polar coordinate representation:

$$C_+ = a_+ e^{i\alpha_+}, C_- = a_- e^{i\alpha_-}, \tag{A.7}$$

where $a_\pm$ is the magnitude of $C_\pm$. The normalization condition on the total probability implies a condition on the complex coefficients, C_m:

$$\begin{aligned} 1 &= < \Psi | \Psi > \\ &= \int \Psi^\dagger \Psi dV \\ &= \Psi^\dagger \Psi V \\ &= V \sum_m \sum_n C_n^* C_m \psi_n^\dagger \psi_m e^{\frac{i}{\hbar}(E_n - E_m)t}, \end{aligned} \tag{A.8}$$

which leads to:

$$V \sum_m |C_m|^2 = 1. \tag{A.9}$$

A combination of Eqs. A.9 and A.7 leads to:

$$1 = V(a_+^2 + a_-^2), \tag{A.10}$$

which is satisfied by $a_+ = \frac{1}{\sqrt{V}} \cos\theta$, and $a_- = \frac{1}{\sqrt{V}} \sin\theta$, for some fixed $\theta \in [0, 2\pi)$. Hence, Eq. A.6 becomes:

$$\begin{aligned} < \mu_y(t) > &= \hbar\gamma V \text{Real}\big\{ i a_- a_+ e^{-i\alpha_- + i\alpha_+ + i\omega_0 t} \big\} \\ &= \hbar\gamma \text{Real}\big\{ i \sin\theta \cos\theta e^{-i\alpha_- + i\alpha_+ + i\omega_0 t} \big\} \\ &= \gamma\hbar \sin\theta \cos\theta \big\{ -\sin(-\alpha_- + \alpha_+ + \omega_0 t) \big\} \\ &= \frac{\gamma\hbar}{2} \sin 2\theta \sin(\alpha_- - \alpha_+ - \omega_0 t) \\ &= \frac{\gamma\hbar}{2} \sin 2\theta \sin(\phi_0 - \omega_0 t), \end{aligned} \tag{A.11}$$

where $\phi_0 = \alpha_- - \alpha_+$. We have completed our proof of Eq. 3.19.

A.5 Derivation of the z-Component of the Moment Expectation in a Time Varying RF Field

The purpose of this appendix is to derive Eq. 3.26. Using Eqs. 3.6 and 3.7, the Hamiltonian for a proton at rest in the given RF field (Eq. 3.21) is:

$$\begin{aligned} H(t) &= -\vec{\mu} \cdot \vec{B}(t) \\ &= -\gamma \vec{S} \cdot \vec{B}(t) \\ &= -\frac{\gamma\hbar}{2} \vec{\sigma} \cdot \vec{B}(t) \\ &= -\frac{\gamma\hbar}{2} \big[\sigma_z B_0 + (\sigma_x \cos\omega_{rf} t) B_1 - (\sigma_y \sin\omega_{rf} t) B_1 \big]. \end{aligned} \tag{A.12}$$

The Schrödinger equation (Eq. 3.16) becomes:

$$\begin{aligned} i\hbar\frac{\partial\Psi}{\partial t} &= H\Psi \\ &= -\frac{\gamma\hbar}{2}\big[\sigma_z B_0 + (\sigma_x\cos\omega_{rf}t)B_1 - (\sigma_y\sin\omega_{rf}t)B_1\big]\Psi \\ &= -\frac{\gamma\hbar}{2}\left[\begin{pmatrix} B_0 & 0 \\ 0 & -B_0 \end{pmatrix} + \begin{pmatrix} 0 & B_1\cos\omega_{rf}t \\ B_1\cos\omega_{rf}t & 0 \end{pmatrix}\right. \\ &\quad \left. - \begin{pmatrix} 0 & -iB_1\sin\omega_{rf}t \\ iB_1\sin\omega_{rf}t & 0 \end{pmatrix}\right]\Psi \\ &= -\frac{\hbar}{2}\begin{pmatrix} \omega_0 & \omega_1 e^{+i\omega_{rf}t} \\ \omega_1 e^{-i\omega_{rf}t} & \omega_0 \end{pmatrix}\Psi. \end{aligned} \tag{A.13}$$

The constant $\omega_{=}\gamma B_1$ is a parameter set by the scanner operator. We assume a general form for the wave function:

$$\Psi = \begin{pmatrix} \psi_1(t)e^{+i\omega_0 t/2} \\ \psi_2(t)e^{-i\omega_0 t/2} \end{pmatrix}. \tag{A.14}$$

We may set the acquisition parameter, ω_{rf}, equal to the hydrogen precession frequency for the given applied magnetic field, ω_0. Substitution of Eq. A.14 into Eq. A.13 yields the following set of coupled differential equations:

$$\frac{d\psi_1}{dt} = \frac{i\omega_1}{2}\psi_2(t) \tag{A.15}$$

$$\frac{d\psi_2}{dt} = \frac{i\omega_1}{2}\psi_1(t). \tag{A.16}$$

Solutions to these equations may be found through differentiation:

$$\frac{d^2\psi_1}{dt^2} = \frac{i\omega_1}{2}\frac{d\psi_2}{dt} = -\frac{\omega_1^2}{4}\psi_1(t) \tag{A.17}$$

$$\frac{d^2\psi_2}{dt^2} = \frac{i\omega_1}{2}\frac{d\psi_1}{dt} = -\frac{\omega_1^2}{4}\psi_2(t). \tag{A.18}$$

These equations yield general solutions for the wave functions:

$$\psi_1(t) = c_1\cos\frac{\omega_1 t}{2} + c_2\sin\frac{\omega_1 t}{2} \tag{A.19}$$

$$\psi_2(t) = c_3\cos\frac{\omega_1 t}{2} + c_4\sin\frac{\omega_1 t}{2}. \tag{A.20}$$

To find the complex coefficients, we may substitute these solutions into Eq. A.15 and Eq. A.16:

$$\begin{aligned} \frac{d\psi_1}{dt} &= \frac{d}{dt}\big(c_1\cos\frac{\omega_1 t}{2} + c_2\sin\frac{\omega_1 t}{2}\big) \\ &= \frac{\omega_1}{2}\big(-c_1\sin\frac{\omega_1 t}{2} + c_2\cos\frac{\omega_1 t}{2}\big) \\ &= \frac{i\omega_1}{2}\big(-ic_2\cos\frac{\omega_1 t}{2} + ic_1\sin\frac{\omega_1 t}{2}\big). \end{aligned}$$

By comparing this relation to Eqs. A.15 and A.20, we see that:

$$c_3 = -ic_2 \ , \quad c_4 = ic_1. \tag{A.21}$$

We may obtain a more convenient form for these coefficients from the normalization condition on the wave function:

$$\begin{aligned} 1 &= <\Psi|\Psi> \\ &= \int \Psi^* \Psi dV \\ &= V\big[|\psi_1(t)|^2 + |\psi_2(t)|^2\big] \end{aligned} \tag{A.22}$$

Substitution of Eqs. A.19, A.20 into this relation yields (and using Eq. A.21):

$$1 = V\big[|c_1|^2 + |c_2|^2\big] \tag{A.23}$$

This relation is satisfied by:

$$c_1 = \frac{1}{\sqrt{V}} \cos\frac{\Theta}{2} \mathrm{e}^{-i\phi_1}, \qquad c_2 = \frac{1}{\sqrt{V}} \sin\frac{\Theta}{2} \mathrm{e}^{-i\phi_2}. \tag{A.24}$$

The expectation value of the z-component of the angular momentum, $<\mu_z>$ may now be calculated:

$$\begin{aligned} <\mu_z(t)> &= <\Psi|\mu_z|\Psi> \\ &= \int_V \Psi^* \mu_z \Psi dV \\ &= \frac{\gamma\hbar V}{2} \Psi^* \sigma_z \Psi \\ &= \frac{\gamma\hbar V}{2} \big[\psi_1^*(t), -\psi_2^*(t)\big] \big[\psi_1(t), \psi_2(t)\big]^T \\ &= \frac{\gamma\hbar V}{2} \big[|\psi_1(t)|^2 - |\psi_2(t)|^2\big]. \end{aligned} \tag{A.25}$$

Using Eqs. A.19 and A.20, we obtain:

$$\begin{aligned} <\mu_z(t)> &= \frac{\gamma\hbar V}{2}\Big\{ \big[|c_1|^2 \cos^2(\frac{\omega_1 t}{2}) + |c_2|^2 \sin^2(\frac{\omega_1 t}{2}) + \\ &\quad (c_1 c_2^* + c_1^* c_2) \cos\frac{\omega_1 t}{2} \sin\frac{\omega_1 t}{2}\big] - \\ &\quad \big[|c_2|^2 \cos^2(\frac{\omega_1 t}{2}) + |c_1|^2 \sin^2(\frac{\omega_1 t}{2}) - (c_1 c_2^* + c_1^* c_2) \cos\frac{\omega_1 t}{2} \sin\frac{\omega_1 t}{2}\big]\Big\} \\ &= \frac{\gamma\hbar V}{2}\Big\{ \big(|c_1|^2 - |c_2|^2\big)\big(\cos^2\frac{\omega_1 t}{2} - \sin^2\frac{\omega_1 t}{2}\big) - \\ &\quad 2\big(c_1 c_2^* + c_1^* c_2\big)\big(\cos\frac{\omega_1 t}{2} \sin\frac{\omega_1 t}{2}\big)\Big\} \\ &= \frac{\gamma\hbar V}{2}\Big[\big(|c_1|^2 - |c_2|^2\big)\cos\omega_1 t - \big(c_1 c_2^* + c_1^* c_2\big)\sin\omega_1 t\Big]. \end{aligned} \tag{A.26}$$

Use of Eq. A.24 yields:

$$\begin{aligned} < \mu_z(t) > &= \frac{\gamma\hbar V}{2}\Big[\frac{1}{V}\Big(\cos^2\frac{\Theta}{2} - \sin^2\frac{\Theta}{2}\Big)\cos\omega_1 t - \\ &\quad \frac{1}{V}\Big(\cos\frac{\Theta}{2}\sin\frac{\Theta}{2}\mathrm{e}^{i\phi_1 - i\phi_2} + \cos\frac{\Theta}{2}\sin\frac{\Theta}{2}\mathrm{e}^{i\phi_2 - i\phi_1}\Big)\sin\omega_1 t\Big] \\ &= \frac{\gamma\hbar}{2}\Big[\cos\Theta\cos\omega_1 t - \frac{1}{2}\sin\Theta\sin(\omega_1 t)(\mathrm{e}^{i\phi_1 - i\phi_2} + \mathrm{e}^{i\phi_2 - i\phi_1})\Big] \\ &= \frac{\gamma\hbar}{2}\Big[\cos\Theta\cos\omega_1 t - \sin\Theta\sin(\omega_1 t)\cos(\phi_1 - \phi_2)\Big] \\ &= \frac{\gamma\hbar}{2}\Big[\cos\Theta\cos\omega_1 t - \sin\Theta\sin(\omega_1 t)\cos\Phi\Big], \end{aligned} \tag{A.27}$$

where $\Phi = \phi_1 - \phi_2$. We have therefore completed our proof of Eq. 3.26.

A.6 Derivation of the z-Component of the Equilibrium Magnetization

The purpose of this appendix is to prove Eq. 3.34 from Eq. 3.32. Recall that Eq. 3.32 provides the z-component of the thermal equilibrium value of M_z for a magnetic number m_s:

$$M_0 = \rho \sum_{m_s=-s}^{s} P(\epsilon(m_s))\mu_z(m_s),$$

where ρ is the density of spins per unit volume in the homogeneous isochromat volume V. Using Eq. 3.8, we have that:

$$\epsilon = -m_s\hbar\omega_0 \ , \text{ and } \ \mu_z = m\gamma\hbar. \tag{A.28}$$

Thus,

$$M_0 = \frac{N\gamma\hbar}{V}\frac{\sum_{m_s=-s}^{s} m_s e^{m_s u}}{\sum_{m_s=-s}^{s} e^{m_s u}}, \quad u = \frac{\hbar\omega_0}{kT}. \tag{A.29}$$

This expression may be simplified by making an approximation. We see that u represents a ratio of two energies: the nuclear magnetic energy of the system, $\hbar\omega_0$, and its thermal energy, kT. In MR, the magnetic energies are significantly smaller than room-temperature thermal energies[2]:

$$\hbar\omega_0 \ll kT \quad \Rightarrow \quad u \ll 1 \quad \Rightarrow \quad \mathrm{e}^{m_s u} \approx 1 + m_s u.$$

[2] A simple calculation yields $u \approx 1.0214 \times 10^{-5}$ radians at $T = 300$K and $B_0 = 1.5$T.

We therefore use a Taylor expansion of $e^{m_s u}$ to simplify Eq. A.29:

$$\begin{aligned}
M_0 &= \rho\gamma\hbar\frac{\sum_{m_s=-s}^{s} m_s e^{m_s u}}{\sum_{m_s=-s}^{s} e^{m_s u}}, \quad \rho = \frac{N}{V} \\
&= \rho\gamma\hbar\frac{\sum_{m_s=-s}^{s} m_s(1+m_s u)}{\sum_{m_s=-s}^{s}(1+m_s u)}, \quad \text{because } \hbar\omega \ll kT \\
&= \rho\gamma\hbar\frac{u\sum_{m_s=-s}^{s} m_s^2}{(2s+1)}, \quad \text{because } \sum_{m_s=-s}^{s} m_s = 0 \\
&= \rho\frac{\gamma^2\hbar^2}{kT}\frac{2\sum_{m_s=0}^{s} m_s^2}{(2s+1)}B_0 \\
&= \rho\frac{\gamma^2\hbar^2}{kT}\frac{2s(s+1)(2s+1)/6}{(2s+1)}B_0 \\
&= \rho\frac{s(s+1)\gamma^2\hbar^2}{3kT}B_0. \qquad \text{(A.30)}
\end{aligned}$$

For a proton, $s = 1/2$, and this expression simplifies to:

$$M_0 \approx \rho\frac{\gamma^2\hbar^2}{4kT}B_0. \qquad \text{(A.31)}$$

We have therefore completed our proof of Eq. 3.34.

A.7 Derivation of the Demodulated Signal

The purpose of this appendix is to derive Eq. 3.42 from Eq. 3.39. This derivation is a mathematical model of the process of demodulation, or phase-sensitive detection (PSD). Our derivation is carried out in *one spatial dimension*, but can very easily be extended to higher dimensions, if desired.

PSD is carried out in a series of steps. The signal is first split into two channels, each channel is multiplied by either a sine or cosine function, and finally, an LPF is applied to each channel. The demodulated signal may be derived by multiplying Eq. 3.39 (for

simplicity, consider only one spatial dimension) by $2[\sin(\omega_0 t) + i\cos(\omega_0 t)]$:

$$\begin{aligned}
2[\sin(\omega_0 t) + i\cos(\omega_0 t)]s_m(t) &= 2[\sin(\omega_0 t) + i\cos(\omega_0 t)] \int M_\perp(x,0)\sin(\omega_0 t + \phi_G)dx \\
&= 2\int M_\perp(x,0)[\sin(\omega_0 t)\sin(\omega_0 t + \phi_G) + \\
&\quad i\cos(\omega_0 t)\sin(\omega_0 t + \phi_G)]dx \\
&= \int M_\perp(x,0)[\cos(\phi_G) - \cos(2\omega_0 t + \phi_G) + \\
&\quad i\sin(2\omega_0 t + \phi_G) - i\sin(-\phi_G)]dx \\
&= \int M_\perp(x,0)[\cos(\phi_G) - \cos(2\omega_0 t + \phi_G) + \\
&\quad i\sin(2\omega_0 t + \phi_G) + i\sin(\phi_G)]dx \\
&= \int M_\perp(x,0)[\mathrm{e}^{i\phi_G} - \mathrm{e}^{-i(2\omega_0 t + \phi_G)}]dx. \qquad \text{(A.32)}
\end{aligned}$$

Convolving with the low-pass filter (LPF) $A(t)$ yields $s_d(t)$:

$$\begin{aligned}
s_d(t) &= A(t) * \left\{2[\sin(\omega_0 t) + i\cos(\omega_0 t)]s_m(t)\right\} \\
&= A(t) * \int M_\perp(x,0)[\mathrm{e}^{i\phi_G} - \mathrm{e}^{-i(2\omega_0 t + \phi_G)}]dx.
\end{aligned}$$

Because the LPF, $A(t)$, is engineered to remove the high frequency term[3], $\mathrm{e}^{-i(2\omega_0 t + \phi_G)}$, we may simplify this expression:

$$s_d(t) \approx A(t) * \int_{-\infty}^{+\infty} M_\perp(x,0)\mathrm{e}^{i\phi_G}dx. \qquad \text{(A.33)}$$

Furthermore, as $A(t)$ is designed to be a LPF, it will (ideally) not distort the low frequency component of the signal, so that

$$s_d(t) \approx \int_{-\infty}^{+\infty} M_\perp(x,0)\mathrm{e}^{i\phi_G}dx. \qquad \text{(A.34)}$$

Thus, the expression for the demodulated signal as expressed in Eq. 3.42 has been obtained. The processing steps that have been used in this derivation are displayed in Fig. 3.2.

A.8 Derivation of the Slice Encoding Equation

The purpose of this section is to derive Eq. 3.45 from the Bloch equations using a standard set of RF pulses and gradient fields. This appendix requires the use of Appendix A.9 to derive the Bloch equations in the rotating frame with a rotating $B_1(t)$ RF field.

[3] Applying a LPF, $A(t)$, to Eq. A.32 easily removes the second sinusoidal term, as $\omega_0 t$ is typically a few orders of magnitude greater than $\phi_G(x,t)$, for all (x,t).

We begin with the Bloch equation for a given (x, y, z) coordinate:

$$\frac{d\vec{M}}{dt}(t) = \vec{M} \times \gamma \vec{B}(t),$$

and consider a particular radiofrequency (RF) pulse applied in the transverse plane:

$$\begin{aligned}\vec{B}_1(t) &= \frac{\omega_1(t)}{\gamma}[\cos(\omega_{rf}t)\hat{i} - \sin(\omega_{rf}t)\hat{j}] \\ &= B_1(t)[\cos(\omega_{rf}t)\hat{i} - \sin(\omega_{rf}t)\hat{j}].\end{aligned}$$

This RF pulse, together with the main magnetic field[4], $\vec{B}_0 = B_0\hat{k}$, leads to the following first order system of differential equations:

$$\begin{aligned}\frac{d\vec{M}}{dt}(t) &= \vec{M} \times \gamma\big[B_1(t)(\cos(\omega_{rf}t)\hat{i} - \sin(\omega_{rf}t)\hat{j}) + B_0\hat{k}\big] \\ &= \begin{pmatrix} 0 & \omega_0 & \omega_1(t)\sin\omega_{rf}t \\ -\omega_0 & 0 & \omega_1(t)\cos\omega_{rf}t \\ -\omega_1(t)\sin\omega_{rf}t & -\omega_1(t)\cos\omega_{rf}t & 0 \end{pmatrix} \begin{pmatrix} M_x \\ M_y \\ M_z \end{pmatrix}. \qquad \text{(A.35)}\end{aligned}$$

This system of differential equations can be solved to determine the trajectory of the net magnetization in the presence of the applied fields. The solution to this system describes a magnetization vector being tipped away from the z-axis and into the transverse plane, allowing for a signal to be measured (as we will soon see).

To solve this system, it is convenient to change the coordinate system into one rotating about the z-axis at frequency ω_{rf}. After changing to a rotating coordinate system, we obtain the following differential system (see Appendix A.9):

$$\begin{aligned}\frac{d\vec{M}_{rot}(t)}{dt} &= \vec{M}_{rot} \times \big[\omega_1(t)\hat{i}_{rot} + \big(\omega_0 - \omega_{rf}\big)\hat{k}_{rot}\big] \\ &= \begin{pmatrix} 0 & \omega_0 - \omega_{rf} & 0 \\ -(\omega_0 - \omega_{rf}) & 0 & \omega_1(t) \\ 0 & -\omega_1(t) & 0 \end{pmatrix} \begin{pmatrix} M_{x,rot}(t) \\ M_{y,rot}(t) \\ M_{z,rot}(t) \end{pmatrix}. \qquad \text{(A.36)}\end{aligned}$$

It is helpful to note that if we set $\omega_{rf} = \omega_0$, and if the RF pulse is designed to be constant over a time interval $t \in [0, T_{rf}]$, the solution to this system is simply:

$$\begin{pmatrix} M_{x,rot}(t) \\ M_{y,rot}(t) \\ M_{z,rot}(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\omega_1 t & \sin\omega_1 t \\ 0 & -\sin\omega_1 t & \cos\omega_1 t \end{pmatrix} \begin{pmatrix} M_{x,rot}(0) \\ M_{y,rot}(0) \\ M_{z,rot}(0) \end{pmatrix}, \quad t \in [0, T_{rf}].$$

This solution describes a rotation[5] of the vector $\vec{M}_{rot}(t)$ about $\hat{i}_{rot}(t)$.

[4]We are assuming that the main magnetic field is spatially invariant and static. See [34], chapters 20 and 27, for discussion of magnetic field inhomogeneities and main magnet coil design.

[5]The RF pulse can be thought of as a screwdriver, rotating $\vec{M}_{rot}(t)$ away from the z-axis.

Note that our solution is *still not spatially dependent*. However, the differential system from which it was derived may be changed by adding a gradient, $\vec{G}(z,t) = G_z(t)z\hat{k}$, that will permit spatial localization. With this gradient, the system becomes:

$$\begin{aligned}\frac{d\vec{M}_{rot}(z,t)}{dt} &= \vec{M}_{rot} \times \gamma\big[B_1(t)\hat{i}_{rot} + \big(B_0 + G_z z - \omega_{rf}/\gamma\big)\hat{k}_{rot}\big] \\ &= \begin{pmatrix} 0 & \omega_0 - \omega_{rf} + \gamma G_z z & 0 \\ -(\omega_0 - \omega_{rf} + \gamma G_z z) & 0 & \omega_1(t) \\ 0 & -\omega_1(t) & 0 \end{pmatrix} \begin{pmatrix} M_{x,rot} \\ M_{y,rot} \\ M_{z,rot} \end{pmatrix}.\end{aligned}$$

Again, if we set $\omega_{rf} = \omega_0$, then our system may be simplified:

$$\frac{d\vec{M}_{rot}(z,t)}{dt} = \begin{pmatrix} 0 & \gamma G_z z & 0 \\ -\gamma G_z z & 0 & \omega_1(t) \\ 0 & -\omega_1(t) & 0 \end{pmatrix} \begin{pmatrix} M_{x,rot} \\ M_{y,rot} \\ M_{z,rot} \end{pmatrix}.$$

We have completed our proof of Eq. 3.45.

A.9 The Bloch Equation in the Rotating Frame

The purpose of this appendix is to derive Eq. A.36. Let us define vectors $\vec{M}_{rot} = [M_{x,rot}, M_{y,rot}, M_{z,rot}]$ and $\vec{B}_{rot} = [B_{x'}, B_{y'}, B_{z'}]$, which represent vectors $\vec{M}$ and $\vec{B}$ in the rotating frame:

$$\begin{aligned}\vec{M} &= M_x\hat{i} + M_y\hat{j} + M_z\hat{k} \\ &= M_{x,rot}\hat{i}_{rot} + M_{y,rot}\hat{j}_{rot} + M_{z,rot}\hat{k}_{rot} \\ &= \vec{M}_{rot},\end{aligned} \tag{A.37}$$

and:

$$\begin{aligned}\vec{B} &= B_x\hat{i} + B_y\hat{j} + B_z\hat{k} \\ &= B_{x'}\hat{i}_{rot} + B_{y'}\hat{j}_{rot} + B_{z'}\hat{k}_{rot} \\ &= \vec{B}_{rot}.\end{aligned} \tag{A.38}$$

Differentiating A.37 with respect to time yields[6]:

$$\begin{aligned}\frac{d\vec{M}(t)}{dt} \quad = \quad & \frac{dM_{x,rot}}{dt}\hat{i}_{rot} + \frac{dM_{y,rot}}{dt}\hat{j}_{rot} + \frac{dM_{z,rot}}{dt}\hat{k}_{rot} + \\ & M_{x,rot}\frac{d\hat{i}_{rot}}{dt} + M_{y,rot}\frac{d\hat{j}_{rot}}{dt} + M_{z,rot}\frac{d\hat{k}_{rot}}{dt}\end{aligned} \tag{A.39}$$

[6]For our particular set of applied fields, we see that $\frac{d\hat{k}_{rot}}{dt}$ is actually zero. We have kept this term in our derivations simply to maintain generality.

If we introduce the vector $\vec{\Omega}_{rf} = \omega_{rf}\hat{k}$, we can easily verify that:

$$\frac{d\hat{i}_{rot}}{dt} = \hat{i}_{rot} \times \vec{\Omega}_{rf} \ , \ \frac{d\hat{j}_{rot}}{dt} = \hat{j}_{rot} \times \vec{\Omega}_{rf} \ , \ \frac{d\hat{k}_{rot}}{dt} = \hat{k}_{rot} \times \vec{\Omega}_{rf}. \tag{A.40}$$

For convenience, we introduce the following notation for a primed derivative:

$$\Big(\frac{d\vec{M}_{rot}(t)}{dt}\Big)' = \frac{dM_{x,rot}}{dt}\hat{i}_{rot} + \frac{dM_{y,rot}}{dt}\hat{j}_{rot} + \frac{dM_{z,rot}}{dt}\hat{k}_{rot}. \tag{A.41}$$

Thus, Eq. A.39 becomes:

$$\begin{aligned}
\frac{d\vec{M}(t)}{dt} &= \Big(\frac{d\vec{M}_{rot}(t)}{dt}\Big)' + M_{x,rot}\hat{i}_{rot} \times \vec{\Omega}_{rf} + M_{y,rot}\hat{j}_{rot} \times \vec{\Omega}_{rf} + M_{z,rot}\hat{k}_{rot} \times \vec{\Omega}_{rf} \\
&= \Big(\frac{d\vec{M}_{rot}(t)}{dt}\Big)' + \big(M_{x,rot}\hat{i}_{rot} + M_{y,rot}\hat{j}_{rot} + M_{z,rot}\hat{k}_{rot}\big) \times \vec{\Omega}_{rf} \\
&= \Big(\frac{d\vec{M}_{rot}(t)}{dt}\Big)' + \vec{M}_{rot}(t) \times \vec{\Omega}_{rf}.
\end{aligned} \tag{A.42}$$

Therefore:

$$\begin{aligned}
\Big(\frac{d\vec{M}_{rot}(t)}{dt}\Big)' &= \frac{d\vec{M}(t)}{dt} - \vec{M}_{rot}(t) \times \vec{\Omega}_{rf} \\
&= \vec{M}(t) \times \gamma\vec{B}(t) - \vec{M}_{rot}(t) \times \vec{\Omega}_{rf} \\
&= \vec{M}_{rot}(t) \times \gamma\vec{B}_{rot}(t) - \vec{M}_{rot}(t) \times \vec{\Omega}_{rf} \\
&= \vec{M}_{rot}(t) \times \gamma\Big(\vec{B}_{rot}(t) - \vec{\Omega}_{rf}/\gamma\Big).
\end{aligned} \tag{A.43}$$

If the total applied magnetic field, $\vec{B}_{rot}(t)$ is given by:

$$\begin{aligned}
\vec{B}(t) &= \vec{B}_1(t) + \vec{B}_0 \\
&= B_1(t)\big(\cos(\omega_{rf}t)\hat{i} + \sin(\omega_{rf}t)\hat{j}\big) + B_0\hat{k} \\
&= B_1(t)\hat{i}_{rot} + B_0\hat{k} \\
&= \vec{B}_{rot}(t),
\end{aligned}$$

then Eq. A.43 becomes:

$$\begin{aligned}
\Big(\frac{d\vec{M}_{rot}(t)}{dt}\Big)' &= \vec{M}_{rot}(t) \times \gamma\Big(\vec{B}_{rot}(t) - \vec{\Omega}_{rf}/\gamma\Big) \\
&= \vec{M}_{rot}(t) \times \Big(\omega_1(t)\hat{i}_{rot} + (\omega_0 - \omega_{rf})\hat{k}\Big).
\end{aligned} \tag{A.44}$$

We have therefore completed our proof of Eq. A.36.

A.10 Derivation of the Fourier Transform Model of the MR Signal Equation

This appendix is simply a derivation of Eq. 3.50 from Eq. 3.47:

$$s_d(t) = \int\int M_{z_0}(x,y)\exp[i\phi_G(x,y,t)]dxdy.$$

Here we use a two (spatial) dimensional model, but our derivation can easily be extended to one or three spatial dimensions.

Consider again the Larmor equation, which relates the spin frequency to the applied magnetic field at time t and position (x, y),

$$\omega(x, y, t) = \gamma B(x, y, t). \tag{A.45}$$

From Eq. 3.48, we can obtain a relationship between spin frequency and the gradient fields:

$$\begin{aligned} \omega_G(x, y, t) &= \gamma B_G(x, y, t) \\ &= \gamma(G_x(t)x + G_y(t)y) \\ &= \gamma \vec{r} \cdot \vec{G}(t). \end{aligned} \tag{A.46}$$

Finally, to make a connection between the *phase*, $\phi(x, y, t)$, and spin *frequency*, we integrate Eq. A.46 with respect to time to obtain:

$$\begin{aligned} \phi_G(x, y, t) &= -\int_0^t \omega_G(x, y, \tau) d\tau \\ &= -\int_0^t \gamma B_G(x, y, \tau) d\tau \\ &= -\gamma \vec{r} \cdot \int_0^t \vec{G}(\tau) d\tau, \end{aligned} \tag{A.47}$$

where the components of the time integrated gradient fields are:

$$\int_0^t \vec{G}(t')dt' = \big[\int_0^t G_x(t')dt', \int_0^t G_y(t')dt' \big]. \tag{A.48}$$

We may now obtain the desired expression for the demodulated signal in terms of the magnetic field gradients by substituting Eq. A.47 into Eq. 3.47:

$$\begin{aligned} s_d(t) &= \int\int M_{z_0}(x, y) \exp[i\phi_G(x, y, t)] dx dy \\ &= \int\int M_{z_0}(x, y) \exp[-i\gamma \vec{r} \cdot \int_0^t \vec{G}(\tau) d\tau] dx dy. \end{aligned}$$

Using a simple change of variables (Eq. 3.51), yields the desired Eq. 3.50.

A.11 One-dimensional k-Space Model of the Merged Data

We will now derive Eq. 5.13. In the spatial domain, the merged data, $g(j\Delta x)$ can be expressed as (Eq. 5.12):

$$\begin{aligned} g(j\Delta x/2) &= \begin{cases} l(j\Delta x/2, 0), & j = 0, 2, 4, \cdots, J-2 \\ l\big((j-1)\Delta x/2, \Delta x/2\big), & j = 1, 3, 4, \cdots, J-1 \end{cases} \\ &= l_z(j\Delta x, 0) + l_z\big((j-1)\Delta x, \Delta x/2\big), \end{aligned} \tag{A.49}$$

where $J = 2N_l$, and:

$$l_z(j\Delta x, 0) = \begin{cases} l(j\Delta x/2, 0), & j = 0, 2, 4, \cdots, J-2 \\ 0, & j = 1, 3, 4, \cdots, J-1 \end{cases}$$

$$l_z(j\Delta x, \Delta x/2) = \begin{cases} l(j\Delta x/2, \Delta x/2), & j = 0, 2, 4, \cdots, J-2 \\ 0, & j = 1, 3, 4, \cdots, J-1 \end{cases}.$$

We will denote the DFT of $l_z(j\Delta x, r)$ by $L_z(m\Delta k, r)$, $m = -N_l, -N_l+1, \cdots, N_l-1$. We require an expression for $L_z(m\Delta k, r)$ in terms of $L(m\Delta k, r)$. Recall that by using a discrete set of x values, we have made the assumption that $L(m\Delta k, r)$ is periodic, with period $N_l\Delta k$:

$$L(m\Delta k, r) = L\big((m+N_l)\Delta k, r\big), \qquad m = -N_l/2, -N_l/2+1, \cdots, N_l/2-1.$$

It can be shown that by performing the IDFT on $L(m\Delta k, r)$ with $2N_l$ points, rather than N_l points yields $l_z(m\Delta x, r)$ (see Appendix A.12 or [39, 62] for derivation):

$$\sum_{m=-N_l}^{N_l-1} L(m\Delta k, r) \exp\big(+2\pi i j \frac{m}{2N_l}\big) = l_z(j\Delta x, r), \qquad j = 0, 1, 2, \cdots 2N_l - 1. \tag{A.50}$$

It follows that:

$$L_z(m\Delta k, r) = L(m\Delta k, r), \qquad m = -N_l, -N_l+1, \cdots, N_l-1.$$

Since $L(m\Delta k, r)$ has a period of $N_l\Delta k$, it follows that $L_z(m\Delta k, r)$ also has a period of $N_l\Delta k$. Using the discrete Fourier shift theorem and taking the DFT of Eq. A.49 yields:

$$\begin{aligned} G(m\Delta k) &= L_z(m\Delta k, 0) + \mathrm{e}^{2\pi i \frac{\Delta x}{2} m\Delta k} L_z(m\Delta k, \Delta x/2), \;\; m \in [-N_l, N_l-1] \\ &= \begin{cases} L_z(m\Delta k, 0) + \mathrm{e}^{2\pi i \frac{\Delta x}{2} m\Delta k} L_z(m\Delta k, \Delta x/2), & m \in [-N_l, -N_l/2-1] \\ L_z(m\Delta k, 0) + \mathrm{e}^{2\pi i \frac{\Delta x}{2} m\Delta k} L_z(m\Delta k, \Delta x/2), & m \in [-N_l/2, N_l/2-1] \\ L_z(m\Delta k, 0) + \mathrm{e}^{2\pi i \frac{\Delta x}{2} m\Delta k} L_z(m\Delta k, \Delta x/2), & m \in [N_l/2, N_l-1] \end{cases} \\ &= \begin{cases} L_z(m\Delta k, 0) + \mathrm{e}^{2\pi i \frac{m}{2N_l}} L_z(m\Delta k, \Delta x/2), & m \in [-N_l, -N_l/2-1] \\ L_z(m\Delta k, 0) + \Big[L_z(m\Delta k, 0) - dL(m\Delta k, \Delta x/2)\Big], & m \in [-N_l/2, N_l/2-1] \\ L_z(m\Delta k, 0) + \mathrm{e}^{2\pi i \frac{m}{2N_l}} L_z(m\Delta k, \Delta x/2), & m \in [N_l/2, N_l-1] \end{cases}. \end{aligned}$$

Eqs. 1.5 and 5.10 were used in the last simplification. For notational convenience, we introduce the following sets:

$$\begin{aligned}\mathcal{N}_1 &= [-N_l, -N_l/2 - 1] \\ \mathcal{N}_2 &= [-N_l/2, N_l/2 - 1] \\ \mathcal{N}_3 &= [N_l/2, N_l - 1].\end{aligned}$$

Using that $L_z(m\Delta k, \Delta x/2)$ is periodic with period $N_l\Delta k$, we may then write:

$$\begin{aligned}G(m\Delta k) &= \begin{cases} L_z((m+N_l)\Delta k, 0) + \mathrm{e}^{2\pi i \frac{m}{2N_l}} L_z((m+N_l)\Delta k, \Delta x/2), & m \in \mathcal{N}_1 \\ 2L_z(m\Delta k, 0) - dL(m\Delta k, \Delta x/2), & m \in \mathcal{N}_2 \\ L_z((m-N_l)\Delta k, 0) + \mathrm{e}^{2\pi i \frac{m}{2N_l}} L_z((m-N_l)\Delta k, \Delta x/2), & m \in \mathcal{N}_3 \end{cases} \\ &= \begin{cases} L_z((m+N_l)\Delta k, 0) + \mathrm{e}^{2\pi i \frac{m+N_l-N_l}{2N_l}} L_z((m+N_l)\Delta k, \Delta x/2), & m \in \mathcal{N}_1 \\ 2L_z(m\Delta k, 0) - dL(m\Delta k, \Delta x/2), & m \in \mathcal{N}_2 \\ L_z((m-N_l)\Delta k, 0) + \mathrm{e}^{2\pi i \frac{m+N_l-N_l}{2N_l}} L_z((m-N_l)\Delta k, \Delta x/2) & m \in \mathcal{N}_3 \end{cases} \\ &= \begin{cases} L_z((m+N_l)\Delta k, 0) + \mathrm{e}^{2\pi i \frac{m+N_l}{2N_l}} \mathrm{e}^{2\pi i \frac{-N_l}{2N_l}} L_z((m+N_l)\Delta k, \Delta x/2), & m \in \mathcal{N}_1 \\ 2L_z(m\Delta k, 0) - dL(m\Delta k, \Delta x/2), & m \in \mathcal{N}_2 \\ L_z((m-N_l)\Delta k, 0) + \mathrm{e}^{2\pi i \frac{m-N_l}{2N_l}} \mathrm{e}^{2\pi i \frac{+N_l}{2N_l}} L_z((m-N_l)\Delta k, \Delta x/2), & m \in \mathcal{N}_3 \end{cases} \\ &= \begin{cases} L_z((m+N_l)\Delta k, 0) - \mathrm{e}^{2\pi i \frac{m+N_l}{2N_l}} L_z((m+N_l)\Delta k, \Delta x/2), & m \in \mathcal{N}_1 \\ 2L_z(m\Delta k, 0) - dL(m\Delta k, \Delta x/2), & m \in \mathcal{N}_2 \\ L_z((m-N_l)\Delta k, 0) - \mathrm{e}^{2\pi i \frac{m-N_l}{2N_l}} L_z((m-N_l)\Delta k, \Delta x/2), & m \in \mathcal{N}_3 \end{cases}.\end{aligned}$$

Using that $L_z((m+N_l)\Delta k, r) = L_z(m\Delta k, r)$ for all r, and $L_z(m\Delta k, r) = L(m\Delta k, r)$ for $m \in [-N_l/2, N_l/2 - 1]$ and all r, we can write:

$$G(m\Delta k) = \begin{cases} dL((m+N_l)\Delta k, 0), & m \in [-N_l, -N_l/2 - 1] \\ 2L(m\Delta k, 0) - dL(m\Delta k, \Delta x/2), & m \in [-N_l/2, N_l/2 - 1] \\ dL((m-N_l)\Delta k, 0), & m \in [N_l/2, N_l - 1] \end{cases}.$$

We have completed our proof of Eq. 5.13.

A.12 DFT of the Merged Data

The purpose of this appendix is to prove Eq. A.50. Let $l_1(m)$, $m = 0, 1, 2, \cdots, N_l$, be a discrete signal. We design a new sequence $l_{1,z}(n)$ with N_h samples, where $2N_l = N_h$, by interleaving zeros in $l_1(m)$:

$$l_{1,z}(n) = \begin{cases} l_1(\frac{n}{2}) & n = 0, 2, 4, \cdots, N_h - 2 \\ 0 & n = 1, 3, 5, \cdots, N_h - 1 \end{cases}. \tag{A.51}$$

To obtain the discrete Fourier Transform (DFT) of $l_{1,z}(n)$, we use the inverse DFT (IDFT) of $l_1(\frac{n}{2})$:

$$l_1(\frac{n}{2}) = \sum_{k=0}^{N_l-1} L_1(k)\exp(+i2\pi\frac{n}{2}\frac{k}{N_l}), \qquad n = 0, 2, 4, \cdots, N_h - 2. \tag{A.52}$$

The vector $L_1(k)$ is the DFT of $l_1(n)$, where $k = 0, 1, 2, \cdots, N_l$. To calculate the DFT of $l_{1,z}(n)$, we use the weighting factor $w(n)$:

$$\begin{aligned} w(n) &= \begin{cases} 1 & n \text{ even} \\ 0 & n \text{ odd} \end{cases} \\ &= \frac{1}{2}[1 + \exp(i\pi n)]. \end{aligned} \tag{A.53}$$

Combining Eqs. A.51, A.52, and A.53, we obtain:

$$\begin{aligned} l_{1,z}(n) &= \begin{cases} l_1(\frac{n}{2}), & n = 0, 2, 4, \cdots, N_h - 2 \\ 0, & n = 1, 3, 5, \cdots, N_h - 1 \end{cases} \\ &= l_1(\frac{n}{2})w(n), \qquad n = 0, 1, 2, \cdots, N_h - 1 \\ &= \frac{1}{2}\sum_{k=0}^{N_l-1} L_1(k)\exp(+i2\pi\frac{n}{2}\frac{k}{N_l})[1 + \exp(i\pi n)], \qquad n = 0, 1, 2, \cdots, N_h - 1 \\ &= \frac{1}{2}\sum_{k=0}^{N_l-1} L_1(k)[\exp(+i2\pi\frac{nk}{2N_l}) + \exp(+i2\pi\frac{n}{2N_l}(k + N_l))], \qquad n = 0, 1, 2, \cdots, N_h - 1 \end{aligned} \tag{A.54}$$

However, $2N_l = N_h$, and $L_1(k)$ is periodic with period N_l (this can be verified from the definition of the DFT of $l_1(n)$), so that $L_1(k) = L_1(k - N_l)$. Therefore, for $n = 0, 1, 2, \cdots, N_h - 1$, Eq. A.54 becomes:

$$\begin{aligned} l_{1,z}(n) &= \frac{1}{2}\sum_{k=0}^{N_l-1} L_1(k)\{\exp(+i2\pi\frac{nk}{N_h}) + \exp[+i2\pi\frac{n}{N_h}(k + N_l)]\} \\ &= \frac{1}{2}\sum_{k=0}^{N_l-1} L_1(k)\exp(+i2\pi\frac{nk}{N_h}) + \frac{1}{2}\sum_{k=0}^{N_l-1} L_1(k)\exp[+i2\pi\frac{n}{N_h}(k + N_l)] \\ &= \frac{1}{2}\sum_{k=0}^{N_l-1} L_1(k)\exp(+i2\pi\frac{nk}{N_h}) + \frac{1}{2}\sum_{k=N_l}^{N_h-1} L_1(k - N_l)\exp(+i2\pi\frac{n}{N_h}k) \\ &= \frac{1}{2}\sum_{k=0}^{N_h-1} L_1(k)\exp(+i2\pi\frac{nk}{N_h}). \end{aligned} \tag{A.55}$$

Thus, if $L_{1,z}(k)$ is the DFT of $l_{1,z}$, then:

$$L_{1,z}(k) = \frac{1}{2}L_1(k), \qquad \forall\ k \in \mathbb{Z}. \tag{A.56}$$

However, the period of $L_{1,z}$ is N_h, while the period of L_1 is only N_l. In other words, the spectrum of $2l_{1,z}$ is the same as that of l_1, but the *extents* of their spectrums are different by a factor of two (assuming that $\Delta k = 1$). Thus, interleaving zeros in the discrete signal l_1 stretches its spectral extent from N_l to N_h.

A.13 A k-Space Interpretation for Why dL is Small

We will now offer an explanation for why dL can be small that involves the k-space representation of this quantity. Often in NMR, the bandwidth of the anti-aliasing filter is much larger than the bandwidth of the measured signal [29]. In this case, the convolution integral may be approximated by $s(k)$:

$$\int_{-\delta k}^{\delta k} s(k_m - \kappa) A(\kappa / BW_{read}) d\kappa \approx s(k_m),$$

and $dL(k_m, r_n)$ becomes:

$$\begin{aligned}
dL(k_m, r_n) &= L(k_m, 0) - R(k_m, -r_n) L(k_m, r_n) \\
&= \int_{-\delta k}^{\delta k} s(k_m - \kappa) R(k_m - \kappa, 0) A(\kappa) d\kappa \\
&\qquad - R(k_m, -r_n) \int_{-\delta k}^{\delta k} s(k_m - \kappa) R(k_m - \kappa, r_n) A(\kappa) d\kappa \\
&= \int_{-\delta k}^{\delta k} s(k_m - \kappa) A(\kappa) d\kappa \\
&\qquad - R(k_m, -r_n) \int_{-\delta k}^{\delta k} s(k_m - \kappa) R(k_m - \kappa, r_n) A(\kappa) d\kappa \\
&\approx s(k_m) - R(k_m, -r_n) s(k_m) R(k_m, r_n) \\
&= 0.
\end{aligned}$$

Numerical simulations and experiments with real MRI data were carried out to study the extent to which this approximation is valid. The results of these experiments are described throughout Chapter 5.

A.14 Spatial Domain Interpretation for Why E_{dL} is Small

The purpose of this section is to prove Eq. 5.16a, which is derived in the limit as δk goes to infinity. We will first consider k to be a continuous parameter, and later discuss the effects of finite sampling. The expression for the energy of $dL(k, r_n)$ for this case can be

written as:

$$\begin{aligned} E_{dL}(r_n) &= \int_{-\infty}^{\infty} |dL(k, r_n)|^2 dk \\ &= \int_{-\infty}^{\infty} |dl(x, r_n)|^2 dx \\ &= \int_{-\infty}^{\infty} |a(x) - a(x - r_n)|^2 |\hat{\rho}(x)|^2 dx. \end{aligned}$$

Parseval's energy theorem [8] was used in this derivation. The function $\hat{\rho}(x)$ is the proton density convolved with a sinc function

$$\hat{\rho}(x) = 2k_{max} \int_{-\infty}^{\infty} \rho(x') \text{sinc}\big(2k_{max}(x - x')\big) dx',$$

and $\rho(x)$ is the one-dimensional proton density of the object. If we again use $A(k) = \text{sinc}\big(\frac{k}{\Delta k}\big)$, we can easily obtain Eq. 5.16a.

We have shown that the E_{dL} can be small relative to the energy of $L(k, 0)$ for continuous k. Sampling dL prior to calculating its energy will not increase the value of E_{dL}, i.e.:

$$\int_{-\infty}^{\infty} |dL(k, r_n)|^2 dk \leq \int_{-\infty}^{\infty} |dL(k_m, r_n)|^2 dk, \qquad m = 0, 1, 2, \cdots N_l - 1.$$

A.15 Continuity of Contraction Maps

The purpose of this appendix is to prove that:

$$d_X(\bar{x}_f, \bar{x}_g) \leqslant \frac{1}{1 - c_f} d_{Con(X)}(f(x), g(x)), \ \forall \ x \in X,$$

which is a result of Theorem 2, in Section 6.1.2. To prove this inequality, we first introduce the point $z = f(\bar{x}_g)$. Then,

$$\begin{aligned} d_X(\bar{x}_f, \bar{x}_g) &\leqslant d_X(\bar{x}_f, z) + d_X(z, \bar{x}_g), \quad (\textit{by the triangle inequality}) \\ &= d_X(\bar{x}_f, f(\bar{x}_g)) + d_X(f(\bar{x}_g), \bar{x}_g) \\ &\leqslant c_f d_X(x_f, x_g) + d_X(f(\bar{x}_g), \bar{x}_g), \quad (\textit{by contractivity}) \\ &= c_f d_X(x_f, x_g) + d_X(f(\bar{x}_g), g(\bar{x}_g)) \\ &= c_f d_X(x_f, x_g) + d_{Con(X)}(f, g), \ (\textit{by definition}). \end{aligned}$$

By rearranging this inequality, we obtain the desired result:

$$d_X(\bar{x}_{f(x)}, \bar{x}_{g(x)}) \leqslant \frac{1}{(1 - c_f)} d_{Con(X)}(f, g).$$

A.16 Contractivity Condition for Fractal Transform

The purpose of this appendix is to prove Eq. 6.5:

$$\| Tu - Tv \|_p \;\leq\; C_p \| u - v \|_p, \quad C_p = \sum_{m=1}^{N} |s_m|^{1/p} |\alpha_m|.$$

Recall the definition of the fractal transform, Eq. 6.3:

$$(\hat{T}u)(x) = \sum_{m=1}^{N} \left[\alpha_m u(w_m^{-1}(x)) + \beta_m\right] I_{X_m}(x).$$

Substitution yields:

$$\begin{aligned}
\| Tu - Tv \|_p =& \| \sum_{m=1}^{N} \left[\alpha_m u(w_m^{-1}(x)) + \beta_m\right] I_{X_m} - \sum_{m=1}^{N} \left[\alpha_m v(w_m^{-1}(x)) + \beta_m\right] I_{X_m} \|_p \\
=& \| \sum_{m=1}^{N} \alpha_m u(w_m^{-1}(x)) I_{X_m} - \sum_{m=1}^{N} \alpha_m v(w_m^{-1}(x)) I_{X_m} \|_p \\
=& \Big(\int \Big\{ \sum_{m=1}^{N} \alpha_m [u(w_m^{-1}(x)) - v(w_m^{-1}(x))] I_{X_m} \Big\}^p dx \Big)^{1/p}.
\end{aligned}$$

Making the substitution $y = w_m^{-1}(x)$ implies $x = w_m(y) = s_m y + a_m$, and $dx = s_m dy$:

$$\begin{aligned}
\| Tu - Tv \|_p &= \Big(\int \Big\{ \sum_{m=1}^{N} \alpha_m [u(y) - v(y)] I_{X_m} \Big\}^p s_m dy \Big)^{1/p} \\
&\leq \sum_{m=1}^{N} |s_m|^{1/p} |\alpha_m| \Big(\int \Big\{ [u(y) - v(y)] I_{X_m} \Big\}^p dy \Big)^{1/p} \\
&= C_p \| u - v \|_p .
\end{aligned}$$

We have completed our proof of Eq. 6.5. We see that the IFSM operator has the Lipschitz constant

$$C_p = \sum_{m=1}^{N} |s_m|^{1/p} |\alpha_m|. \tag{A.57}$$

A.17 Derivation of the IFTS Operator

The purpose of this appendix is to derive Eq. 7.6, which is found in Section 7.2. This equation describes relationship between the Fourier Transforms U and V. To make this connection, we first substitute the expression for the fractal transform (Eq. 6.3) into the

definition of the Fourier Transform (Eq. 7.5):

$$\begin{aligned} V(k) &= \sum_{m=1}^{N_{maps}} \alpha_m \int_{-\infty}^{\infty} \mathrm{e}^{-2\pi i k x} u(w_m^{-1}(x))\, dx \\ &\quad + \sum_{m=1}^{N_{maps}} \beta_m \int_{-\infty}^{\infty} \mathrm{e}^{-2\pi i k x} I_{X_m}(x)\, dx. \end{aligned} \tag{A.58}$$

We assume that u vanishes outside the interval $X = [-x_{max}, x_{max}]$. In the first set of integrals, we make the change of variable $y = w_m^{-1}(x)$, implying that $x = w_m(y) = s_m y + a_m$ and $dx = s_m\, dy$. The first set of integrals becomes

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-2\pi i k x} u(w_m^{-1}(x))\, dx &= s_m \int_X e^{-2\pi i k (s_m y + a_m)} u(y)\, dy \\ &= s_m e^{-2\pi i a_m k} \int_{-\infty}^{\infty} e^{-2\pi i s_m k y} u(y)\, dy \\ &= s_m e^{-2\pi i a_m k} U(s_m k). \end{aligned} \tag{A.59}$$

We now consider the integral:

$$\int_{-\infty}^{\infty} e^{-2\pi i k x} I_{X_m}(x)\, dx. \tag{A.60}$$

To evaluate this integral, we express $I_{X_m}(x)$ in terms of the rect function:

$$\begin{aligned} I_{X_m}(x) &= \begin{cases} 1, & x \in X_m \\ 0, & x \notin X_m \end{cases} \\ &= \begin{cases} 1, & |x - a_m| \le s_m x_{max} \\ 0, & |x - a_m| > s_m x_{max} \end{cases} \\ &= \Pi\Big(\frac{x - a_m}{2 s_m x_{max}}\Big) \\ &= \Pi\Big(\frac{x - a_m}{\sigma_m}\Big),\ \sigma_m = 2 s_m x_{max}. \end{aligned}$$

Here we used that X_m is the interval centered at a_m with full width of $2 s_m x_{max}$, and the following definition for $\Pi(x)$:

$$\Pi(x) = \begin{cases} 1, & |x| \le 1/2 \\ 0, & |x| > 1/2 \end{cases},$$

whose Fourier Transform is known. Eq. A.60 becomes:

$$\begin{aligned}
\int_{-\infty}^{\infty} \mathrm{e}^{-2\pi ikx} I_{X_m}(x)\ dx &= \int_{-\infty}^{\infty} \Pi\Big(\frac{x-a_m}{\sigma_m}\Big)\mathrm{e}^{-2\pi ikx}\ dx,\ \sigma_m = 2s_m x_{max} \\
&= \int_{-\infty}^{\infty} \Pi\Big(\frac{x'}{\sigma_m}\Big)\mathrm{e}^{-2\pi ik(x'+a_m)}\ dx',\ x' = x - a_m \\
&= \mathrm{e}^{-2\pi ia_m k}\int_{-\infty}^{\infty} \Pi\Big(\frac{x'}{\sigma_m}\Big)\mathrm{e}^{-2\pi ikx'}\ dx' \\
&= \sigma_m \mathrm{e}^{-2\pi ia_m k}\int_{-\infty}^{\infty} \Pi(x'')e^{-2\pi i\sigma_m kx''}\ dx'',\ x'' = \frac{x'}{\sigma_m} \\
&= \sigma_m \mathrm{e}^{-2\pi ia_m k}\frac{\sin(\pi\sigma_m k)}{\pi\sigma_m k} \\
&= \sigma_m \mathrm{e}^{-2\pi ia_m k}\mathrm{sinc}(\sigma_m k). \qquad \text{(A.61)}
\end{aligned}$$

Combining Eqs. A.58, A.59, and A.61, yields

$$\begin{aligned}
V(k) &= \sum_{m=1}^{N_{maps}} \alpha_m \int_{-\infty}^{\infty} \mathrm{e}^{-2\pi ikx} u(w_m^{-1}(x))\ dx + \sum_{m=1}^{N_{maps}} \beta_m \int_{-\infty}^{\infty} \mathrm{e}^{-2\pi ikx} I_{X_m}(x)\ dx \\
&= \sum_{m=1}^{N_{maps}} \alpha_m s_m e^{-2\pi ia_m k} U(s_m k) + \sum_{m=1}^{N_{maps}} \beta_m \sigma_m \mathrm{e}^{-2\pi ia_m k}\mathrm{sinc}(\sigma_m k) \\
&= \sum_{m=1}^{N_{maps}} \mathrm{e}^{-2\pi ia_m k}\Big[\alpha_m s_m U(s_m k) + \beta_m \sigma_m \mathrm{sinc}(\sigma_m k)\Big].
\end{aligned}$$

We have completed our proof of Eq. 7.6.

A.18 Proof That the IFTS Operator Maps $L^2(\mathbb{R})$ to Itself

The purpose of this appendix is to prove that $\hat{M}$ maps $L^2(\mathbb{R})$ to itself, which is item 1 on page 72. Using the definition of the IFTS operator,

$$\begin{aligned}
||V||_2 &= ||MU||_2 \\
&= \Big|\Big|\sum_{m=1}^{N_{maps}} \mathrm{e}^{-2\pi ia_m k}\Big[\alpha_m s_m U(s_m k) + \beta_m \sigma_m \mathrm{sinc}(\sigma_m k)\Big]\Big|\Big|_2 \\
&\leq \sum_m \Big[|\alpha_m||s_m|\Big|\Big|U(s_m k)\Big|\Big|_2\Big] + \sum_m \Big[|\beta_m||\sigma_m|\Big|\Big|\mathrm{sinc}(\sigma_m k)\Big]\Big|\Big|_2, \qquad \text{(A.62)}
\end{aligned}$$

by using the triangle inequality. From a change of variable,

$$\begin{aligned}
||U(s_m k)||_2 &= \Big(\int \big[U(s_m k)\big]^2 dk\Big)^{1/2} \\
&= \Big(\int \big[U(t)\big]^2 \frac{dt}{s_m}\Big)^{1/2},\ t = s_m k \\
&= |s_m|^{-1/2}||U(k)||, \qquad \text{(A.63)}
\end{aligned}$$

and similarily

$$||\text{sinc}(\sigma_m k)||_2 = |\sigma_m|^{-1/2}||\text{sinc}(k)||_2. \tag{A.64}$$

We therefore have the result

$$||V|| \leq \sum_m \Big[|\alpha_m||s_m|^{1/2}\Big]||U|| + \sum_m \Big[|\beta_m||\sigma_m|^{1/2}\Big]||\text{sinc}(k)|| < \infty, \tag{A.65}$$

because U and the sinc function are both in $L^2(\mathbb{R})$. Therefore V is in $L^2(\mathbb{R})$, and $\hat{M}$ maps functions in $L^2(\mathbb{R})$ to itself.

A.19 Derivation of $C_{IFTS,p}$

The purpose of this appendix is to prove that the IFTS operator has the Lipschitz constant (see Eq. 7.7)

$$\| \hat{M}U(k) - \hat{M}V(k) \|_p \leq C_{IFTS,p} \| U(k) - V(k) \|_p$$
$$C_{IFTS,p} = \sum_{m=1}^{N} |s_m|^{1+1/p}|\alpha_m|.$$

Recall the definition of the IFTS operator, Eq. 7.6:

$$V(k) = \sum_{m=1}^{N_{maps}} \text{e}^{-2\pi i a_m k}\Big[\alpha_m s_m U(s_m k) + \beta_m \sigma_m \text{sinc}(\sigma_m k)\Big].$$

Substitution yields

$$\begin{aligned}
\| \hat{M}U(k) - \hat{M}V(k) \|_p =& \| \sum_{m=1}^{N_{maps}} \text{e}^{-2\pi i a_m k}\Big[\alpha_m s_m U(s_m k) + \beta_m \sigma_m \text{sinc}(\sigma_m k)\Big] - \\
& \sum_{m=1}^{N_{maps}} \text{e}^{-2\pi i a_m k}\Big[\alpha_m s_m V(s_m k) + \beta_m \sigma_m \text{sinc}(\sigma_m k)\Big] \|_p \\
=& \| \sum_{m=1}^{N_{maps}} \text{e}^{-2\pi i a_m k}\alpha_m s_m \Big[U(s_m k) - V(s_m k)\Big] \|_p \\
=& \Big(\int \Big\{ \sum_{m=1}^{N_{maps}} \text{e}^{-2\pi i a_m k}\alpha_m s_m \Big[U(s_m k) - V(s_m k)\Big] \Big\}^p dk \Big)^{1/p} \\
\leq& \Big(\int \sum_{m=1}^{N_{maps}} \Big\{ \text{e}^{-2\pi i a_m k}\alpha_m s_m \Big[U(s_m k) - V(s_m k)\Big] \Big\}^p dk \Big)^{1/p} \\
=& \sum_{m=1}^{N_{maps}} |\alpha_m s_m| \Big(\int \Big[U(s_m k) - V(s_m k)\Big]^p dk \Big)^{1/p}.
\end{aligned}$$

Making the substitution $y = s_m k$ implies $k = y/s_m$, so that

$$\begin{aligned} \| \hat{M}U(k) - \hat{M}V(k) \|_p &\leq \sum_{m=1}^{N_{maps}} |\alpha_m s_m| |s_m|^{-1/p} \Big(\int \Big[U(y) - V(y) \Big]^p dy \Big)^{1/p} \\ &\leq \sum_{m=1}^{N_{maps}} |\alpha_m| |s_m|^{1-1/p} \| U - V \|_p \\ &= C_{IFTS,p} \| U - V \|_p . \end{aligned}$$

We have arrived at our proof of Eq. 7.7. We see that the Lipschitz constant for the IFTS operator

$$C_{IFTS,p} = \sum_{m=1}^{N_{maps}} |\alpha_m| |s_m|^{1-1/p},$$

is equal to the Lipschitz constant for the IFSM operator (Eq. A.57) only for the $p = 2$ case.

A.20 The Mathematical Connection Between Self-Similar and Autoregressive Based Methods

Here we provide a short but key derivation of Eq. 8.7. Using Eqs. 4.9 and 8.3:

$$\begin{aligned} \mathbf{r} &= [r(0), r(1), r(2), ..., r(N_P - 1)] \\ &= [- \sum_{m=1}^{2N_c} a_m s(N_m), - \sum_{m=1}^{2N_c} a_m s(N_m + 1), \ldots, - \sum_{m=1}^{2N_c} a_m s(N_S - 1 - m)] \\ &= - \sum_{m=1}^{2N_c} a_m [s(N_m), s(N_m + 1), \ldots, s(N_S - 1 - m)] \\ &= - \sum_{m=1}^{2N_c} a_m \mathbf{d}_m. \end{aligned}$$

where $N_m = N_S - N_P - m$.

A.21 Derivation of the One-Dimensional Linear Prediction Equation

The purpose of this appendix is to derive Eq. 9.3 from Eq. 9.2. This particular proof can be found in the literature [83, 97], but is referred to in several sections of this thesis. For example, the 1D piecewise constant and piecewise linear, as well as the 2D piecewise

constant models (discussed in Sections 4.1.3, and 9.2), make reference to the following derivation.

To begin, we first define a new variable

$$z_m = e^{-\sigma_m},\ m = 1, 2, 3, \ldots, 2N_c - 1, \tag{A.66}$$

using the unknown σ_m values. Eq. 9.2 becomes:

$$s_{N_C}(n) = \sum_{m=1}^{2N_c} h_m z_m^n,\ h_m, \sigma_m \in \mathbb{C},\ n = 0, 1, 2, \ldots, N_s - 1. \tag{A.67}$$

We also construct the polynomial

$$p(z) = \Pi_{m=1}^{2N_c}(z - z_m), \tag{A.68}$$

and the polynomial coefficients a_m, such that

$$p(z) = \Pi_{m=1}^{2N_c}(z - z_m) = \sum_{m=0}^{2N_c} a_m z^m, \tag{A.69}$$

where $a_{2N_c} = 1$. We see that $p(z_l) = 0$, for any integer $l = 0, 1, 2, \ldots, 2N_c$, and will use this property shortly.

We may construct the following sum:

$$\sum_{l=0}^{2N_c} a_l s_{N_C}(n - l),\ n = 2N_c, 2N_c + 1, 2N_c + 2, \ldots, N_s - 1. \tag{A.70}$$

We find that this new entity is always equal to zero, because

$$\begin{aligned}
\sum_{l=0}^{2N_c} a_l \hat{s}(n - l) &= \sum_{l=0}^{2N_c} a_l \sum_{m=1}^{2N_c} h_m z_m^{n-l} \\
&= \sum_{l=0}^{2N_c} a_l \sum_{m=1}^{2N_c} h_m e^{-\sigma_m(n-l)} \\
&= \sum_{m=1}^{2N_c} h_m e^{-\sigma_m n} \sum_{l=0}^{2N_c} a_l e^{\sigma_m l} \\
&= \sum_{m=1}^{2N_c} h_m e^{-\sigma_m n} p(z_m) \\
&= 0,
\end{aligned}$$

because $p(z_m) = 0$ for any integer $m \in [0, 2N_C]$, as noted above. We therefore have that

$$\begin{aligned}
0 &= \sum_{l=0}^{2N_c} a_l s_{N_C}(n - l) \\
&= a_0 s_{N_C}(n) + \sum_{l=1}^{2N_c} a_l s_{N_C}(n - l).
\end{aligned}$$

Simple rearrangement yields (assume $a_0 \neq 0$)

$$s_{N_C}(n) = -\sum_{l=1}^{2N_c} \frac{a_l}{a_0} s_{N_C}(n-l),\ n = 2N_c, 2N_c+1, 2N_c+2, \ldots, N_s - 1.$$

A simple change of variables yields Eq. 9.3.

A.22 Derivation of the One-Dimensional Piecewise Linear Model in Complex Trigonometric Polynomial Form

The purpose of this appendix is to prove Eq. 9.10. Starting with Eq. 9.9:

$$\mathsf{P}_{N_C}(k_x) = \mathrm{sinc}^2(\Delta x k_x)\Delta x \sum_{m=1}^{N_C} c_m e^{i\nu},\ \nu = -2\pi x_m k_x.$$

Using the identity $\sin\theta = \frac{1}{2i}(e^{i\theta} - e^{-i\theta})$, this becomes

$$\begin{aligned}
\mathsf{P}_{N_C}(k_x) &= \frac{1}{\Delta x \pi^2 k_x^2} \sin^2(\pi k_x) \sum_{m=1}^{N_C} c_m e^{i\nu} \\
&= \frac{-1}{4\Delta x \pi^2 k_x^2} (e^{i\theta} - e^{-i\theta})^2 \sum_{m=1}^{N_C} c_m e^{i\nu},\ \theta = \pi k_x \\
&= \frac{-1}{4\Delta x \pi^2 k_x^2} (e^{2i\theta} + e^{-2i\theta} - 2) \sum_{m=1}^{N_C} c_m e^{i\nu} \\
&= \frac{-1}{4\Delta x \pi^2 k_x^2} \sum_{m=1}^{N_C} c_m [e^{i(\nu+2\theta)} + e^{i(\nu-2\theta)} - 2e^{i\nu}] \\
&= \frac{-1}{4\Delta x \pi^2 k_x^2} \sum_{m=1}^{3N_C} d_m e^{i\phi_m k_x},
\end{aligned}$$

where

$$\phi_m = \begin{cases} \nu + 2\theta = -2\pi(x_m + \Delta x), & m = 1, 2, \ldots, N_C \\ \nu - 2\theta = -2\pi(x_m - \Delta x), & m = N_C+1, N_C+2, \ldots, 2N_C \\ \nu = -2\pi x_m, & m = 2N_C+1, 2N_C+2, \ldots, 3N_C \end{cases}$$

$$d_m = \begin{cases} c_m, & m = 1, 2, \ldots, N_C \\ c_{m-N_C}, & m = N_C+1, N_C+2, \ldots, 2N_C \\ -2c_{m-2N_C}, & m = 2N_C+1, 2N_C+2, \ldots, 3N_C \end{cases}.$$

We have derived Eq. 9.10.

A.23 Derivation of the Two-Dimensional Piecewise Constant Model in Discrete Complex Trigonometric Polynomial Form

The purpose of this appendix is to derive Eq. 9.16, from Eq. 9.14:

$$s_{N_C}(k_x, k_y) = -4\sin(\Delta x \pi k_x)\sin(\Delta y \pi k_y) \sum_{m_1=1}^{N_C} \sum_{m_2=1}^{N_C} c_{m_1,m_2} e^{-2\pi i(x_{m_1}k_x + y_{m_2}k_y)}.$$

For notational convenience, introduce the variables:

$$\theta_x = \pi \Delta x k_x,\ \theta_y = \pi \Delta x k_y,\ \nu_x = -2\pi x_{m_1} k_x,\ \nu_y = -2\pi x_{m_2} k_y.$$

Using the identity $\sin\theta = -0.5i(e^{\theta i} - e^{-\theta i})$, Eq. 9.14 becomes:

$$\begin{aligned}
s_{N_C}(k_x, k_y) &= -4\sin(\theta_x)\sin(\theta_y) \sum_{m_1=1}^{N_C} \sum_{m_2=1}^{N_C} c_{m_1,m_2} e^{i\nu_x + i\nu_y} \\
&= (e^{i\theta_x} - e^{-i\theta_x})(e^{i\theta_y} - e^{-i\theta_y}) \sum_{m_1=1}^{N_C} \sum_{m_2=1}^{N_C} c_{m_1,m_2} e^{i\nu_x + i\nu_y} \\
&= \sum_{m_1=1}^{N_C} \sum_{m_2=1}^{N_C} c_{m_1,m_2} [e^{i(\nu_x+\theta_x)} + e^{i(\nu_x-\theta_x)}][e^{i(\nu_y+\theta_y)} + e^{i(\nu_y-\theta_y)}] \\
&= \sum_{m_1=1}^{2N_C} \sum_{m_2=1}^{2N_C} d_{m_1,m_2} e^{i\zeta_{m_1}k_x} e^{i\xi_{m_2}k_y},
\end{aligned}$$

where

$$\zeta_{m_1} = \begin{cases} \nu_x + 2\theta_x = -2\pi(x_{m_1} + \Delta x), & m_1 = 1, 2, \ldots, N_C \\ \nu_x - 2\theta_x = -2\pi(x_{m_1} - \Delta x), & m_1 = N_C + 1, N_C + 2, \ldots, 2N_C \end{cases}$$

$$\xi_{m_2} = \begin{cases} \nu_y + 2\theta_y = -2\pi(y_{m_2} + \Delta y), & m_2 = 1, 2, \ldots, N_C \\ \nu_y - 2\theta_y = -2\pi(y_{m_2} - \Delta y), & m_2 = N_C + 1, N_C + 2, \ldots, 2N_C \end{cases}.$$

A simple relation can be derived that relates c_{m_1,m_2} to d_{m_1,m_2}. Otherwise, we have completed our proof of Eq. 9.16.

Appendix B

Algorithms and Examples

B.1 Application of Prolate Spheroidal Wave Functions for Extrapolation

The purpose of this section is to provide some insight into how extrapolation can be performed with Prolate Spheroidal Wave Functions (PSWF's). This appendix expands upon the discussion in Section 4.1.1.

The PSWF's can be defined as the solution to the integral equation [61, 76, 91]:

$$\lambda_n \psi_n(t) = R \int_{-k_{max}}^{+k_{max}} \psi_n(\tau) \text{sinc}[R(k-\tau)] d\tau. \tag{B.1}$$

The λ_n's are the eigenvalues, and $0 \leq n \leq \infty$. R is the known spatial extent of the measured data. Several important properties of PSWF's will be presented in order to explain how they may be used for extrapolation (proofs of these properties may be found throughout the literature [21, 76, 91]).

1. Given $k_{max} \in \mathbb{R}, k_{max} > 0$, and $R \in \mathbb{R}, R > 0$, we can find a countably infinite set of real spatially-limited functions $\{\psi_n(k)\}$ with spatial extent L and positive eigenvalues $\lambda_n \in \mathbb{R}, n = 0, 1, 2, ...$, such that $0 < \lambda_n < 1$.

2. The set $\{\psi_n(k)\}$ forms an orthonormal and complete basis in $\mathcal{R}_R \subset L^2(\mathbb{R})$, where $\mathcal{R}_R = \{F(k) | F(k) = \int_{-\frac{R}{2}}^{+\frac{R}{2}} f(x) e^{-i2\pi kx} dx,\ f \in L^2(\mathbb{R})\}$.

3. The set $\{\psi_n(k)\}$ forms an orthogonal and complete basis on $k \in [-k_{max}, +k_{max}]$:

$$\int_{-k_{max}}^{+k_{max}} \psi_m(k) \psi_n(k) dk = \begin{cases} 0, m \neq n \\ \lambda_n, m = n \end{cases}.$$

Now suppose that we are presented with the continuous spatially-limited extrapolation problem, and the spatial extent of $s(k)$ is known. Since $\{\psi_n(k)\}$ forms an orthonormal and complete basis for spatially limited functions in $\mathcal{R}_R$, $s(k)$ can be written as a linear combination of the basis functions $\{\psi_n(k)\}$:

$$s(k) = \sum_{n=0}^{\infty} c_n \psi_n(k), \ \ \forall \ \ k \in \mathbb{R}. \tag{B.2}$$

The functions $\{\psi_n(k)\}$ are known for all k, but the coefficients c_n are unknown and need to be calculated. However, if the c_n can be found, the PSWF series can be used to determine $\hat{s}(k)$ for all k. To determine the set of coefficients c_n, consider the multiplication of Eq. B.2 by $\psi_m(k)$, followed by an integration:

$$\int_{-k_{max}}^{+k_{max}} \psi_m(k)\hat{s}(k)dk = \int_{-k_{max}}^{+k_{max}} \psi_m(k) \sum_{n=0}^{\infty} c_n \psi_n(k)dk, \forall \ k \in \mathbb{R} \tag{B.3}$$

$$= \sum_{n=0}^{\infty} c_n \int_{-k_{max}}^{+k_{max}} \psi_m(k)\psi_n(k)dk \tag{B.4}$$

$$= a_m \lambda_m. \tag{B.5}$$

By rearranging the last equality, and using $\hat{s}(k) = s(k)$ for $k \in [-k_{max}, +k_{max}]$, we obtain

$$c_n = \frac{1}{\lambda_n} \int_{-k_{max}}^{+k_{max}} \psi_m(k)s(k)dk. \tag{B.6}$$

Thus, by using Eqs. B.2 and B.6, we can uniquely determine $\hat{s}(k) \ \ \forall \ \ k$ with knowledge of only $s(k), k \in [-k_{max}, +k_{max}]$, and the spatial extent of $\hat{s}(k)$, R.

This result may seem surprising! However, the PSWF algorithm presented here is unstable. If we perturb the measured data with a small constant, ϵ, we obtain a new expression for the coefficients:

$$\tilde{c}_n = c_n + \frac{\epsilon}{\lambda_n} \int_{-k_{max}}^{+k_{max}} \psi_m(k)dk. \tag{B.7}$$

We denote the perturbed c_n's by $\tilde{c}_n$. For small λ_n, the perturbation dominates the calculation of $\tilde{a}_n$. Furthermore, it can be shown that λ_n can be ordered so that they decrease monotonically for increasing values of n, and that the eigenvalues λ_n decay very rapidly near $n = 2k_{max}R$. Thus, $2k_{max}R$, is roughly the number of PSWF coefficients that are used in practice.

B.2 The Discrete PGA

In this chapter, we present several experiments that apply the discrete PG algorithm on simulated data. A variety of data produced in the spatial and frequency domains were synthesized to characterize the properties of this algorithm. Using these data sets, the relevant parameters that defined the data and used in the DPGA were modified. The results are reported below.

This appendix complements the discussion in Section 4.1.2.

B.2.1 Methods

A series of experiments were carried out on data generated in the spatial domain. In these experiments, a discrete sequence of spatial domain data, $h(n\Delta x)$, was produced, where $x_n = -x_{max} + n\Delta x$, $n = 0, 1, 2, \cdots, N_h - 1$, and $x_{max} = \frac{N_h \Delta x}{2}$. The sequence h was assumed to be periodic with period $N_h \Delta x$. Specifically, h was chosen to be:

$$h(x_n) = \Pi(\frac{x_n}{R}), \qquad n = 0, 1, 2, \cdots, N_h - 1. \tag{B.8}$$

The number R is the spatial support of the signal. The DFT of $h(n\Delta x)$ was used to calculate $H(n\Delta k)$, where Δk is related to N_h and Δx as in Eq. 2.13. Finally, the measured low resolution data, $L(n\Delta k)$, was produced by truncating $H(n\Delta k)$:

$$L(n\Delta k) = \Pi(\frac{n\Delta k}{N_l \Delta k}) H(n\Delta k), \qquad n = 0, 1, 2, \cdots, N_l - 1, \tag{B.9}$$

and $N_l < N_h$. The PG algorithm was implemented by replacing the continuous functions with vectors and the integration with a discrete summation in Eq. 4.4. The extrapolated signal, $L_{e,p}(n\Delta k)$, where p is the iteration number, is given by:

$$\begin{aligned} L_{e,p+1}(n\Delta k) &= L(n\Delta k) + [1 + \Pi(\frac{n\Delta k}{2k_{max}})] R_{est} \mathrm{sinc}(R_{est} n\Delta k) \circledast L_{e,p}(n\Delta k) \\ L_{e,0}(n\Delta k) &= L(n\Delta k). \end{aligned} \tag{B.10}$$

The circled asterisk, $\circledast$, denotes cyclic discrete convolution which may be performed as a multiplication in the spatial domain [8]. In some cases, it may not be possible to know the true spatial extent, R, of the object being imaged, and so an estimate, R_{est}, of the true spatial extent is used in the algorithm.

Several parameters were used to quantify the success of the DPGA. The average absolute distance between $L_{e,n}$ and H in the frequency domain was measured at each iteration:

$$E_k(n) = \frac{1}{N_e} \sum_i |L_{e,n}(i\Delta k) - H(i\Delta k)|. \tag{B.11}$$

The only terms included in Eq. B.11 were those that corresponded to extrapolated data.

Several experiments were conducted to test the performance of the PG algorithm using Eqs. B.8 and B.10. A simple initial experiment was implemented to demonstrate the use of the DPGA to extrapolate frequency data (parameters used for this experiment were $R = R_{est} = 0.5$, $N_l = 16$, $N_h = 24$, $\Delta k = 0.5$). In this first experiment, 10,000 iterations were performed.

Several other experiments were performed to explore the effects on changing R, R_{est}, and Δk, N_l on the convergence of the algorithm. The effect of adding noise to $L(n\Delta k)$ with a normal distribution and variance of σ was also studied. The specific parameters that were used for each of these experiments are given in Table B.1, and in each experiment 10,000 iterations were performed.

Other data sets were generated which were used to further explore the properties of the DPGA. A data set consisting of multiple rect functions was generated:

$$h(x_n) = \Pi(\frac{x_n}{R}) + \Pi(\frac{2x_n}{R}) + \Pi(\frac{4x_n}{R}). \tag{B.12}$$

For this experiment, $R = R_{est} = 0.8$, $N_l = 16$, $N_h = 20$, $\Delta k = 0.5, \sigma = 10^{-4}$.

Finally, the DPGA was also tested on the following signal:

$$H(k_n) = \sin(2\pi\omega k_n). \tag{B.13}$$

Here, $k_n = -k_{max} + n\Delta k$, $n = 0, 1, 2, \cdots, N_h - 1$, and $k_{max} = \frac{N_h \Delta k}{2}$. Parameters used were: $\Delta k = 0.5$, $N_l = 16$, $N_h = 20$, and $\omega = 5$.

B.2.2 Observations

Results obtained from changing R and R_{est} while holding all other experimental variables constant are shown in Figure B.3a. Increasing R resulted in a longer convergence time. Furthermore, the error in the extrapolated data slightly increased as R increased.

As shown in Figure B.3b, the PG algorithm was found to be very sensitive to under estimations of R. Choosing a value of R_{est} that was smaller than $R = 0.6$ yielded a large error. Choosing $R_{est} > R$ resulted in longer convergence times and more error in the extrapolation, but significantly less error than in the case when $R_{est} < R$.

As seen in Figure B.3c, changing the frequency sample spacing had no significant effect on the convergence time or the error in the extrapolated signal.

Another experiment examined the effect of changing the measured frequency extent while holding Δk and $\frac{N_h}{N_l}$ constant. As shown in Figure B.3d, as $N_l \Delta k$ was increased, the error in the extrapolated result increased, and the convergence rate decreased.

When noise was added to the measured low-resolution data, L, two important effects occurred: the error in the extrapolated result increased, and the convergence time decreased (see Figure B.3e). Indeed, the discrete PG algorithm is sensitive to noise.

Experiment Description	N_l	N_h	Δk	$N_l\Delta k$	$N_h\Delta k$	Δx	$N_h\Delta x$	R	R_{est}	σ
Effect of Changing Spatial Support	16	20	0.5	8	10	0.1	2	0.4 0.6 0.8 1.0 1.2 1.4	0.4 0.6 0.8 1.0 1.2 1.4	0
Effect of Changing Estimated Spatial Support	16	20	0.5	8	10	0.1	2	0.6	0.2 0.4 0.6 0.8 1.0 1.2	0
Effect of Changing Frequency Sample Spacing	8 / Δk	10 / Δk	0.1 0.2 0.25 0.4 0.5 1.0	8	10	0.1	1 / Δk	0.4	0.4	0
Effect of Changing Measured Frequency Extent	$2N_l\Delta k$	$2N_h\Delta k$	0.5	16 24 32 40 48 56	$1.25N_l\Delta k$	$(N_h\Delta k)^{-1}$	2	0.4	0.4	0
Effect of Adding Noise	16	20	0.5	8	10	0.1	2	0.4	0.4	0 1e-14 1e-10 1e-6 1e-2 1
Effect of Increasing the Extrapolation Length (or Decreasing Δx)	16	$2N_h\Delta k$	0.5	8	10 12 14 16 18 20	$(N_h\Delta k)^{-1}$	2	0.4	0.4	0

Figure B.1: Parameters used in experiments with the discrete PG algorithm.

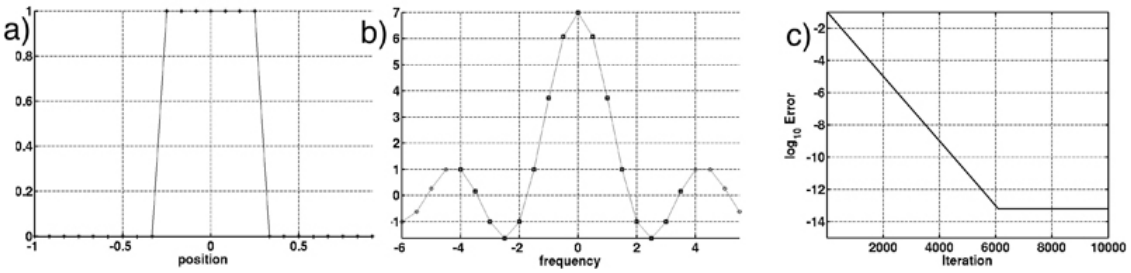


Figure B.2: Simple application of the DPGA: a) a simulated data set was generated in the spatial domain, and b) processed in the frequency domain. The squares represent the measured data, L, and circles represent the extrapolated data, $L_{e,1e4}$. The $\log_{10}$ of $E_k(n)$ as a function of the iteration number is shown in c).

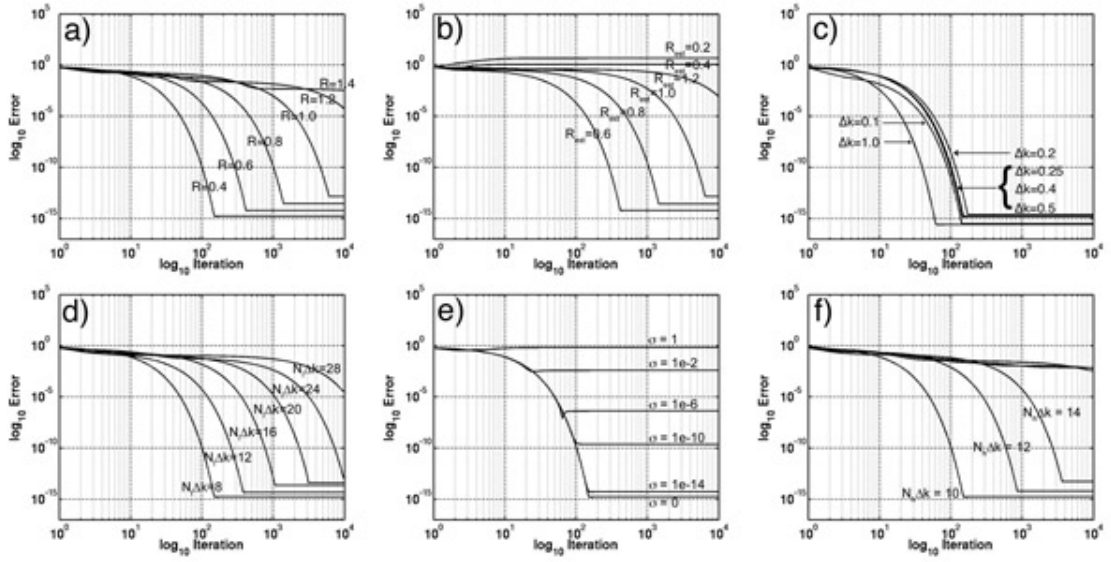


Figure B.3: The $\log_{10}$ of $E_k(n)$ as a function of the $\log_{10}$ of the iteration number calculated from the experiments listed in Fig. B.1 using the DPGA. Graphs are obtained from a) data with varying spatial support lengths, b) varying the estimated spatial support lengths, c) varying the frequency sample spacing, d) varying the measured spectral extent, e) varying the input noise level, and f) varying the extrapolation length.

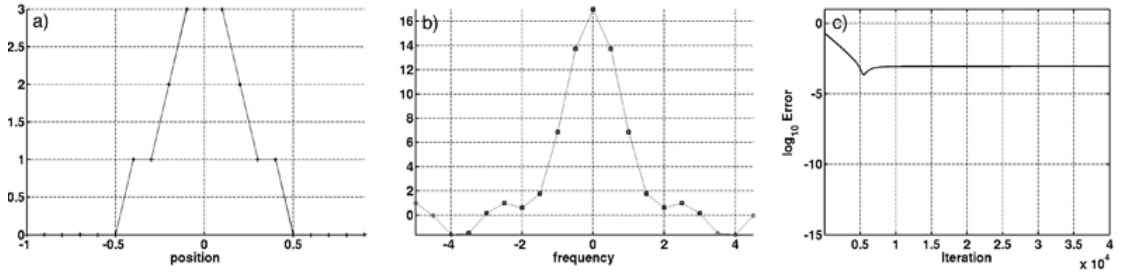


Figure B.4: Application of the DPGA to a data set made from multiple rect functions. Above are results showing: a) the simulated data set in the spatial domain, and b) the discrete series in the frequency domain. The squares represent the measured data, L, and circles represent the extrapolated data, $L_{e,4e4}$. The $\log_{10}$ of $E_k(n)$ as a function of the iteration number is shown in c).

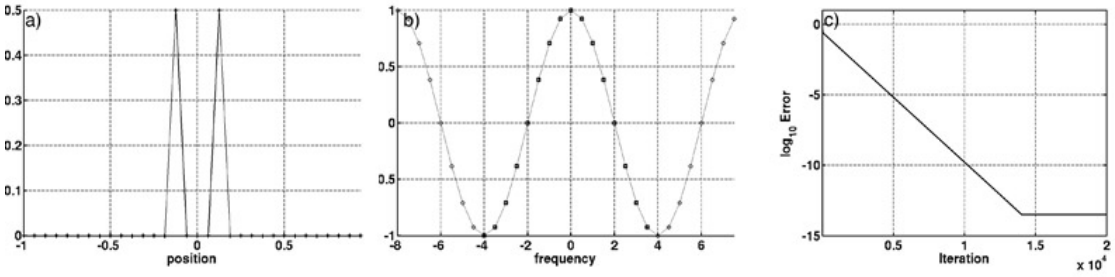


Figure B.5: Application of the DPGA to a frequency domain cosine function. Results show: a) the simulated data set in the spatial domain, and b) the discrete series in the frequency domain. The squares represent the measured data, L, and circles represent the extrapolated data, $L_{e,2e4}$. The $\log_{10}$ of $E_k(n)$ as a function of the iteration number is shown in c).

Finally, Figs. B.4 and B.5 show results of applying the extrapolation algorithm when multiple rect functions were used and when a cosine function were used, respectively. This gives us some indication that the algorithm is capable of handling a variety of spatially limited functions.

B.2.3 Discussion

From the experiments on the single rect function, it is apparent the DPGA is sensitive to the estimated spatial support. Underestimating R can lead to an unacceptable result, while overestimating R increases the convergence rate of the algorithm. This may be important if R is very large relative to the spatial extent, $N_h\Delta x$, as shown in B.3a.

How the data is sampled affects the DPGA's performance. While a decrease in Δk had little to no improvement in the final error, the measured frequency domain extent, $N_l\Delta k$, can affect the overall convergence rate.

Finally, noise in the measured data can make a significant difference in the performance of the DPGA. Figure B.3e shows that the error increases with σ, indicating that a possible future direction may be to investigate techniques that are more noise tolerant.

B.3 Phase-Constrained Reconstruction

The purpose of this appendix is to describe phase-constrained reconstruction algorithms. This discussion adds to the discussion in Section 4.1 of SRMRI techniques that employ a single image.

Phase-constrained techniques incorporate the idea that, neglecting off-resonance artifacts and noise, the reconstructed data should be real. Ideally, the desired spatial domain

signal represents the proton density, $\rho(x)$, of the object being scanned. As noted in Eq. 3.41, $\rho(x)$ is not only real, but also non-negative. These constraints, which are not accounted for in the standard Fourier reconstruction, have been incorporated into a number of MRI reconstruction algorithms [19, 37, 36, 53, 55, 60, 68, 100].

The Margosian Algorithm

A number of reconstruction techniques reconstruct the discrete image by using the assumption that the spatial domain image has strictly no imaginary components, or has zero phase. It can be shown that a real spatial domain data set has a Fourier spectrum with Hermitian symmetry. Using this property of the Fourier Transform, one can acquire data over only non-negative frequencies, and then use the complex conjugate of the known data to fill in the negative frequency data. The image is then formed using an IDFT of the completed data set.

The Margosian algorithm, which was first applied to MRI in 1986 [60], attempts to employ this idea into an MRI reconstruction technique. Off-resonance artifacts such as chemical shift artifacts, magnetic susceptibility artifacts [34, 55], and noise forces the frequency domain data set to lose its conjugate symmetry in general. Reconstructing the corrupted data with an IDFT yields an image that is not real, and thus has a non-zero phase component. The Margosian algorithm attempts to estimate this phase component, and uses this estimate to improve the reconstruction process. However, errors made in the phase estimate have been found to restrict the efficacy of this approach.

For the interested reader, a brief desecription of the Margosian algorithm implementation is provided in Appendix B.4.

Phase-Constrained POCS

The Margosian reconstruction technique has been improved through using additional *a priori* knowledge in the reconstruction process. The phase-constrained projection over convex sets (PCPOCS) algorithm [36, 37, 68] uses a data-consistency constraint that forces the extrapolated and measured data to be equal that on the region where the data was measured in the frequency domain. In fact, the data consistency constraint that the PCPOCS algorithm employs is the same as that used by the PG algorithm.

To implement the PCPOCS algorithm, an asymmetric data set is acquired and a phase estimate is obtained, exactly as was done for the Margosian algorithm. The data is then zero padded to create a symmetrical data set that can be brought into the spatial domain. The data is then brought into the spatial domain with an inverse Fourier Transform, and its phase is set equal to the phase estimate, $\tilde{\phi}(x)$. The PCPOCS algorithm then brings the spatial domain data back into k-space and imposes a data consistency constraint.

Specifically, the data on $k \geq 0$ is set equal to $L(k)$. This data set is then brought back to the spatial domain, phase constrained, and the process repeats itself until a convergence criterion is met.

B.4 The Margosian Algorithm

The purpose of this appendix is to provide a more detailed description of the Margosian algorithm, briefly discussed in Section B.3. To describe the Margosian algorithm, consider a one-dimensional frequency domain signal, $H(k) \in L^2(\mathbb{R})$, that is measured for non-negative frequencies to produce $L(k)$, $k \in [0, \infty)$:

$$L(k) = \begin{cases} 0.5(\text{sgn}(k) + 1)H(k), & k \neq 0 \\ H(0), & k = 0 \end{cases} . \tag{B.14}$$

The signum function is defined as:

$$\text{sgn}(k) = \begin{cases} -1 & k < 0 \\ +1 & k > 0 \end{cases} .$$

We wish to extrapolate the measured data $L(k)$ using the Margosian algorithm to form a new signal that is symmetrical in k-space.

Consider the inverse Fourier Transform of Eq. B.14:

$$\begin{aligned} l(x) &= h(x) * 0.5[\delta(x) + \frac{i}{\pi x}] \\ &= 0.5|h(x)|e^{i\phi(x)} + i0.5[|h(x)|e^{i\phi(x)}] * \frac{1}{\pi x}. \end{aligned} \tag{B.15}$$

The functions $l(x)$ and $h(x)$ are the inverse Fourier Transforms of $L(k)$ and $H(k)$, and $h(x) = |h(x)|e^{i\phi(x)}$. If a reasonable estimate, $\tilde{\phi}(x)$, to $\phi(x)$ can obtained, $\tilde{\phi}(x)$ may be used to phase correct Eq. B.14:

$$\begin{aligned} e^{-i\tilde{\phi}(x)}l(x) &= 0.5|h(x)|e^{i(\phi(x)-\tilde{\phi}(x))} + i0.5e^{-i\tilde{\phi}(x)}[|h(x)|e^{i\phi(x)}] * \frac{1}{\pi x}. \\ &\approx 0.5|h(x)|e^{i(\phi(x)-\tilde{\phi}(x))} + i0.5[|h(x)|e^{i(\phi(x)-\tilde{\phi}(x)}] * \frac{1}{\pi x}. \end{aligned} \tag{B.16}$$

If our phase estimate is close to $\phi(x)$, we may obtain $|h(x)|$, the desired signal, with $2\text{Real}\{e^{-i\tilde{\phi}(x)}l(x)\}$, where $\text{Real}\{f(x)\}$ extracts the real part of $f(x)$.

The Margosian algorithm is limited by the accuracy of the phase estimate which is in turn limited by the degree to which noise and off-resonance sources have corrupted the data. Various algorithms have been proposed to obtain the phase estimate with some success, but the main drawback to this method is the sensitivity to artifacts in the data. For some acquisition protocols (eg - spin echo methods) this algorithm produces useful results. However, for most *medical* applications that can have a rapidly varying phase in the image domain, this algorithm performs very poorly.

Appendix C

Abbreviations

The following is a partial list of abbreviations and acronyms used in this thesis.

ADC	Analog-to-digital converter
AR	Autoregressive
ARMA	Autoregressive moving average
CEPH	Collage Error Probability Histogram
DFT	Discrete Fourier Transform
DPGA	Discrete Papoulis Gerchberg Algorithm
FE	Frequency Encoding
fMRI	Functional Magnetic Resonance Imaging
FOV	Field of View
IDFT	Inverse Discrete Fourier Transform
IFS	Iterated Function System
IFSM	Iterated Function System with Grey Level Maps
IFTS	Iterated Fourier Transform System
LNCS	Lecture Notes in Computer Science
LP	Linear Prediction
LPF	Low Pass Filter
MA	Multiple Acquisitions
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
NMR	Nuclear Magnetic Resonance
PC	Phase Constrained
PE	Phase Encoding

List of abbreviations – continued on next page

PG	Papoulis Gerchberg
PGA	Papoulis Gerchberg algorithm
POCS	Projection onto Convex Sets
PSD	Phase Sensitive Detection
PSF	Point-spread function
PSWF	Prolate Shperoidal Wave Function
RF	Radiofrequency
SD	Standard deviation
SE	Slice Encoding
SL	Support Limited
SNR	Signal-to-noise ratio
SR	Super-resolution
SRMRI	Super-resolution MRI

Appendix D

Mathematical Symbols

The following is a partial list of the functions and variables used in this thesis. Some of the symbols are used for more than one purpose, and their meaning should be interpreted in the context within the section or chapter it is found. Page numbers refer to where the given symbol was defined or first used, if necessary.

α_n	67	a parameter that defines the n^{th} greyscale map
β_n	67	a parameter that defines the n^{th} greyscale map
γ	21	gyromagnetic ratio
$\delta(x)$		the Dirac delta function
δk	50	half-width of the anti-aliasing filer, A
Δ	74	collage distance
Δ_m	82	collage distance between a range block and the m^{th} domain block in $\mathcal{D}$
Δk	50	spatial frequency sample spacing
Δx	1	spatial sample spacing in the x-direction
Δz	31	slice width
θ	22	an arbitrary real constant
Θ	24	an arbitrary real constant
ι	43	a PSF in an SRMRI experiment
λ_n	39	the eigenvalue of the PSWF ψ_n
$\Lambda(x)$	12	the triangle function
$\vec{\mu}$	20	magnetic moment vector
ν_m	9	an arbitrary complex number
ρ	28	the proton density function
σ_m	72	a width parameter used in the IFTS definition

List of mathematical symbols – continued on next page

List of mathematical symbols– continued from previous page

$\vec{\sigma}$	20	Pauli spin operator
ϕ	27	the phase due to a applied magnetic fields
ϕ_n	67	the n^{th} greyscale map defining an IFS
ϕ_G	27	the phase due to a set of applied magnetic field gradients
ϕ_0	22	an arbitrary phase
Φ	24	an arbitrary phase
ψ_m	39	the m^{th} PSWF
$\psi_\pm$	22	spin up/down eigenvectors
Ψ	22	general solution to the time-dependent Schrödinger differential equation
ω		a radial frequency
ω_0		radial frequency due to the main magnetic field, $\omega_0 = \gamma B_0$
ω_{rf}	23	frequency at which the RF field rotates
$\Pi(k)$	12	the rectangle function
$\text{III}(k)$	12	the shah function
a_m	42	AR coefficients in Chapters 4, 8, 9
a_m	67	an IFS parameter in Chapters 6 and 7
$a(x)$	36	the anti-aliasing filter, in the spatial domain, that is employed in a typical MRI acquisition sequence
$A(k)$	36,50	the anti-aliasing filter, in the frequency domain, that is employed in a typical MRI acquisition sequence
$\vec{B}$	19	total applied magnetic field vector
$\vec{B}_{rot}$	110	the total applied magnetic field in the rotating frame
B_0	19	magnitude of the main magnetic field vector
$\vec{B}_1$	23	applied radiofrequency field vector
B_1	23	magnitude of the applied radiofrequency field vector
BW_{read}	3	the bandwidth in the readout direction
c_f	63	contraction factor for a contractive mapping, f
c_m	9	an arbitrary complex number
C_{m_s}	22	an arbitrary complex number
$C(r_n)$	54	angle between $L(k_m, 0)$ and $R(k_m, -r_n)L(k_m, +r_n)$
$\mathbb{C}$		the complex numbers
$Con(X)$	65	The set of all contraction maps in a given metric space, X
d_n	55	a position parameter used for an SR simulation
$\mathcal{D}$	82	the Domain pool, or set of domain blocks
dL	54	the difference between $L(k_m, 0)$ and $R(k_m, -r_n)L(k_m, r_n)$
E_{dL}	54	the l^2 energy of dL

List of mathematical symbols – continued on next page

List of mathematical symbols– continued from previous page

$f(x)$		an arbitrary, possibly complex, function
$F(x)$		the Fourier Transform of $f(x)$
$\mathcal{F}(f)$	11	the Fourier Transform of f
g	43	the merged data set in an SRMRI experiment
G	44	the DFT of g
$\vec{G}$	32	magnetic field gradient
G_x, G_y, G_z	31, 32	magnetic field gradients applied in the x, y, and z-directions
h	43	high resolution data set in an SRMRI experiment
H	21	Hamiltonian of a system
$\hbar$	20	Planck's constant divided by 2π
i		$\sqrt{-1}$
$\hat{i}$		unit vector in the x-direction
$\hat{i}_{rot}$	24	a unit vector rotating in the transverse plane
$I_A(x)$	67	the characteristic function of a set $A \subseteq X$: $I_A(x) = 1$ if $x \in A$ and $I_A(x) = 0$ otherwise
$\vec{J}$	20	nuclear angular momentum
J	20	magnitude of the nuclear angular momentum
J_z	20	z-component of the nuclear angular momentum
$\hat{j}$		unit vector in the y-direction
$\hat{j}_{rot}$	24	a unit vector rotating in the transverse plane
$\hat{k}$		unit vector in the z-direction
$\hat{j}_{rot}$	24	a unit vector, equal to $\hat{k}$
k, k_x, k_y		spatial frequency variables
k, k_x, k_y		spatial frequency variables
k_{max}	4	maximum (positive) frequency measured in a 1D MRI experiment
K	13	an arbitrary real constant
l_m	43	the m^{th} low resolution spatial domain data set, or channel
L_m	50	the m^{th} low resolution data set, or channel in k-space
$\vec{L}$	20	angular momentum for atomic and nuclear systems due to orbital motion
$L^2(\mathbb{R})$	11	the space of all square-integrable functions
$\vec{M}$	19	the net magnetization induced in an object via an MR experiment
$\hat{M}$	72	the IFTS operator
$\vec{M}_{z_0}$	32	the net magnetization, after slice encoding
$\vec{M}_{rot}$	110	the net magnetization in the rotating frame

List of mathematical symbols – continued on next page

List of mathematical symbols– continued from previous page

M_x, M_y, M_z	19	the x, y, and z-components of $\vec{M}$
$M_\perp$	28	magnitude of the transverse component of the net magnetization vector
m		a summation index
m_s	20	a magnetic quantum number
n		a summation index
N		an arbitrary integer, greater than zero
N_{acq}	3	number of averaged acquisitions
N_{bin}	3	number of bins in a joint probability histogram
N_{img}	6	number of images in an SRMRI sequence
N_{GQ}	43	number of points Gaussian quadrature scheme
N_h	43	number of data points in a high resolution data set
N_l	43	number of data points in a low resolution data set
N_x	3	the number of samples acquired in the x-direction
N_y	3	the number of samples acquired in the y-direction
N_z	3	the number of samples acquired in the z-direction
p		a summation index
$p_{i,j}(r_n, N_{bin})$	54	joint probability histogram between $l(j\Delta x, 0)$ and $l(j\Delta x - r_n, r_n)$, using N_{bin} bins
$\mathbf{r}$	82	a range block in 1D SRMRI experiment
r_n	50	the spatial translation used between MASRMRI acquisitions
$\mathcal{R}_R$	13	space of spatially limited functions with extent R
R	13	spatial extent of a 1D obect
$R(k_m, r_n)$	50	a discrete linear phase ramp with slope r_n
R_x, R_y, R_z	28	spatial extents of a 3D obect in the x, y, and z-directions
s	20	spin quantum number
$\vec{S}$	20	quantum mechanical spin angular momentum
S_z	20	z-component of the quantum mechanical spin angular momentum
$s(n)$		a discrete, and possibly complex, signal
s_m	27	in Chapters 3-5, the raw, continuous, real, measured MRI signal
s_m	67	in Chapters 6-9, a scaling parameter that defines IFS maps
$s_{true}(k)$	5	the desired, continuous MR signal defined over all values of the spatial frequency variable k
$s_{k_{max}}(k)$	5	a truncated version of $s_{true}(k)$: $s_{k_{max}}(k) = \Pi(k/2k_{max})s_{true}(k)$

List of mathematical symbols – continued on next page

$\text{sinc}(k)$	12	$\frac{\sin(\pi k)}{\pi k}$
T	26	temperature
$\hat{T}$	67	the fractal transform operator
T_1	26	the spin-lattice relaxation time
T_2	26	the spin-spin relaxation time
T_{rf}	109	the duration of time the RF pulse is switched on
TR	56	the repetition time in an MR acquisition sequence
TE	3	the echo time in an MR acquisition sequence
T_{scan}	3	the total time required to acquire one MR image
U	21	potential energy
V	24	a small volume within an object
w	55	a width parameter used for an SR simulation
w_n	67	the n^{th} affine contraction map defining an IFS
$W_m(x)$	42	a shifted and scaled rect function used in the 1D boxcar model
$\bar{x}$	64	the fixed point of a contractive operator

References

[1] S. Alexander. Multiscale methods in image modelling and image processing. Ph.D. Thesis, Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada, 2005. 10, 83

[2] S. K. Alexander, E. R. Vrscay, and S. Tsurumi. A simple, general model for the affine self-similarity of images. In A. Campilho and M. Kamel, editors, *Image Analysis and Recognition*, Lecture Notes in Computer Science, Volume 5112, pages 192–203. Springer Berlin / Heidelberg, 2008. 68, 69, 80, 82

[3] M. M. Barbieri and P. Barone. A two-dimensional Prony's method for spectral estimation. *IEEE Transactions on Signal Processing*, 40(11):2747–2756, 1992. 83, 94

[4] H. Barkhuijsen, R. De Beer, W. M. M. J. Bovée, and D. Van Ormondt. Retrieval of frequencies, amplitudes, damping factors, and phases from time-domain signals using a least-squares procedure. *Journal of Magnetic Resonance*, 61(3):465–481, 1985. 5, 8, 91

[5] M. F. Barnsley, V. Ervin, D. Hardin, and J. Lancaster. Solution of an inverse problem for fractals and other sets. *Proceedings of the National Academy of Sciences of the United States of America*, 83(7):1975–1977, 1986. 64, 66

[6] A. Baudes, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *SIAM Multiscale Modeling and Simulation*, 4(2):490–530, 2005. 83

[7] S. Borman and R. Stevenson. Spatial resolution enhancement of low-resolution image sequences - a comprehensive review with directions for future research. Technical report, 1998. 6

[8] R. Bracewell. *The Fourier Transform and its applications*. McGraw-Hill, 1978. 2, 10, 11, 12, 14, 36, 52, 117, 128

[9] R. Bracewell. *Two-Dimensional Imaging*. Prentice-Hall, Inc., 1995. 2, 10, 13, 18

[10] J. Cadzow. An extrapolation procedure for band-limited signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 27(1):4–12, 1979. 13, 39, 41

[11] E. Carmi, S. Liu, N. Alon, A. Fiat, and D. Fiat. Resolution enhancement in MRI. *Magnetic Resonance Imaging*, 24(2):133–154, 2006. 6, 44, 45, 53, 61, 95

[12] P. M. Centore and E. R. Vrscay. Continuity of attractors and invariant measures for iterated function systems. *Canadian Journal of Mathematics*, 37(3):315–329, 1994. 64, 65

[13] S. Chaudhuri. *Super-Resolution Imaging*. Kluwer Academic, Boston, 2001. 6

[14] C. Chen and D. Hoult. *Biomedical Magnetic Resonance Technology*, page 262. Adam Hilger, Bristol, 1989. 50, 55

[15] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley and Sons, Inc., 1991. 54

[16] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, 2007. 83

[17] K. R. Davidson and A. P. Donsig. *Real Analysis with Real Applications*. Prentice-Hall, Inc., 2002. 64

[18] A. W. Dowsey, J. Keegan, M. Lerotic, S. Thom, D. Firmin, and G. Yang. Motion-compensated MR valve imaging with COMB tag tracking and super-resolution enhancement. *Medical Image Analysis, Special Issue on the Ninth International Conference on Medical Image Computing and Computer-Assisted Interventions*, 11(5):478–491, 2007. 48, 96

[19] D. A. Feingberg, J. D. Hale, J. C. Watts, L. Kaufman, and A. Mark. Halving MR imaging time by conjugation: demonstration at 3.5 kG. *Radiology*, 161(2):527–531, 1986. 133

[20] Y. Fisher. *Fractal Image Compression*. Springer-Verlag, 1995. 68, 69, 88, 97

[21] C. Flammer. *Spheroidal Wave Functions*. Stanford University Press, Stanford, 1957. 38, 126

[22] B. Forte and E. R. Vrscay. Solving the inverse problem for measures using iterated function systems: A new approach. *Advances in Applied Probability*, 27(3):800–820, 1995. 75

[23] B. Forte and E. R. Vrscay. *Inverse problem methods for generalized fractal transforms*, volume 159 of *Fractal Image Encoding and Analysis, NATO ASI Series F: computer and system sciences*, pages 169–199. Springer Verlag, 1998. 75

[24] B. Forte and E. R. Vrscay. *Theory of Generalized Fractal Transforms*, volume 159 of *Fractal Image Encoding and Analysis, NATO ASI Series F: computer and system sciences*, pages 145–168. Springer Verlag, 1998. 65, 66, 67, 69, 80

[25] D. P. Gallegos and D. M. Smith. A NMR technique for the analysis of pore structure: determination of continuous pore size distributions. *Journal of Colloid and Interface Science*, 122(1):143–153, 1988. 91

[26] C. F. Gerald and P. O. Wheatley. *Applied Numerical Analysis, 5th Edition*. Addison Wesley, 1994. 56, 57

[27] R. W. Gerchberg. Super-resolution through error energy reduction. *Journal of Modern Optics*, 21(9):709–720, 1974. 39, 41

[28] M. Ghazel. Adaptive fractal and wavelet denoising. Ph.D. Thesis, Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada, 2004. 86, 99

[29] G. Giovannetti, V. Hartwig, V. Vitti, G. Gaeta, R. Francesconi, L. Landini, and A. Benassi. Application of undersampling technique for the design of an NMR signals digital receiver. *Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering*, 29B(3):107–114, 2006. 50, 55, 60, 116

[30] H. Greenspan. Super-resolution in medical imaging. *The Computer Journal (Advance Access)*, 2008. 6, 43

[31] H. Greenspan, G. Oz, N. Kiryati, and S. Peled. MRI inter-slice reconstruction using super-resolution. *Magnetic Resonance Imaging*, 20(5):437–446, 2002. 6, 44, 48, 52, 96

[32] H. Greenspan, G. Oz, N. Kiryati, and S. Peled. Super-resolution in MRI. In *IEEE International Symposium on Biomedical Imaging*, pages 943–946, 2002. 6, 48, 96

[33] H. Greenspan, S. Peled, G. Oz., and N. Kiryati. MRI inter-slice reconstruction using super-resolution. In *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention*, Lecture Notes in Computer Science, Volume 2208, pages 1204–1206. Springer Berlin / Heidelberg, 2001. 6, 44, 47, 48, 96

[34] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. V. Venkatesan. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. John Wiley and Sons, Inc., 1999. 3, 9, 14, 15, 19, 22, 24, 25, 27, 28, 31, 32, 50, 109, 133

[35] E. M. Haacke, Z. Liang, and S. H. Izen. Constrained reconstruction: A superresolution, optimal signal-to-noise alternative to the Fourier Transform in magnetic resonance imaging. *Medical Physics*, 16(3):388–397, 1989. 5, 8, 42, 81, 83

[36] E. M. Haacke, J. Mitchell, and D. Lee. Improved contrast at 1.5 Tesla using half-Fourier imaging: Application to spin-echo and angiographic imaging. *Magnetic Resonance Imaging*, 8(1):79–90, 1990. 8, 133

[37] M. E. Haacke, E. D. Lindskog, and W. Lin. A fast, iterative, partial-Fourier technique capable of local phase recovery. *Journal of Magnetic Resonance*, 92(1):126–145, 1991. 8, 133

[38] R. Haberman. *Elementary Applied Partial Differential Equations*. Prentice-Hall, Inc., 1998. 11, 12, 101

[39] M. Hayes. *Schaums Outline of Theory and Problems of Digital Signal Processing*, page 236. McGraw-Hill, 1999. 113

[40] W. S. Hinshaw and A. H. Lent. An introduction to NMR imaging: From the Bloch equation to the imaging equation. *Proceedings of the IEEE*, 71(3):338–350, 1983. 27, 28, 50

[41] D. Hoult. The solution of the Bloch equations in the presence of a varying B1 field - an approach to selective pulse analysis. *Journal of Magnetic Resonance*, 35:69–86, 1979. 31

[42] A. Jain and S. Ranganath. Extrapolation algorithms for discrete signals with application in spectral estimation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(4):830–845, 1981. 13, 39, 40, 41, 100

[43] W. Kalender, W. Seissler, E. Klotz, and P. Vock. Spiral volumetric ct with single-breath-hold technique, continuous transport, and continuous scanner rotation. *Radiology*, 176:181–183, 1990. 6

[44] M. Kasal, J. Halámek, V. Husek, M. Villa, U. Ruffina, and P. Cofrancesco. Signal processing in transceivers for nuclear magnetic resonance and imaging. *Review of Scientific Instruments*, 65(6):1897–1902, 1994. 55, 60

[45] S. M. Kay and S. L. Marple Jr. Spectrum analysis – A modern perspective. *Proceedings of the IEEE*, 69(11):1380–1419, 1981. 91

[46] P. Kornprobst, G. Malandain, O. Faugeras, R. Peeters, T. Vieville, S. Mierisova, S. Suanez, and P. Van Hecke. Superresolution in MRI and its influence in statistical analysis. Technical report, 2002. 6, 44, 48, 96

[47] P. Kornprobst, R. Peeters, M. Nikolova, R. Deriche, M. Ng, and P. Hecke. A superresolution framework for fMRI sequences and its impact on resulting activation maps. In R. Ellis and T. Peters, editors, *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003: 6th International Conference,*

Montréal, Canada, Proceedings, Part II, Lecture Notes in Computer Science, Volume 2879, pages 117–125, 2003. 6, 44, 48, 96

[48] D. Kramer, A. Li, I. Simovsky, C. Hawryszko, J. Hale, and L. Kaufman. Applications of voxel shifting in magnetic resonance imaging. *Investigative Radiology*, 25(12):1305–1310, 1990. 52

[49] R. Kumaresan. On the zeros of the linear prediction-error filter for deterministic signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(1):217–220, 1983. 90

[50] R. Kumaresan and D. Tufts. Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 30(6):833–840, 1982. 83

[51] C. Lanczos. *Applied Analysis*. Prentice Hall, 1957. 91

[52] Y. Li. Determining NMR relaxation times for porous media: Theory, measurement, and the inverse problem. M. Math Thesis, Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada, 2007. 91

[53] Z. P. Liang, F. E. Boada, R. T. Constable, E. M. Haacke, P. C. Lauterbur, and M. R. Smith. Constrained reconstruction methods in MR imaging. *Reviews of Magnetic Resonance in Medicine*, 4:67–185, 1992. 4, 6, 8, 27, 38, 41, 42, 43, 45, 61, 78, 133

[54] Z. P. Liang, E. M. Haacke, and C. W. Thomas. High-resolution inversion of finite Fourier Transform data through a localized polynomial approximation. *Inverse Problems*, 5:831–847, 1989. 42

[55] Z. P. Liang and P. C. Lauterbur. *Principles of Magnetic Resonance Imaging, A Signal Processing Perspective*. IEEE Press, 2000. 4, 6, 8, 27, 28, 29, 38, 42, 43, 45, 61, 133

[56] P. R. Locher. Computer simulation of selective excitation in NMR imaging. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 289(1037):537–542, 1980. 31

[57] C. H. Love. *Abscissas and Weights for Gaussian Quadrature for N=2 to 100, and N=125, 150, 175, and 200*, volume 98 of *National Bureau of Standards Monograph*. 1966. 56, 57

[58] J. Makhoul. Linear prediction: A tutorial review. *Proceedings of the IEEE*, 63(4):561–580, 1975. 42

[59] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1999. 11, 12, 14

[60] P. Margosian, F. Schmitt, and D. Purdy. Faster MR imaging: Imaging with half the data. *Health Care Instrumentation*, 1:195–197, 1986. 8, 133

[61] R. J. Marks II. *Introduction to Shannon Sampling and Interpolation Theory*. Springer-Verlag, 1991. 11, 13, 14, 15, 39, 126

[62] G. Mayer. Synthetic Aperture MRI. M.Sc. Thesis, Department of Electrical Engineering, The University of Calgary, Calgary, Alberta, Canada, 2003. 6, 44, 45, 52, 53, 113

[63] G. Mayer and E. R. Vrscay. Mathematical analysis of "phase ramping" for super-resolution magnetic resonance imaging. In A. Campilho and M. Kamel, editors, *Image Analysis and Recognition*, Lecture Notes in Computer Science, Volume 4141, pages 82–93. Springer Berlin / Heidelberg, 2006. 10, 44, 45, 49, 53, 54, 96

[64] G. Mayer and E. R. Vrscay. Measures of information for super-resolution MRI. In *ESMRMB 2006*, volume 19 of *Magnetic Resonance in Materials, Biology, Physics and Medicine*, pages 60–61, September 21-23 2006. 10, 50, 96

[65] G. Mayer and E. R. Vrscay. Iterated Fourier transform systems: A method for frequency extrapolation. In A. Campilho and M. Kamel, editors, *Image Analysis and Recognition*, Lecture Notes in Computer Science, Volume 4633, pages 728–793. Springer Berlin / Heidelberg, 2007. 10, 70, 81, 97

[66] G. Mayer and E. R. Vrscay. Measuring information gain for frequency encoded super-resolution MRI. *Magnetic Resonance Imaging*, 25(7):1058–1069, 2007. 6, 10, 50, 95, 96

[67] G. S. Mayer, E. R. Vrscay, M. L. Lauzon, B. G. Goodyear, and J. R. Mitchell. Self-similarity of images in the Fourier domain, with applications to MRI. In *Image Analysis and Recognition*, Lecture Notes in Computer Science, Volume 5112, pages 43–52. Springer Berlin / Heidelberg, 2008. 10, 81, 82, 89, 98

[68] G. McGibney, M. R. Smith, S. T. Nichols, and A. Crawley. Quantitative evaluation of several partial Fourier reconstruction algorithms used in MRI. *Magnetic Resonance in Medicine*, 30(1):51–59, 1993. 5, 8, 42, 45, 61, 78, 133

[69] M. Meadows. Precession and Sir Joseph Larmor. *Concepts in Magnetic Resonance*, 11(4):239–241, 1999. 21

[70] J. R. Mitchell, G. Mayer, M. L. Lauzon, and H. Zhu. Synthetic aperture MRI, Patent Issued August 5 2005, Filed 25 June 2004. 6, 44, 45, 53

[71] D. Mitsouras, A. Edelman, L. P. Panych, F. A. Jolesz, and G. P. Zientara. Super-resolved MRI by back-projection. In *Proceedings of the International Society for Magnetic Resonance in Medicine*, Volume 9, page 764, 2001. 6, 47, 48, 52, 96

[72] K. Ng, R. Deriche, P. Kornprobst, and M. Nikolova. Half-quadratic regularization for MRI image restoration. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP 03)*, Proceedings. Volume 6, pages 585–588. 6, 44

[73] L. Ning. *Fractal Imaging.* Academic Press, 1997. 97

[74] D. Nishimura. *Principles of Magnetic Resonance Imaging.* 1996. 3

[75] R. Otazo, R. Jordan, F. Lin, and S. Posse. Superresolution parallel MRI. In *IEEE International Confernece on Image Processing (ICIP 2007)*, Volume 3, pages 153–156, 2007. 6, 29, 96

[76] A. Papoulis. *The Fourier Integral and its Applications.* McGraw-Hill, 1962. 2, 10, 11, 14, 126

[77] A. Papoulis. A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Transactions on Circuits and Systems*, 22(9):735–742, 1975. 5, 13, 39, 40, 41

[78] J. Pauly, D. Nishimura, and A. Macovski. A k-space analysis of small-tip-angle excitation. *Journal of Magnetic Resonance*, 81:43–56, 1989. 31

[79] R. R. Peeters, P. Kornprobst, M. Nikolova, S. Sunaert, T. Vieville, G. Malandain, R. Deriche, O. Faugeras, M. Ng, and P. Van Hecke. The use of super-resolution techniques to reduce slice thickness in functional MRI, special issue: High-resolution image reconstruction, II. *International Journal of Imaging Systems and Technology*, 14(3):131–138, 2004. 6, 44, 48, 96

[80] S. Peled, H. Greenspan, G. Oz, Y. Yeshurun, and N. Kiryati. High resolution MRI from 2D and 3D superresolution applied to single-shot images. In *Proceedings of the International Society for Magnetic Resonance in Medicine*, Volume 9, page 165, 2001. 6, 47, 48, 52, 96

[81] S. Peled and Y. Yeshurun. Superresolution in MRI: Application to human white matter fiber tract visualization by diffusion tensor imaging. *Magnetic resonance in medicine*, 45(1):29–35, 2001. 6, 44, 49, 52

[82] S. Peled and Y. Yeshurun. Superresolution in MRI – perhaps sometimes. *Magnetic resonance in medicine*, 48(2):409, 2002. 6, 7, 44, 46, 52, 96

[83] R. Prony. Essai expérimental et analytique. *Paris J. de l"Ecole Polytechnique*, 1:24–76, 1795. 42, 89, 122

[84] K. Pruessmann, M. Weiger, M. Scheidegger, and P. Boesiger. Sense: Sensitivity encoding for fast mri. *Magnetic Resonance in Medicine*, 42:952–962, 1999. 7

[85] S. Ranganath and A. Jain. Two-dimensional linear prediction models – part I: Spectral factorization and realization. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(1):280–299, 1985. 83

[86] E. Roullot, A. Herment, I. Bloch, M. Nikolova, and E. Mousseaux. Regularized reconstruction of 3D high-resolution magnetic resonance images from acquisitions of anisotropically degraded resolutions. In *15th International Conference on Pattern Recognition (ICPR'00)*, Volume 3, page 3350, 2000. 6

[87] M. Sabri and W. Steenaart. An approach to band-limited signal extrapolation: The extrapolation matrix. *IEEE Transactions on Circuits and Systems*, 25(2):74–78, 1978. 39, 41

[88] J. Sanz and T. Huang. Discrete and continuous band-limited signal extrapolation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(5):1276–1285, 1983. 13, 39, 41

[89] J. Sanz and T. Huang. Some aspects of band-limited signal extrapolation: Models, discrete approximations, and noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(6):1492–1501, 1983. 39, 41

[90] K. Scheffler. Superresolution in MRI? *Magnetic resonance in medicine*, 48(2):408, 2002. 6, 7, 44, 46, 47, 49, 52, 96

[91] D. Slepian and H. Pollack. Prolate spheroidal wave functions, Fourier analysis and uncertainty – I. *Bell System Technical Journal*, pages 43–63, 1961. 5, 13, 38, 126

[92] M. R. Smith, S. T. Nichols, R. M. Henkelman, and M. L. Wood. Application of autoregressive moving average parametric modelling in magnetic resonance image reconstruction. *IEEE Transactions on Medical Imaging*, 5(3):132–139, 1986. 5, 8, 42

[93] V. Tom, T. Quatieri, M. Hayes, and J. McClellan. Convergence of iterative nonexpansive signal reconstruction algorithms. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(5):1052–1058, 1981. 39, 41

[94] H. Ur and D. Gross. Improved resolution from subpixel shifted pictures. *CVGIP: Graphical Models and Image Processing*, 54(2):181–186, 1992. 13

[95] E. R. Vrscay. A generalized class of fractal-wavelet transforms for image representation and compression. *Canadian Journal of Electrical and Computer Engineering*, 23(1):69–84, 1998. 81

[96] E. R. Vrscay. A simple model for affine self-similarity of images and its applications. In *Proceedings of LNLA 2008, The 2008 International Workshop on Local and Non-Local Approximation in Image Processing*, pages 61–76, 2008. 69

[97] L. Weiss and R. N. McDonough. Prony's method, z-transforms, and Padé approximation. *SIAM Review*, 5(2):145–149, 1963. 42, 91, 94, 122

[98] J. Woods. *Multidimensional Signal, Image, and Video Processing and Coding*. Academic Press, 2006. 17

[99] D. Youla. Generalized image restoration by the method of alternating orthogonal projections. *IEEE Transactions on Circuits and Systems*, 25(9):694–702, 1978. xii, 40, 41

[100] X. Zhang, E. Yacoub, and X. Hu. New strategy for reconstructing partial-Fourier imaging data in functional MRI. *Magnetic Resonance in Medicine*, 46(5):1045–1048, 2001. 133

[101] D. Zwillinger. *Handbook of Integration*. A K Peters, 1992. 56