Complex Bases, Number Systems and Their Application to Fractal-Wavelet Image Coding

by

Daniel G. Piché

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Applied Mathematics

Waterloo, Ontario, Canada, 2002

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A THESIS

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 explores new approaches to the analysis of functions by combining tools from the fields of complex bases, number systems, iterated function systems (IFS) and wavelet multiresolution analyses (MRA).

The foundation of this work is grounded in the identification of a link between two-dimensional non-separable Haar wavelets and complex bases. The theory of complex bases and this link are generalized to higher dimensional number systems. Tilings generated by number systems are typically fractal in nature. This often yields asymmetry in the wavelet trees of functions during wavelet decomposition. To acknowledge this situation, a class of extensions of functions is developed. These are shown to be consistent with the Mallat algorithm. A formal definition of local IFS on wavelet trees (LIFSW) is constructed for MRA associated with number systems, along with an application to the inverse problem.

From these investigations, a series of algorithms emerge, namely the Mallat algorithm using addressing in number systems, an algorithm for extending functions and a method for constructing LIFSW operators in higher dimensions. Applications to image coding are given and ideas for further study are also proposed.

Background material is included to assist readers less familiar with the varied topics considered. In addition, an appendix provides a more detailed exposition of the fundamentals of IFS theory.

Acknowledgements

In completing this work, I relied upon the support and assistance of many individuals. First, I would like to thank my supervisor Edward Vrscay. My choice of leaving Waterloo to start a company in Winnipeg, while still continuing my degree, must have caused him concern. However, he was always understanding of those many times when I had to put my research on hold in order to focus my energies on building my company. His support, which he demonstrated in countless ways, was vital to completing my degree. My eternal thanks, Ed.

I would also like to thank William Gilbert for introducing me to complex bases. My research really started when one day I found an article he wrote that he had given me while I was a Masters student. Will has always helped me find new ideas of study, including an impromptu discussion in early June 2002, that spawned weeks of investigations and a complete reworking of my results for general number systems.

Sincere thanks are extended to my readers, Doug Hardin, Sue Ann Campbell, Stan Lipshitz and Will and Ed, for their assistance in making this a better work. Thanks also to Helen Warren, our graduate secretary, for getting paperwork in and my mail sent to me during my long absence.

I would like to acknowledge the financial support given to me from various sources: NSERC (in the form of a PGS B scholarship and through Edward Vrscay's research grant), the Department of Applied Mathematics and Faculty of Mathematics (in the form of teaching assistantships) and the University of Waterloo (in the form of UW Graduate Scholarships). J'aimerais d'autant plus remercier profondément la Fondation Baxter & Alma Ricard de leur grande générosité, m'ayant accordé une de leurs premières bourses d'excellence et de leadership en 1999. La création de la Fondation est vraiment ce qui m'a permi de compléter mon doctorat ces derniers trois ans, durant une période où je n'avais vraiment pas d'autres sources de revenu disponibles, et où je devais consacrer mes énergies entièrement à mes études et au développement de mon entreprise. Ce soutien ne peut être sous estimé.

Thanks to my team and friends at myLocalStore. Thanks to my brother Robert and friend Christian Dandeneau for their excellent timing in choosing to start the company two years into my degree. Thanks also to Barry DeJaeger and Gary Slotnikov for sticking with us since the beginning, even though it has been tougher than we all bargained for. I know it has been worth it for all of us. Thanks also to Monelle for letting Chris play with us these past few years.

J'aimerais remercier mes parents, Michelle et René. Ils m'ont toujours encouragé et ont toujours démontré la confiance dans mes habiletés et dans ma persévérance de compléter ce projet. Je remercie aussi mon amour, Andréanne Bouchard, pour son soutien et son amour pendant les sept dernières années.

I extend sincere thanks to all those individuals at the Canadian Mathematical Society, including Graham Wright, John Borwein, Monique Bouchard, Alan Kelm, Kathy Heinrich,

Richard Kane, the members of the Student Committee, including Susan Cooper and Robert Juričevič, the many organizers of the Canadian Undergraduate Mathematics Conference and my friend Benoit Charbonneau for helping me stay involved in the mathematical community and getting all my many other projects accomplished. Thanks also to Corina Drapaca for helping me during the transition to Winnipeg in early 2000, by taking over my TA duties that winter.

There are many others who have been supportive of my efforts. I extend my deepest thanks to them all.

À Andéanne.

Cette thèse est une exploration.
Sa création a été soutenue par
intérêt et amour des
mathématiques.

Contents

List of Tables

List of Figures

Introduction

The idea of finding connections between different, seemingly unrelated things, is fascinating. Connections impact every aspect of the world in which we live. The works of Burke, for instance, explore the connections of many aspects of humanity's evolution [12, 13, 14]. This thesis is a result of exploring connections.

Inspiration for this work was found in the following areas of mathematics: fractals, iterated function systems (IFS), complex bases, number systems, and wavelet multiresolution analyses (MRA).

Although mathematicians have studied these areas for many years, and in some cases centuries, it is only in the past few decades, since the development and wide-scale availability of computers, that they have flourished. Indeed, computers have demonstrated many *real-world* applications for these complex areas in such diverse fields as finance, physics, biology, chemistry, cryptography, medicine and image coding, to name but a few.

The main text is comprised of this introduction, three chapters, a summary and an appendix. The first chapter presents mathematical background of the various topics explored in this thesis. The second chapter contains the results of the research. Readers familiar with the background material may begin with Chapter 2, referring to Chapter 1 as needed for occasional clarification. The third chapter presents sample applications of

the results to the field of image compression. This choice of application allows a visual illustration of the key ideas discussed. The summary contains a brief discussion of the findings and a sequence of possible questions for further study. The appendix presents additional background to the theory of IFS.

In more detail, Chapter 1 contains four sections. The first brief section lists a few definitions and results of functional analysis. This is principally for notational purposes since an understanding of functional analysis is required to comprehend the majority of this work. Section 1.2 describes the basic theory of IFS, starting with the fundamental concepts of contractivity of functions and attractors. Major concepts and results such as the Banach Contraction Mapping Principle, continuity of fixed points, and the inverse problem are given. Subsequently, the concepts of IFS, first developed by Barnsley and Demko, and IFS on grey-level maps (IFSM), introduced by Forte and Vrscay, are presented. The section continues with a discussion of the inverse problem for IFSM and its formal solution. It concludes with the theory of local IFSM (LIFSM) and its associated inverse problem. For the curious reader, Appendix A contains additional background discussions on the theory of IFS.

Section 1.3 presents an introduction to complex bases. First, the concepts of Gaussian integer, valid base, positional notation and the fundamental tile of a valid base are introduced. Certain key results regarding the existence of valid bases are presented, including an equivalence result of Gilbert. This is followed by a discussion of various algorithms to construct representations of Gaussian integers in valid bases.

The final section in Chapter 1 covers basic results of wavelet analysis. The discussion is restricted to the area of multiresolution analyses (MRA) over self-similar lattice

tilings. This includes the definitions of acceptable dilation, complete residue system and self-similarity. Then the Mallat reconstruction and decomposition algorithms are presented, followed by a discussion of local IFS on wavelet trees (LIFSW).

Chapter 2 contains the discoveries made during this exploration. The foundation of this work is grounded in the identification of a link between two-dimensional non-separable Haar wavelets and complex bases. This link is presented in Section 2.1. It is shown there that the link indeed generates an MRA. In Section 2.2, a notation is developed for the scaling and wavelet coefficients of an MRA associated with a complex base. Subsequently, it is demonstrated that this translation is consistent with the Mallat algorithm. A corollary of this result is that the decomposition algorithm of a function in an MRA associated with a complex base must terminate. Some preliminary work from this section was originally published in [66, 76].

In Section 2.3, the theory of complex bases and their link to wavelets are generalized to higher dimensional number systems. A definition of a number system is made and a number of the properties of complex bases are shown to hold for number systems, including an extension of the equivalence result of Gilbert. An investigation of diagonalizable operators is conducted in an attempt to generalize this result further, leading to a conjecture.

Tilings generated by number systems are typically fractal in nature. This often yields asymmetry in the wavelet trees of functions during wavelet decomposition. A notation is developed to discuss this situation. Subsequently, a class of extensions of functions is developed in Section 2.4. An algorithm for generating the extensions is given. This algorithm is shown to be well defined and the resulting extensions are shown to be consistent with the Mallat algorithm.

Section 2.5 presents a rigorous generalization of LIFSW to number systems. This section begins by discussing the issues caused by the fractal nature of the tilings. Notation is introduced to help in the discussion. As a result, an LIFSW associated with a number system is defined. Detailed explanations of the various aspects of the LIFSW operator are given. In Section 2.6, the inverse problem is studied for these operators, along with a method for constructing LIFSW approximations in higher dimensions.

The third chapter contains a few applications of the results to image compression. The definition of a discrete image is presented in the first section along with a short discussion on the comparison of images. The second section describes a general approach to image compression, including sample specific methods that are used (pruning, thresholding and LIFSW). The third section contains numerous examples of applying these methods to grey-scale images. These examples provide an interesting illustration of the behaviour of the fundamental tiles of complex bases.

Throughout the text, examples are provided to assist the reader in understanding key ideas and results. Typically, proofs are omitted for background results, although if the proofs are short or have not been found in the literature, they are sometimes included. Choice references for the various topics covered include the following:

Topology [49, 90]	Real Analysis [77, 89]
Functional Analysis [1, 2, 11, 18, 71, 72]	Matrix Analysis [8, 42]
Fractals and Self-Similarity [23, 43, 58, 63]	IFS [4, 5, 17]
IFSM, IFSW [27, 28, 29, 85, 86]	Image Coding [44, 73]
Complex Bases [30, 31, 32, 33, 34, 35]	Number Systems [50, 51, 55]
Wavelets [9, 15, 20, 62, 64, 68, 69, 87]	MRA [36, 37, 61]
Operator Theory and Wavelets [10, 19, 22, 48]	Tilings [3, 38, 52, 53, 56, 81]

An index is included at the end of the thesis to make it easier to find key definitions and concepts. Given the large number of definitions, notations and abbreviations used, a glossary and abbreviations list precede the index. Page numbers in **boldface** indicate the definition of the item or a significant reference to it. These three indexes will also facilitate the use of this work as a reference manual.

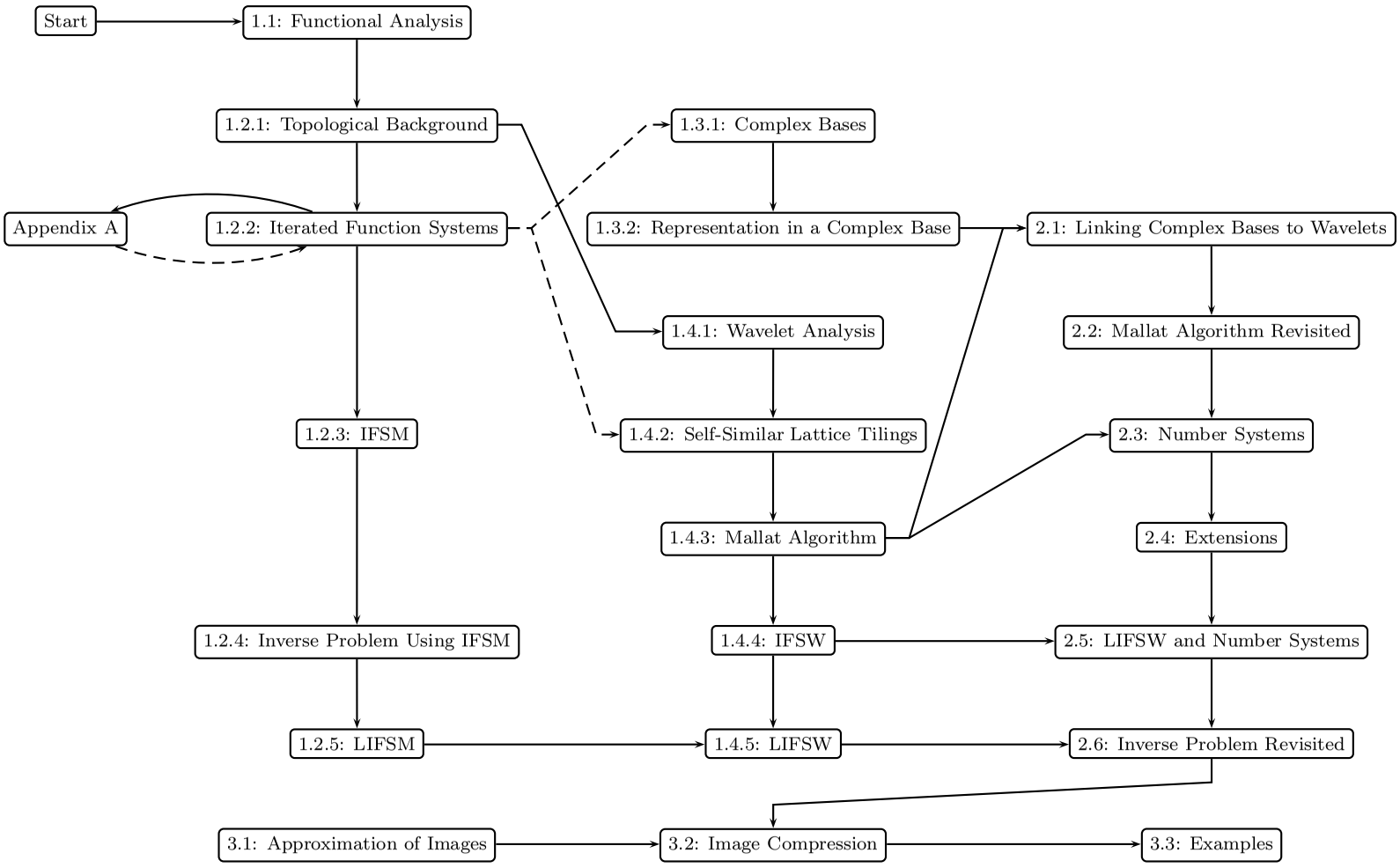
To assist the reader further in navigating the text, a diagram illustrating the dependencies of the various parts of the thesis is included after this introduction. A solid line pointing from a section indicates that reading this section is important for understanding the section to which the line points. A dashed line indicates that reading the section may be helpful for understanding the other section, but is not required.

After reading this thesis, one should discover that there are more questions left to answer than there are answers given. It is hoped that the ultimate contribution of this work will be the many seeds of future exploration that it can sow.

Daniel G. Piché

dgpiche@links.uwaterloo.ca

Section Dependencies



Chapter 1

Mathematical Background

This first chapter lays the foundation of background material to assist the reader through the results presented in the subsequent chapters.

1.1 Functional Analysis

This brief section presents a few definitions and results of functional analysis, principally for notational purposes. For general background in this topic, see [18, 77].

Definition 1.1.1 Let $L^2(\mathbb{R}^n)$ be the space of all square integrable functions from $\mathbb{R}^n$ to $\mathbb{R}$, that is, the space of all Lebesgue integrable functions $f : \mathbb{R}^n \to \mathbb{R}$, such that

$$\int_{\mathbb{R}^n} |f|^2 < \infty.$$

Definition 1.1.2 The *inner product* of two functions $f, g \in L^2(\mathbb{R}^n)$ is defined by $\langle f, g \rangle \equiv \int_{\mathbb{R}^n} fg$. The *norm* of f is $||f|| \equiv \sqrt{\langle f, f \rangle}$. Two functions f and g are *orthogonal* if their inner product is zero. This is denoted by $f \perp g$. A subset of $L^2(\mathbb{R}^n)$ is called *orthonormal* if its elements are pairwise orthogonal, and each has norm one.

Definition 1.1.3 A *basis* of $L^2(\mathbb{R}^n)$ is a maximal orthonormal set in $L^2(\mathbb{R}^n)$.

Theorem 1.1.4 If $\{f_n : n \in \mathbb{Z}\} \subset L^2(\mathbb{R}^n)$ is a basis of $L^2(\mathbb{R}^n)$, then each $f \in L^2(\mathbb{R}^n)$ can be written uniquely as $f = \sum_{n\in\mathbb{Z}} c_n f_n$, where $c_n = \langle f, f_n \rangle$.

Our study will involve specifically *wavelet bases*, which are described in Section 1.4. Further results of functional analysis will be presented there as needed.

1.2 Iterated Function Systems

An important concept in our study is that of iterated function systems (IFS). In this section, we present an overview of the major concepts and results of IFS and their application to the study of functions. A more detailed review of the topics in this section was given in [75]. That work contains detailed proofs of the material contained here.

1.2.1 Topological Background

We begin with some basic notation and definitions from metric spaces. Other topics not covered here may be found in [49, 90].

Notation 1.2.1 We will use the following notation to denote certain classical sets:

$\mathbb{N}$	=	$\{0, 1, 2, \ldots\}$;
$\mathbb{N}^+$	=	$\{1, 2, \ldots\}$;
$\mathbb{Z}$	=	the integers;
$\mathbb{R}$	=	the set of real numbers; and
$\mathbb{C}$	=	the set of complex numbers.

Notation 1.2.2 Throughout this section, (X, d) will denote a metric space where X is the set and d is the metric. Special properties, such as completeness, will be specified as needed. We will denote a sequence in X by $(x_n)_{n\in A}$, where $A \subset \mathbb{N}$. A sequence will be written as (x_n) if the range of the subscripts is clear from the context.

Definition 1.2.3 A metric space (X, d) is *complete* if and only if every Cauchy sequence converges in X with respect to the metric d.

Definition 1.2.4 A function $f : X \to X$ is said to be *Lipschitz* if and only if there exists an $s \in [0, \infty)$ such that $\forall x, y \in X$ we have

$$d(f(x), f(y)) \leq sd(x, y).$$

We call s a *Lipschitz constant* of f. If there exists such an $s < 1$, we say f is *contractive* or is a *contraction* and call s a *contractivity factor* of f. In this case we say that f has contractivity at least s. We write $Con(X, d)$ for the set of all contractive maps $f : X \to X$. If $X = \mathbb{R}^n$, write simply $Lip(\mathbb{R}^n)$ and $Con(\mathbb{R}^n)$.

Proposition 1.2.5 Let $f \in Con(X, d)$. Define c_f by

$$c_f = \inf\{s : s \text{ is a contractivity factor of } f\}.$$

Then c_f is a contractivity factor of f and is called the *contractivity* of f.

Contractive maps are the building blocks of IFS.

Example 1.2.6 Consider the function $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = \frac{1}{2}x + \frac{1}{2}$ $\forall x \in \mathbb{R}$. Then for $x, y \in \mathbb{R}$,

$$\begin{aligned} |f(x) - f(y)| &= \left| \left(\frac{1}{2}x + \frac{1}{2} \right) - \left(\frac{1}{2}y + \frac{1}{2} \right) \right| \\ &= \frac{1}{2}|x - y|. \end{aligned}$$

Therefore, f is contractive with contractivity $\frac{1}{2}$.

Notation 1.2.7 For $x \in X$, we define the *n-fold composition* of a function f at x recursively by

$$f^{\circ 1}(x) = f(x);$$
$$f^{\circ n+1}(x) = f(f^{\circ n}(x)).$$

We call $f^{\circ n}(x)$ the *n-th iterate of f at x*.[1] The symbol $\circ$ will also be used to denote the composition of two functions as follows: $v \circ w$.

Definition 1.2.8 We say that $y \in X$ is the *attractor* of $f : X \to X$ if and only if

$$\lim_{n\to\infty} f^{\circ n}(x) = y \quad \forall x \in X.$$

Example 1.2.9 Consider the function f from Example 1.2.6. Then for any $x \in \mathbb{R}$, we have

$$\begin{aligned} f(x) &= \frac{1}{2}x + \frac{1}{2}; \\ f^{\circ 2}(x) &= \frac{1}{2}\left(\frac{1}{2}x + \frac{1}{2}\right) + \frac{1}{2} \\ &= \frac{x}{2^2} + \frac{1}{2} + \frac{1}{4} \end{aligned}$$

and for a general $n > 1$,

$$f^{\circ n}(x) = \frac{x}{2^n} + \sum_{i=1}^{n} \frac{1}{2^i}.$$

Hence, $\lim_{n\to\infty} f^{\circ n}(x) = 1$ $\forall x \in \mathbb{R}$ and $x = 1$ is the attractor of f.

Definition 1.2.10 Let $f : X \to X$. If for some $x \in X$, $f(x) = x$, we call x a *fixed point* of f.

[1] We have chosen here to use the notation $f^{\circ n}$, instead of f^n, to remain consistent with the IFS literature. However, in the sections on complex bases and number systems, the $\circ$ will be omitted since function composition will be understood. There, we will use instead the notation Φ^n (for example, see Theorem 1.3.15).

Example 1.2.11 Consider the function f from Example 1.2.6. Then

$$\begin{aligned} f(x) = x &\implies \frac{1}{2}x + \frac{1}{2} = x \\ &\implies x = 1. \end{aligned}$$

Hence $x = 1$ is a fixed point of f.

This is not a coincidence, as the next proposition shows.

Proposition 1.2.12 If a continuous function $f : X \to X$ has an attractor $x \in X$, then x is a fixed point of f.

Proof Suppose $x \in X$ is the attractor of f. Then, since f is continuous,

$$x = \lim_{n\to\infty} f^{\circ n}(x) = f\left(\lim_{n\to\infty} f^{\circ n-1}(x)\right) = f(x). \qquad \blacksquare$$

The fundamental result upon which the entire theory of iterated function systems is founded is the *Banach Contraction Mapping Principle*, or *BCMP* for short [90].

Theorem 1.2.13 (Banach Contraction Mapping Principle) Suppose (X, d) is a complete metric space and let $f \in Con(X, d)$ with contractivity factor s. Then f has a unique fixed point $\bar{x}_f \in X$. Furthermore, $\bar{x}_f$ is the attractor of f.

If f is contractive, we write $\bar{x}_f$ to denote its fixed point.

Intuitively one would hope that if the given maps f and g are close to each other, then their respective fixed points $\bar{x}_f$ and $\bar{x}_g$ should also be close. This is indeed the case and is important for fractal-based methods of approximation. When one wants to describe a point by a function, for which the point is the attractor, it is often too difficult to describe the function completely. One must therefore choose a function close to the exact function that in turn will have its attractor being close to the original point of interest. The following result demonstrates this fact and is found in [17].

Theorem 1.2.14 (Continuity of Fixed Points) Define $F\colon Con(X,d) \to X$ by $F(f) = \bar{x}_f$ for each $f \in Con(X,d)$. Define also for $f, g \in Con(X,d)$ the functions

$$\bar{d}(f,g) = \sup_{x \in X} d(f(x), g(x))$$

and

$$d_m(f,g) = \min\{\bar{d}(f,g), 1\}.$$

Then F is continuous with respect to d_m and $d_m(f,g)$ is a metric on $Con(X,d)$. Furthermore, if (X,d) is compact, then F is continuous with respect to $\bar{d}$ and $\bar{d}$ is a metric on $Con(X,d)$.

Corollary 1.2.15 If (X,d) is a compact metric space and $f, g \in Con(X,d)$, then

$$d(\bar{x}_f, \bar{x}_g) < \frac{1}{1-c}\bar{d}(f,g),$$

where $c = \min(c_f, c_g)$.

Now, suppose that we are given $x \in X$. A natural question that was first asked in IFS theory is whether or not it is always possible to find a contractive operator $f \in Con(X,d)$ whose fixed point $\bar{x}_f$ is x. In simple cases one can guess at such a function (as in Example 1.2.6). One can also take the constant function $f(y) = x$ for all $y \in X$. As mentioned above, the goal is to approximate x using a function which is easy to describe, and a constant function would often require the complete description of x. However, this defeats the purpose of trying to express x – in possibly a more concise manner – as the fixed point of a contractive operator f. We also expect that, in general, this is not possible and that one must be satisfied in finding fixed points $\bar{x}_i$ of contractive operations f_i that are *approximations* to x. Even in this case, however, we are faced with the problem of

finding such fixed points $\bar{x}_i$.

This problem is called the *Inverse Problem of Approximation by Fixed Points of Contraction Maps*, or the *Inverse Problem* for short. It is generally stated as follows:

Question 1.2.16 Given (Y, d_Y) a metric space, $y \in Y$ and $\epsilon > 0$, can we find a non constant $f \in Con(Y, d_Y)$ such that $d_Y(y, \bar{y}_f) < \epsilon$?

Detailed discussions can be found in [29] and [84]. Indeed, whether such an f can be constructed, or whether it even exists is uncertain at this stage. In other words, the question is "Is $\{\bar{x}_f : f \in Con(Y, d_Y)\}$ dense in Y"? In practice, Y could be any one of a large number of relevant spaces: compact subsets of $[0, 1]^n$; probability measures on $[0, 1]$; $L^p(\mathbb{R})$; fuzzy set functions. These questions do not have easy answers.

Before commenting on this question, an additional question that arises is: "Given $y \in Y$, $f \in Con(Y, d_Y)$, how close is y to $\bar{y}_f$"? The following proposition lends an answer:

Proposition 1.2.17 Let y, Y and f be as above. Then

$$d_Y(y, \bar{y}_f) \leq \frac{1}{1 - c_f} d_Y(y, f(y)).$$

This is often called the *Collage Theorem*. It is important in helping identify the functions to use in an IFS in order to approximate the attractor.

The Collage Theorem is fundamental to the theory of IFS because it states that if $f(y)$ is close to y, then $\bar{y}_f$ is also close to y. Of course, if $c_f \approx 1$, the right hand side of the inequality might not be very small. Thus, this gives some insight into finding our desired function. We should find an $f \in Con(Y, d_Y)$ which takes y close to itself. We remember from the BCMP that $\bar{y}_f$ is the attractor of f if Y is complete. Hence we can iterate f to retrieve $\bar{y}_f$ and get the desired approximation to y. Therefore, the Inverse Problem is often

formulated as follows:

Question 1.2.18 (Inverse Problem) Let (Y, d_Y) be a complete metric space, and let $y \in Y$. Given $\epsilon > 0$, can we find a non constant $f \in Con(Y, d_Y)$ such that $d_Y(y, f(y)) < \epsilon$?

A formal solution to this problem was given in [28] in the case of IFS on grey-level maps. This will be important in our study of approximations of images and is discussed in Section 1.2.4.

1.2.2 Iterated Function Systems

The concept we wish to present in this section is that of iterated function systems (IFS). These were first developed by Hutchinson [43]. They were independently discovered by Barnsley and Demko [5] who gave them their name. We give a brief introduction here and again refer the reader to [75] for a more comprehensive discussion.

The following definition of an iterated function system, or IFS is found in [4]. A motivation for this definition is given in Appendix A.

Definition 1.2.19 An *iterated function system*, or *IFS*, consists of a complete metric space (X, d) together with a finite set of contraction mappings w_n: $X \to X$ with respective contractivity factors c_n, $n = 1, 2, \ldots, N$.[2] Such an IFS, denoted by $\mathbf{w}$, where $\mathbf{w}= \{w_n : n = 1, 2, \ldots, N\}$, is called an *N-map IFS*. The IFS is said to have *contractivity* $c = \max\{c_n : n = 1, 2, \ldots, N\}$.

The contractivity of an IFS is meaningful in a certain space associated with (X, d). A brief discussion is presented in Appendix A.

Barnsley and Demko [5] first noticed that a natural object such as a leaf could be approximated as the attractor of a rather simple IFS. Indeed, the appendix of [5] contains

[2]This definition can be extended to an infinite number of maps [58].

the first seeds of the inverse problem for IFS. The next significant demonstration appeared in [7], where Barnsley and co-workers presented a striking computer-generated picture of a spleenwort fern. As well, they presented the "collage theorem" as a method of solving the inverse problem.

The next major advance was the paper of Jacquin in which a fractal block coding method for images was proposed. These papers spawned a great deal of research on the possible use of IFS methods for image coding and compression – so-called "fractal image compression". Accounts of this research can be found in the books by Barnsley [4], Barnsley and Hurd [6], Fisher [26, 25] and Lu [59].

Before continuing, we note the following classes of functions from which useful IFS maps are often chosen:

Notation 1.2.20 Define the following three sets:

$$Con_1(X,d)=\{w \in Con(X,d) : w \text{ is 1-1}\};$$
$$Sim(X,d) =\{w : d(w(x),w(y)) = cd(x,y) \text{ for some } c \in [0,1], \forall x,y \in X\}; \text{ and}$$
$$Sim_1(X,d)=Sim(X,d) \cap Con_1(X,d).$$

The mathematical theory behind the application of IFS methods to grey-level (grey-scale) images was developed by Forte and Vrscay [27, 28, 29]. They formulated an IFS-type method which allows the construction of grey-scale images. A good deal of work, however, was necessary to arrive at this result. For example, they began by extending the concept of IFS on sets (defined earlier in [5, 43]) to IFS on characteristic functions of sets – the image functions used here were simple black-and-white (bitmap) images. The range of these image functions were then extended from [0,1] to $\mathbb{R}$. As well, the probabilities associated with the IFS maps were replaced by grey-level maps that operated on the image function

values. However, the metric first used on the image functions was the very restrictive metric of *fuzzy sets* (Hausdorff metric between level sets of two image functions). This metric was then weakened by looking at the measure of symmetric differences of the level sets. Forte and Vrscay then showed that in the special case that the measure was the Lebesgue measure, the L^1 metric was obtained. The next natural step was to extend the IFS method to L^p. The reader is referred to [27, 28, 29] for detailed discussions.

Consider the class of *grey-level images* supported on a metric space X. We write $\mathcal{F}(X)$ for the set of grey-level images on X. In practical applications, images are usually considered to be nonnegative functions so that their range will be defined as $\mathbb{R}^+ = [0, \infty)$ or some appropriate subset. However, in more general mathematical treatments (e.g. [28]), the range of $\mathcal{F}(X)$ can be extended to $\mathbb{R}$. The set X is thought of as the "pixel space", usually I or I^2 with the usual Euclidean metric.

Definition 1.2.21 Let (X, d) be a metric space. Let $\mathbf{w} = \{w_k\}_{k=1}^N$ where $w_k \in Con_1(X, d)$ for $k = 1, 2, \ldots, N$. Then, let $\mathbf{\Phi} = \{\phi_k\}_{k=1}^N$ where $\phi_k : \mathbb{R} \to \mathbb{R}$ for $k = 1, 2, \ldots, N$. The pair $(\mathbf{w}, \mathbf{\Phi})$ is called an *iterated function system with grey-level maps*, or *IFSM* for short.

Define the IFSM operator $T_{(\mathbf{w},\mathbf{\Phi})}: \mathcal{F}(X) \to \mathcal{F}(X)$, on $u \in \mathcal{F}(X)$, by

$$T_{(\mathbf{w},\mathbf{\Phi})}u(x) = \sum_{k=1}^{N}{}' \phi_k(u(w_k^{-1}(x))) \quad \forall x \in X, \tag{1.1}$$

where $\sum'$ indicates that the sum runs over the indices k with $x \in w_k(X)$. We use the convention that an empty sum has a value of 0.

Notice that when u is displaced in space, its grey-level values are also modified. In other words, the action of the operator T on an image function u is to produce a set of N copies of u over smaller regions $w_i(X)$. Furthermore, the grey level values of each of these copies are modified by the grey-level maps ϕ_i. An example of this procedure is presented

in Example 1.2.22. For simplicity, the w_i and ϕ_i are usually taken to be affine maps. That is, $w_i(x) = s_i x + a_i$ and $\phi_i(t) = \alpha_i t + \beta_i, x \in X, t \in \mathbb{R}$.

With appropriate conditions on the IFSM (w, ϕ), the operator T, often called a *fractal transform operator*, is contractive on $\mathcal{F}(X)$. For example, if $\mathcal{F}(X) = L^p(X)$, then a sufficient condition for contractivity is that $\sum_i |s_i \alpha_i^p| < 1$. Hence, by the Banach Contraction Mapping Principle, T has a unique fixed point function $\bar{u} = T\bar{u}$.

Example 1.2.22 To demonstrate the action of an IFSM operator T, consider the following function:

$$u(x) = e^{\left(-\frac{1}{2}x\right)} \cos(\sin(5x))$$

on the unit interval $[0, 1]$. It is illustrated in Figure 1.1 on page 19. Now, consider an IFSM with two IFS component maps

$$w_1(x) = 0.6x, \quad w_2(x) = 0.6x + 0.4$$

with grey-scale maps

$$\phi_1(x) = 0.75x, \quad \phi_2(x) = 0.5x + 0.5.$$

Let us decompose the action of T on u. First, the two "component" curves for the IFSM are given by $u(w_1^{-1}(x))$ and $u(w_2^{-1}(x))$. In each case, the domain of $u \circ w_i^{-1}$ is $w_i([0, 1])$. Here, $w_1([0, 1]) = [0, 0.6]$ and $w_2([0, 1]) = [0.4, 1]$. The component functions are therefore

$$u(w_1^{-1}(x)) = e^{\left(-\frac{5}{6}x\right)} \cos\left(\sin\left(\frac{25}{3}x\right)\right), \qquad x \in [0, 0.6]$$

and

$$u(w_2^{-1}(x)) = e^{\left(-\frac{5}{6}x + \frac{1}{3}\right)} \cos\left(\sin\left(\frac{25}{3}x - \frac{10}{3}\right)\right), \quad x \in [0.4, 1].$$

These component functions are shown in the top right corner of Figure 1.1. The component functions are copies of u, scaled and translated along the horizontal axis. Next, the

component functions are modified by the grey-level maps as follows:

$$\phi_1(u(w_1^{-1}(x))) = \frac{3}{4}e^{\left(-\frac{5}{6}x\right)}\cos\left(\sin\left(\frac{25}{3}x\right)\right), \qquad x \in [0, 0.6]$$

and

$$\phi_2(u(w_2^{-1}(x))) = \frac{1}{2}e^{\left(-\frac{5}{6}x+\frac{1}{3}\right)}\cos\left(\sin\left(\frac{25}{3}x - \frac{10}{3}\right)\right) + \frac{1}{2}, \quad x \in [0.4, 1].$$

These modified component functions, which are the component functions scaled along the vertical axis, are shown in the bottom left corner of Figure 1.1. Finally, the modified component functions are added together to give the action of T on u, as illustrated in the fourth image in Figure 1.1.

1.2.3 IFSM on $L^p(X, \mu)$

Let $(\mathbf{w}, \mathbf{\Phi})$ be an IFSM on a complete metric space (X, d) where $\mathbf{w} = \{w_1, w_2, \ldots, w_N\}$, $w_k \in Con_1(X)$ and $\mathbf{\Phi} = \{\phi_1, \phi_2, \ldots, \phi_N\}$, $\phi_k : \mathbb{R} \to \mathbb{R}$. When it is understood that a specific IFSM is being considered, write T for $T_{(\mathbf{w},\mathbf{\Phi})}$, to denote the associated IFSM operator.

The theory of IFSM was developed for $L^p(X, \mu)$ by Forte and Vrscay in [28]. [3] We present a few of their results here.

Proposition 1.2.23 Let $(\mathbf{w}, \mathbf{\Phi})$ be an N-map IFSM and let T be the associated IFSM operator. Suppose:

i) $\forall u \in L^p(X, \mu), u \circ w_k^{-1} \in L^p(X, \mu), 1 \leq k \leq N$; and

ii) $\phi_k \in Lip(\mathbb{R}), 1 \leq k \leq N$.

[3] The space $L^p(X, \mu)$ is the Banach space of all p-integrable functions on the measure space X with measure μ and norm $||f||_{L^p} = (\int |f|^p)^{1/p}$.

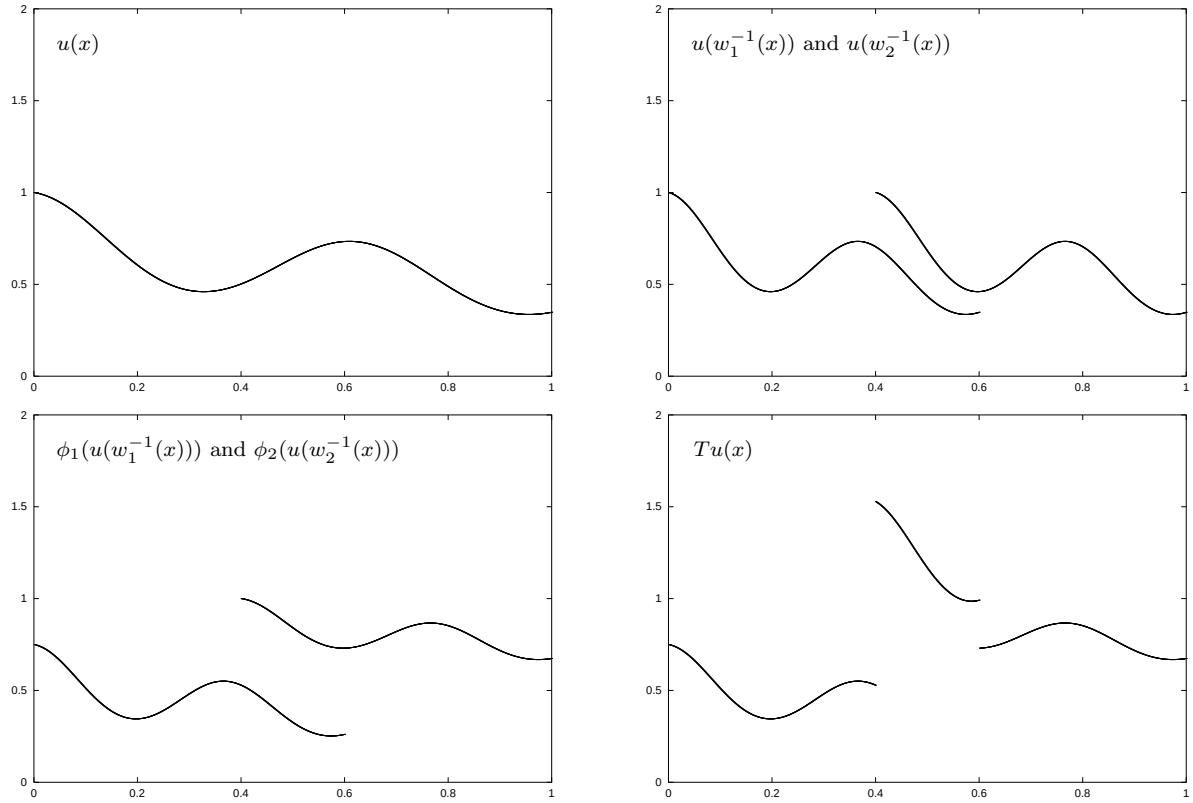


Figure 1.1: Illustration of the action of an IFSM operator T on a function u. The original function is shown in the top left corner. To its right are the component functions. The image on the bottom left shows the component functions modified by the grey-level maps. The last image shows the action of T on u.

Then for $1 \leq p \leq \infty, T : L^p(X,\mu) \to L^p(X,\mu)$.

The following shows that under certain conditions $T_{(\mathbf{w},\mathbf{\Phi})}$ is contractive. Let $M(X)$ denote the set of finite measures on $B(X)$, the Borel sets of X.

Proposition 1.2.24 Let $(\mathbf{w},\mathbf{\Phi})$ be an N-map IFSM such that $\phi_k(t) = a_k \in \mathbb{R}$ for all $t \in \mathbb{R}$ and $1 \leq k \leq N$. Then $\forall p \in [1,\infty)$ and $\mu \in M(X)$, the associated IFSM operator T is contractive on $L^p(X,\mu)$, with contractivity factor $c_T = 0$. Furthermore, its fixed point $\bar{u}_T$ is

$$\bar{u}_T = \sum_{k=1}^{N} a_k \chi_{\hat{w}_k(X)}.$$

Proposition 1.2.25 Let $X \subset \mathbb{R}^D, D \in \mathbb{N}^+$, and let $\mu = m^{(D)}$ be the Lebesgue measure on $\mathbb{R}^D$ and let d be the usual Euclidean metric. Let $(\mathbf{w},\mathbf{\Phi})$ be an N-map IFSM such that, for $1 \leq k \leq N$:

i) $w_k \in Sim_1(X,d)$ is a similarity with contractivity factor c_k; and

ii) $\phi_k \in Lip(\mathbb{R})$, with Lipschitz constant K_k.

Then for $p \in [1,\infty)$ and $u, v \in L^p(X,\mu)$, we have

$$||Tu - Tv||_p \leq C(D,p)||u - v||_p,$$

where

$$C(D,p) = \sum_{k=1}^{N} c_k^{D/p} K_k.$$

Hence, if $C(D,p) < 1$, T is contractive on $L^p(X,\mu)$ and has a unique, attracting fixed point. It is not necessary that all of the IFS maps be contractive (in the base space X) for T to be contractive. The contractivity of the ϕ_k (in the grey-level range) can contribute in this aspect [28].

Example 1.2.26 Let $X = [0,1]$ and μ be the Lebesgue measure on X. Let $w_i(x) = \frac{1}{3}(x + i - 1), i = 1,2,3$. Let $\phi_1(t) = \frac{1}{2}t, \phi_2(t) = \frac{1}{2}, \phi_3(t) = \frac{1}{2}t + \frac{1}{2}$, for $t \in \mathbb{R}$. The fixed point

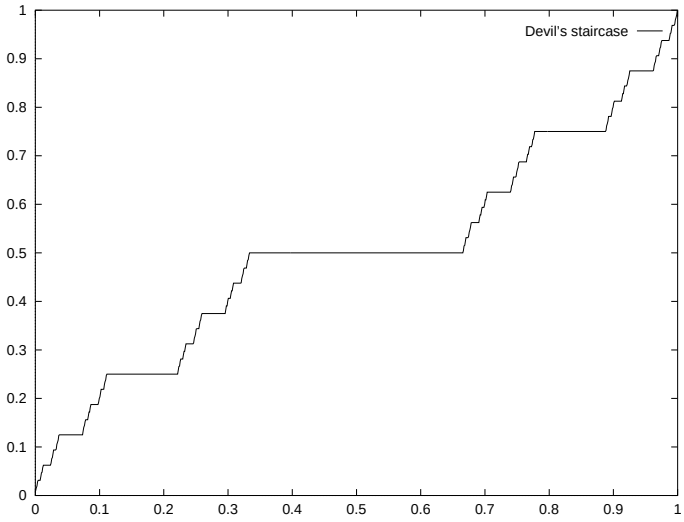


Figure 1.2: The Devil's staircase. This is also the distribution function $F(x) = \int_0^x d\mu$ of the Cantor-Lebesgue measure μ.

of this IFSM is the *Devil's staircase*. It is constant almost everywhere on X, increasing on $\mathcal{C}$, yet its image is all of X. The attractor $\bar{u}$ is shown in Figure 1.2.

Given two N-map IFSM $(\mathbf{w}, \mathbf{\Phi}_i)$, $i = 1, 2, \ldots, N$, where $\mathbf{\Phi}_i = \{\phi_{i1}, \phi_{i2}, \ldots, \phi_{iN}\}$, define the distance between the grey-level components by

$$d^N_{\mathbf{\Phi}}(\mathbf{\Phi}_1, \mathbf{\Phi}_2) = \sup_{1 \le k \le N} \sup_{t \in \mathbb{R}} |\phi_{1k}(t) - \phi_{2k}(t)|.$$

This function will be a metric when X is compact. The following result from [28] establishes the continuity of fixed points for IFSM (c.f. Theorem 1.2.14).

Proposition 1.2.27 Let $(\mathbf{w}, \mathbf{\Phi}_1)$ be an N-map IFSM with fixed point $\bar{u}_1 \in L^p(X, \mu)$. Then given $\epsilon > 0$, $\exists \delta > 0$ such that for all N-map IFSM $(\mathbf{w}, \mathbf{\Phi}_2)$ with $d^N_{\mathbf{\Phi}}(\mathbf{\Phi}_1, \mathbf{\Phi}_2) < \delta$, then $||\bar{u}_1 - \bar{u}_2||_p < \epsilon$, where $\bar{u}_2$ is the fixed point of $(\mathbf{w}, \mathbf{\Phi}_2)$.

1.2.4 Inverse Problem Using IFSM

In this section we present the formal solution to the Inverse Problem for IFSM. Consider the following formulation for IFSM:

Question 1.2.28 For $v \in L^p(X, \mu)$ and $\epsilon > 0$, can we find an IFSM $(\mathbf{w}, \mathbf{\Phi})$ with associated operator T such that $||v - Tv||_p < \epsilon$?

A formal solution was obtained in [28] by constructing sequences of N-map IFSM $(\mathbf{w}^N, \mathbf{\Phi}^N)$, $N = 1, 2, 3, \ldots$, where $\mathbf{w}^N$ is chosen from a fixed, infinite set $\mathcal{W}$ of contraction maps.

Definition 1.2.29 Let $\mathcal{W} = \{w_1, w_2, \ldots\}$ be an infinite set of contraction maps on X. Then $\mathcal{W}$ *generates a μ-dense and non-overlapping*, or μ-d-n, family of subsets of X if $\forall \epsilon > 0$ and $\forall B \subset X$, there exists a finite set of integers $i_k \geq 1$, $1 \leq k \leq N$ such that:

i) $A = \cup_{k=1}^{N} w_{i_k}(X) \subset B$;

ii) $\mu(B \backslash A) < \epsilon$; and

iii) $\mu(w_{i_k}(X) \cap w_{i_l}(X)) = 0$, whenever $k \neq l$.

Example 1.2.30 Let $X = [0, 1]$ with Lebesgue measure. Let $w_{ij}(x) = 2^{-i}(x + j - 1)$, $i = 1, 2, \ldots$, $1 \leq j \leq 2^i$. For each $i \geq 1$, the set of maps $\{w_{ij}, 1 \leq j \leq 2^i\}$ is a set of 2^i contractions of $[0, 1]$ which tile $[0, 1]$. Then $\mathcal{W} = \{w_{ij}\}$ is μ-d-n. Figure 1.3 illustrates the idea.

Now, suppose $\mathcal{W} = \{w_i\}$, with $w_i \in Con_1(X, d)$, generates a μ-d-n family of subsets of X. Let

$$\mathbf{w}^N = \{w_1, w_2, \ldots, w_N\}, \quad N = 1, 2, \ldots,$$

denote the N-map truncations of $\mathcal{W}$. Assume that for each $k \in \mathbb{N}^+$, $\phi_k \in Lip(\mathbb{R})$ is the associated grey-level map of w_k, and let

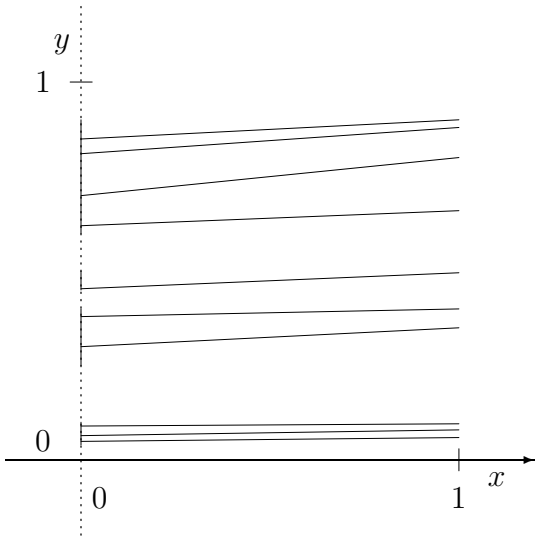


Figure 1.3: Illustration of the sets A and B of the formal solution to the inverse problem. The set B is the union of the solid lines on the vertical axis. The set A is the union of the lines (projected onto the vertical axis).

$$\mathbf{\Phi}^N = \{\phi_1, \phi_2, \ldots, \phi_N\}.$$

Let $T^N : L^p(X,\mu) \to L^p(X,\mu)$ be the associated IFSM operator of $(\mathbf{w}^N, \mathbf{\Phi}^N)$. Then the following result holds:

Theorem 1.2.31 Let $v \in L^p(X,\mu)$, $1 \leq p < \infty$ and $\mathcal{W}$ as above. Then

$$\lim_{N\to\infty} \inf ||v - T^N v||_p = 0. \tag{1.2}$$

Proof A proof can be found in [28]. ■

This result and Example 1.2.30 can be used to build an algorithm for the construction of IFSM approximations of target functions $v \in L^p(X,\mu)$. Then, given an N-map IFSM $(\mathbf{w}, \mathbf{\Phi})$ on (X,d) with associated operator T, we have the squared L^2 collage distance

$$\Delta^2 = ||v - Tv||_2^2$$

$$= \int_X \left(\sum_{k=1}^{N}{}' \phi_k(v(w_k^{-1}(x))) - v(x) \right)^2 d\mu(x). \tag{1.3}$$

With the formal solution (1.2) in mind, we assume the IFS maps w_k are fixed and search for grey-level maps ϕ_k which minimize Δ^2 for the given target v. This is the key idea for IFSM [28].

For computational simplicity, assume the maps w_k and ϕ_k are affine. The pair $(\mathbf{w}, \mathbf{\Phi})$ will be called an *affine IFSM*. Assuming that $\phi_k(t) = \alpha_k t + \beta_k \ \forall t \in \mathbb{R}, k = 1, 2, \ldots, N$, then

$$Tu(x) = \sum_{k=1}^{N}{}' \left[\alpha_k u(w_k^{-1}(x)) + \beta_k \chi_{w_k(X)}(x) \right]. \tag{1.4}$$

If $X \subset \mathbb{R}^D$, then by Proposition 1.2.25 on page 20, $\forall u, v \in L^p(X, \mu)$,

$$||Tu - Tv||_p \leq C(D, p)||u - v||_p$$

with $C(D, p) = \sum_{k=1}^{N} c_k^{D/p} \alpha_k$, and c_k is the contractivity of w_k. Hence, if $C(D, p) < 1$, T is contractive on $L^p(X, \mu)$ and has a unique fixed point $\bar{u}_T$.

Example 1.2.32 If $\beta_k = 0$ for $1 \leq k \leq N$, then $\bar{u}_T \equiv 0$.

Example 1.2.33 If $X = [0, 1]$, $w_k(x) = a_k x + b_k, a_k \neq 0, 1 \leq k \leq N$, and T is contractive with fixed point $\bar{u}_T$, then by Equation (1.4),

$$\begin{aligned} \bar{u}_T(x) &= \sum_{k=1}^{N}{}' \alpha_k \bar{u}_T \left(\frac{x - b_k}{a_k} \right) + \beta_k \chi_{w_k(X)}(x) \\ &= \sum_{k=1}^{N}{}' \alpha_k \psi_k(x) + \beta_k \phi_k(x). \end{aligned}$$

Therefore, $\bar{u}_T$ is a linear combination of piecewise constant functions ϕ_k, and functions

ψ_k, which are dilations and translations of $\bar{u}_T$. In general, if $f : \mathbb{R} \to \mathbb{R}$, then given $a \in \mathbb{R}$, $f(a\cdot)$ is called the *dilation* of f by a, and $f(\cdot - a)$ is called the *translation* of f by a. This idea is reminiscent of the wavelets relations (see Section 1.4).

By the following theorem, it is sufficient in practical situations to study the subclass of affine IFSM [28]. This makes the generation of IFSM much simpler in practice.

Theorem 1.2.34 Let $X = \mathbb{R}^D$ and let $\mu \in M(X)$. Given $p \geq 1$, let $L^p_A(X, \mu) \subset L^p(X, \mu)$ be the set of fixed points of contractive N-map affine IFSM on X. Then $L^p_A(X, \mu)$ is dense in $L^p(X, \mu)$.

Now, suppose $(\mathbf{w}, \mathbf{\Phi})$ is an N-map affine IFSM with:

i) $w_k \in Con_1(X, d)$ with contractivity factors $c_k > 0$ for $1 \leq k \leq N$;

ii) $\cup_{k=1}^N w_k(X) = X$; and

iii) $\phi_k : \mathbb{R} \to \mathbb{R}$, where $\phi_k(t) = \alpha_k t + \beta_k$, $t \in \mathbb{R}$, $1 \leq k \leq N$.

The squared L^2 collage distance of Equation (1.3) becomes

$$\begin{aligned}\Delta^2 =& < v - Tv, v - Tv > \\ =& \sum_{k=1}^{N}\sum_{l=1}^{N}(< \psi_k, \psi_l > \alpha_k\alpha_l + 2 < \psi_k, \chi_l > \alpha_k\beta_l + < \chi_k, \chi_l > \beta_k\beta_l) \\ & - 2\sum_{k=1}^{N}(< v, \psi_k > \alpha_k + < v, \chi_k > \beta_k) + < v, v >,\end{aligned}$$

where $\psi_k = v \circ w_k^{-1}$ and $\chi_k = \chi_{w_k(X)}$. Then Δ^2 can be written as a *quadratic form* in the parameters α_k and β_k as

$$\Delta^2 = \mathbf{x}^T A\mathbf{x} + \mathbf{b}^T\mathbf{x} + c,$$

where $\mathbf{x}^T = (\alpha_1, \ldots, \alpha_N, \beta_1, \ldots, \beta_N) \in \mathbb{R}^{2N}$. Minimizing Δ^2 is a *quadratic programming* (QP) problem in the α_k and β_k. A detailed discussion is given in [28, 84].

We end this section with a few examples of approximations which demonstrate the application of IFSM. The second example reveals a shortcoming of this method.

Example 1.2.35 Let $u(x) = \sin(x)$ and $X = [0, 1]$. The approximations of u are given in Figure 1.4 on the next page, with the ratio $1 : N$ indicating the number of maps w_k used. The maps w_k map X to evenly divided subintervals. For example, consider 2 maps, $w_1(x) = x/2$ and $w_2(x) = x/2 + 1/2$. Table 1.1 gives the L^2 distance between u and the approximations. In this case, the results are quite nice. This happens since the parts of the function on the subintervals are similar to the function on the entire interval.

Number of maps	Distance	File size (bytes)	Computation time (sec.)
u	0.0	30878	n/a
2	0.0199362	42	0.120
4	0.0191687	82	0.130
16	0.0188445	320	0.140

Table 1.1: IFSM compression of $\sin(x)$.

Example 1.2.36 Let $u(x) = \sin(\pi x)$ and $X = [0, 1]$. Consider the approximations in Figure 1.5 on page 28. Table 1.2 on page 29 gives the L^2 distance between u and the approximations. Here, it is difficult to get a good approximation since the function lacks self-similarity on the subintervals. The problem described arises in most situations since one cannot hope to have global self-similarity, that is subsets of an image are not similar to the entire image but may be to other subsets. This situation can be remedied by considering local IFSM, which are discussed next.

In both of the above examples, and Example 1.2.39 on page 32, computations were performed using unoptimized C code compiled and run on a Mac G3 500MHz with 384MB RAM. Computation times were quite small and varied between 0.100 and 0.200 seconds

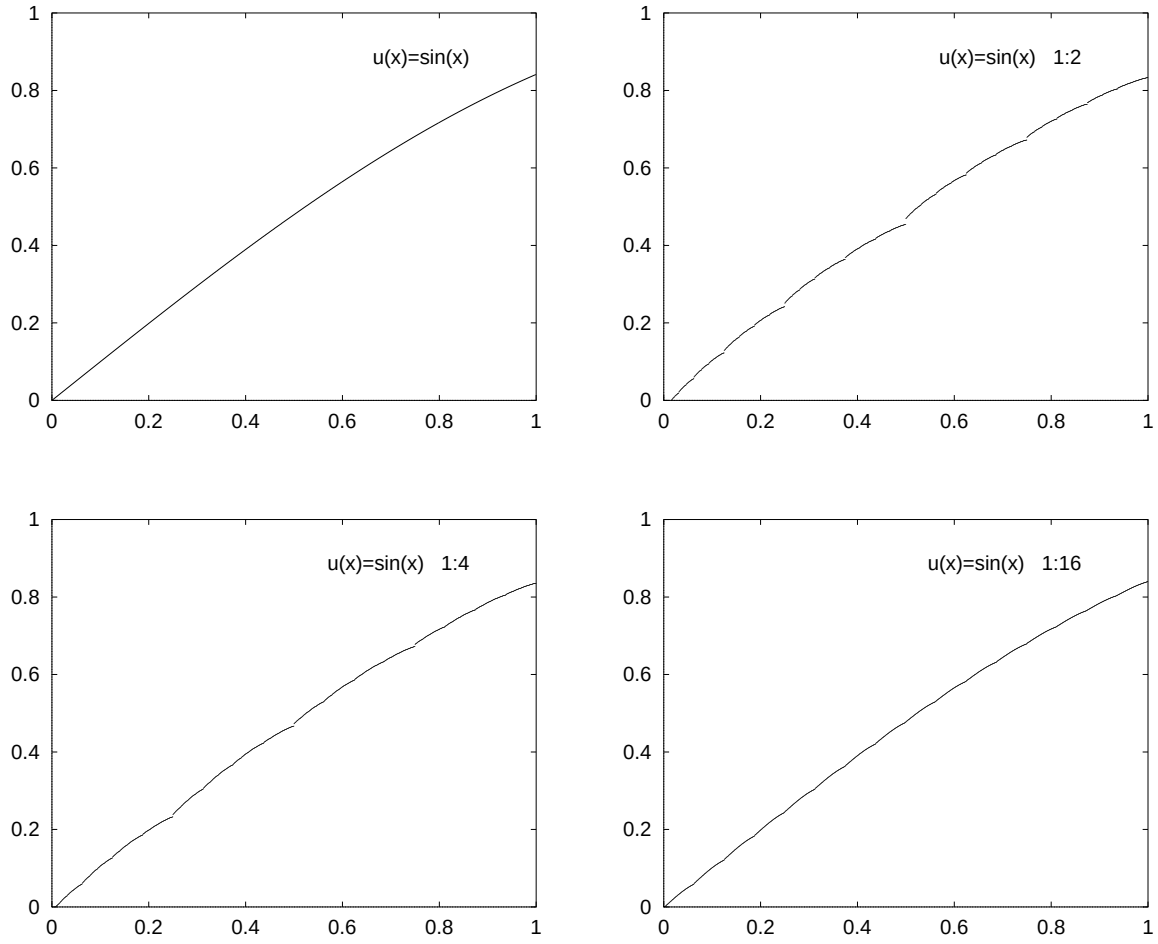


Figure 1.4: IFSM approximations of $u(x) = \sin(x)$.

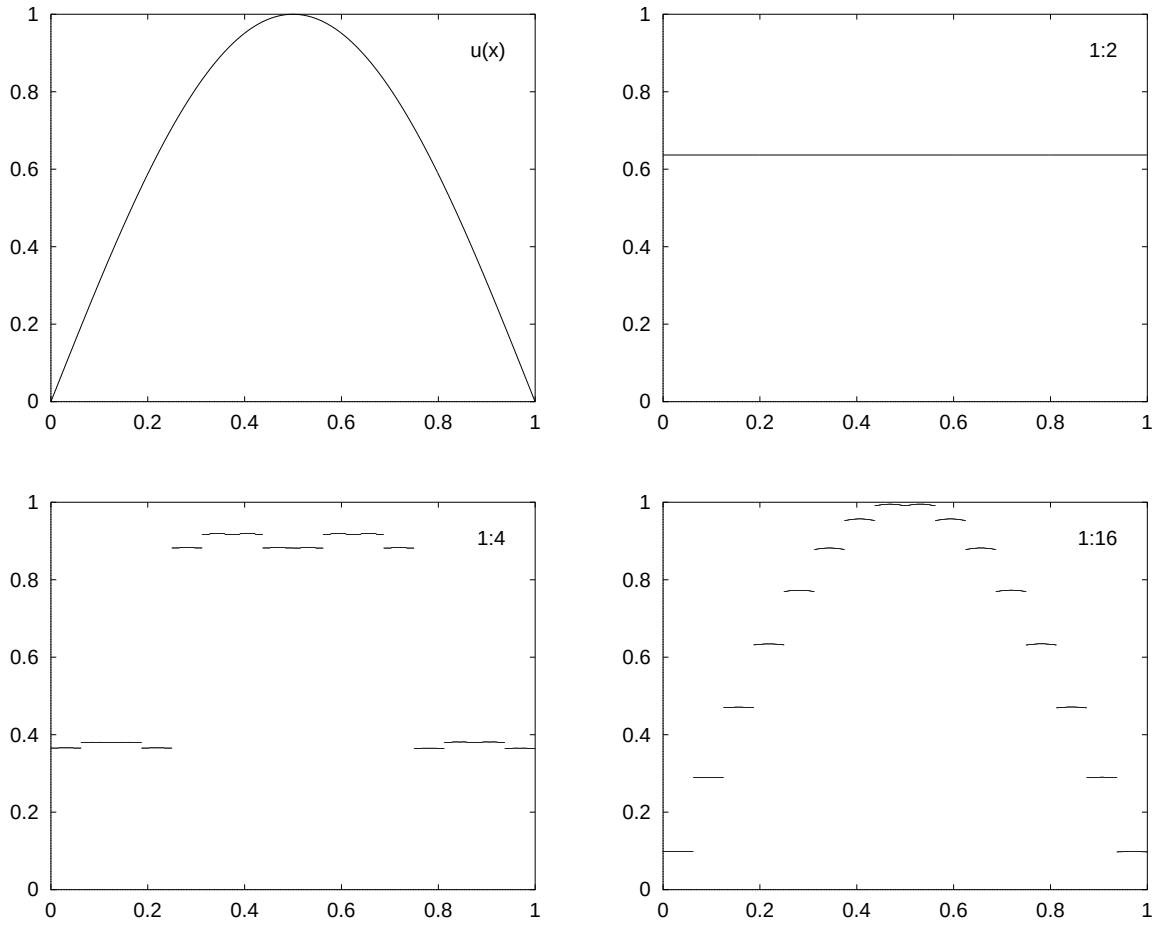


Figure 1.5: IFSM approximation of $u(x) = \sin(\pi x)$ with 2, 4 and 16 range blocks.

Number of maps	Distance	File size (bytes)	Computation time (sec)
u	0.0	30768	n/a
2	0.307759	39	0.130
4	0.158104	82	0.140
16	0.0400456	339	0.160

Table 1.2: IFSM compression of $\sin(\pi x)$.

depending on processor availability at the time, the same computation ranging across the entire interval on sequential executions. It is interesting to note that this represents a ten-fold improvement over computations done with these same examples in 1997 on a DEC Alpha.

1.2.5 LIFSM

A method that, in general, yields better approximations than the IFSM is the method of local IFSM (LIFSM) [28]. In this method, the fractal transform operator T constructs altered copies of *subsets* of the image function u, in contrast to the IFSM method where altered copies of the *entire* image function u are produced.

Definition 1.2.37 Let $X \subset \mathbb{R}^D$ and $\mu = m^{(D)}$. Let $J_k \subset X$, $k = 1, 2, \ldots, N$, such that

i) $\cup_{k=1}^{N} J_k = X$ (covering condition); and

ii) $\mu(J_j \cap J_k) = 0$ when $j \neq k$ (μ-non-overlapping condition).

Suppose also that $\forall J_k$, $\exists I_{j(k)} \subset X$ with an associated map $w_{j(k),k} \in Con(X, d)$ with contractivity factor $c_{j(k),k}$ such that

$$w_{j(k),k}(I_{j(k)}) = J_k.$$

The set J_k is called the *range block* of the *domain block* $I_{j(k)}$.[4] For each $w_{j(k),k}$, let ϕ_k :

[4]Often, one calls the domain blocks "parent" blocks and the range blocks "child" blocks.

$\mathbb{R} \to \mathbb{R}$ be an associated grey-level map. Then define

$$\mathbf{w}_{loc} = \{w_{i(1),1}, \ldots, w_{i(N),N}\} \quad \text{and} \quad \mathbf{\Phi} = \{\phi_1, \ldots, \phi_N\}.$$

The pair $(\mathbf{w}_{loc}, \mathbf{\Phi})$ is called an N-map *local IFSM*, or *LIFSM*. The associated operator $T^{loc}_{(\mathbf{w},\mathbf{\Phi})}\colon \mathcal{F}(X) \to \mathcal{F}(X)$ is defined by

$$T^{loc}_{(\mathbf{w},\mathbf{\Phi})}u(x) = \begin{cases} \phi_k(u(w^{-1}_{i(k),k})) & x \in J_k \setminus \cup_{l\neq k}^{N} J_l(X), \\ 0 & \text{otherwise.} \end{cases}$$

A result similar to Proposition 1.2.25 can then be obtained:

Proposition 1.2.38 Let $X \subset \mathbb{R}^D$ and $\mu = m^{(D)}$. Let $(\mathbf{w}_{loc}, \mathbf{\Phi})$ be an LIFSM as above with $\phi_k \in Lip(\mathbb{R})$ for $1 \leq k \leq N$ and let T^{loc} be the associated LIFSM operator. Then, for $u, v \in L^p(X, \mu)$,

$$||T^{loc}u - T^{loc}||_p \leq C_{loc}(D,p)||u - v||_p,$$

with $C_{loc}(D,p) = \left(\sum_{k=1}^{N} c_{j(k),k} K_k^p\right)^{1/p}$, where $c_{j(k),k}$ is the contractivity of $w_{j(k),k}$. Thus, if $C_{loc}(D,p) < 1$, T^{loc} is contractive on $L^p(X,\mu)$ and has a unique fixed point.

Now, suppose $X \subset \mathbb{R}^D$, $\mu = m^{(D)}$ and $v \in L^2(X,\mu)$. Then, given an N-map LIFSM as above, the squared collage distance is given by

$$\begin{aligned} \Delta^2 &= ||T^{loc}v - v||_2^2 \\ &= \sum_{k=1}^{N} \int_{J_k} [\phi_k(v(w^{-1}_{j(k),k}(x))) - v(x)]^2 dx \\ &= \sum_{k=1}^{N} \Delta^2_{j(k),k}. \end{aligned}$$

It is therefore sufficient to minimize the $\Delta^2_{j(k),k}$ individually for each range block J_k. In the case where the maps ϕ_k are affine, this becomes a QP problem [28].

To apply this idea to the Inverse Problem, consider the following:

i) $X \subset \mathbb{R}^D$, $\mu = m^{(D)}$, d usual Euclidean metric;

ii) $w_k \in Sim_1(X, d)$, with $X = \cup_{k=1}^N X_k$, where $X_k = w_k(X)$ (*covering condition*);

iii) $\mu(X_i \cap X_j) = 0$ when $i \neq j$ (μ-*non-overlapping condition*); and

iv) $\phi_k : \mathbb{R} \to \mathbb{R}$ are affine with $\phi_k(t) = \alpha_k t + \beta_k$, $t \in \mathbb{R}$.

Then, given v as above,

$$\begin{aligned}\Delta_k^2 &\equiv \int_{X_k} [\alpha_k v(w_k^{-1}(x)) + \beta_k - v(x)]^2 d\mu \\ &= c_k^D \int_X [\alpha_k v(x) + \beta_k - v(w_k(x))]^2 d\mu.\end{aligned}$$

As before, with the formal solution of the Inverse Problem in mind, assume that the w_k are fixed, and hence for each k, Δ_k can be viewed as a quadratic form in the parameters α_k and β_k:

$$\begin{aligned}c_k^{-D}\Delta_k^2 = ||v||_2^2\alpha_k^2 + 2\alpha_k\beta_k||v||_1 + \beta_k^2 - 2 < v, v \circ w_k > \alpha_k \\ - 2||v \circ w_k||_1\beta_k + ||v \circ w_k||_2^2.\end{aligned}$$

The problem can be viewed as a least squares minimization of Δ_k with respect to α_k and β_k. Set

$$\frac{\partial \Delta_k^2}{\partial \alpha_k} = \frac{\partial \Delta_k^2}{\partial \beta_k} = 0, \quad k = 1, 2, \ldots, N.$$

Then

$$||v||_2^2\alpha_k + ||v||_1\beta_k =< v \circ w_k, v >,$$
$$||v||_1\alpha_k + \beta_k = ||v \circ w_k||_1,$$

for $k = 1, 2, \ldots, N$.

Then, if $D_v \equiv ||v||_1^2 - ||v||_2^2 \neq 0$, the solutions are given by

$$\alpha_k = D_v^{-1}(< v \circ w_k, v > -||v \circ w_k||_1 \, ||v||_1)$$
$$\beta_k = D_v^{-1}(||v||_2^2 \, ||v \circ w_k||_1 - ||v||_1 < v \circ w_k, v >),$$

for $1 \leq k \leq N$.

When considering images, the condition that $\phi_k : \mathbb{R}^+ \to \mathbb{R}^+$ is needed. This forces $\alpha_k, \beta_k \geq 0$. It is not guaranteed that the α_k and β_k given by the above method will be nonnegative. However, if we consider an image to be a function defined on a compact subset A of $\mathbb{R}$, the condition on the α_k and β_k could be relaxed, with $\phi_k(v(x))$ still being nonnegative on A.

Hence, given $v \in L^2(X, \mu)$, fix N_J range blocks J_k, $1 \leq k \leq N_J$, and N_I domain blocks I_j, $1 \leq j \leq N_I$. For each range block J_k, minimize the distance $\Delta_{j,k}^2$, for each domain block I_j, $1 \leq j \leq N_I$. Then, let $I_{j(k)}$ be the domain block for which $\Delta_{j(k),k}$ is minimized over the domains. The values of $I_{j(k)}$, and the associated parameters α_k and β_k, are then stored, for $1 \leq k \leq N_J$. These values are called an LIFSM approximation of v.

Example 1.2.39 Consider the function $u(x) = \sin(\pi x)$ for $x \in X = [0, 1]$. Some approximations to u using the LIFSM method are shown in Figure 1.6 on page 34. Table 1.3 on the next page gives the L^2 distance between u and the approximations. Comparing these re-

sults with the IFSM case shown in Example 1.2.36 on page 26, one can see the enhancement brought by the LIFSM method.

Domains	Ranges	Distance	File size (bytes)	Computation time (sec)
u	n/a	0.0	30768	n/a
2	4	0.0266135	82	0.160
2	8	0.0144324	162	0.160
2	16	0.00762873	322	0.170
4	16	0.00131272	324	0.170

Table 1.3: LIFSM compression of $\sin(\pi x)$.

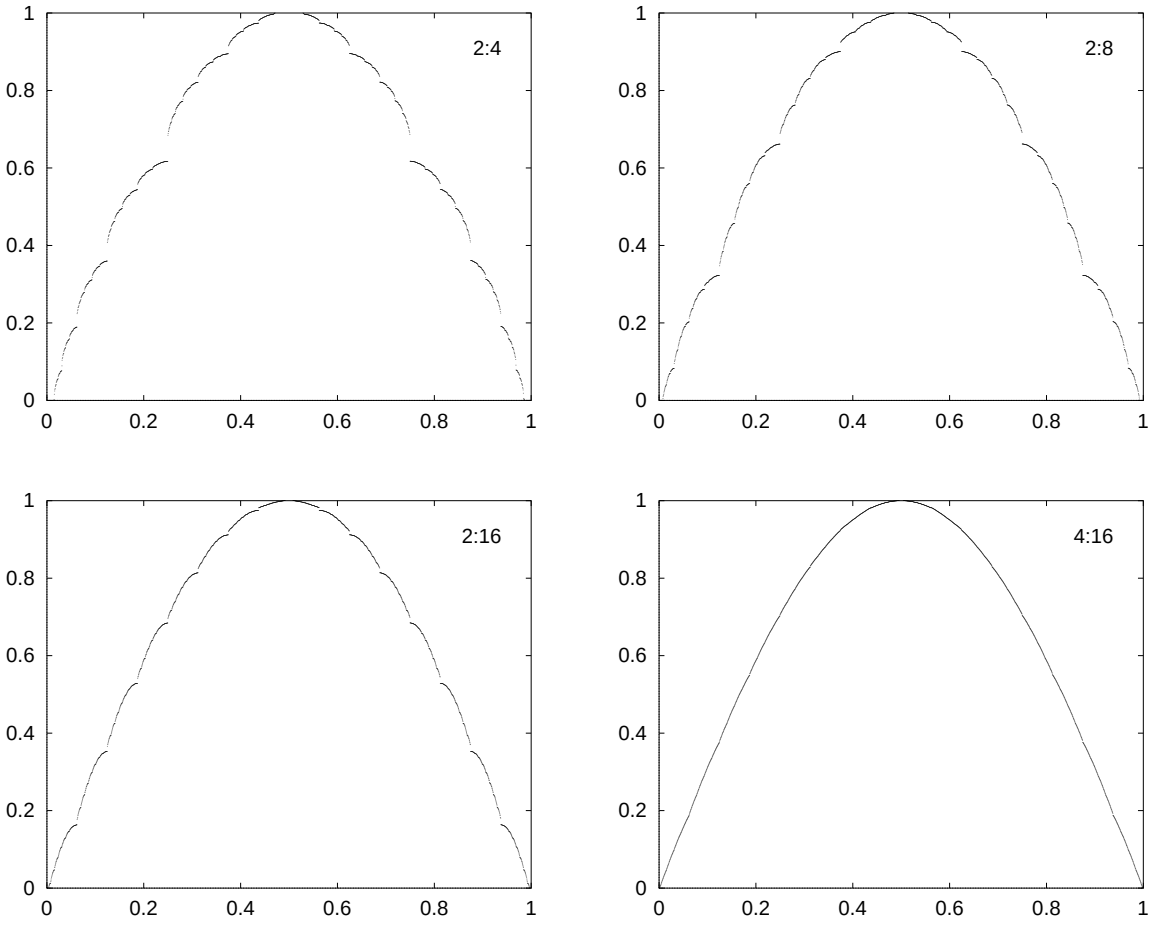


Figure 1.6: LIFSM approximation of $\sin(\pi x)$ with block ratio $(D : R)$ from left to right, top to bottom, 2:4, 2:8, 2:16 and 4:16.

1.3 Complex Bases

In this section, we discuss another kind of base. Here, we use a base and digits from an integer set to represent the numbers in a given system. For more on the theory of complex bases, we refer the reader to some papers by Gilbert [31, 32, 33]. Most of the details for the results that follow can be found there.

1.3.1 Fundamentals of Complex Bases

The concept of representing numbers in a base is a very simple and familiar one. In real numbers, the most common examples are the decimal and binary systems.

Example 1.3.1 Representing positive real numbers with $\mathbb{N}$ as the integers.

Base 10, digits $\{0, 1, 2, \ldots, 9\}$

$$193_{10} = 1 \cdot 10^2 + 9 \cdot 10^1 + 3 \cdot 10^0.$$

Base 2, digits $\{0, 1\}$

$$11010_2 = 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = 26_{10}.$$

The same idea carries over to complex numbers, as was demonstrated by Penny in [74]. In this case, the integers are the Gaussian integers:

Definition 1.3.2 The set of *Gaussian integers*, denoted by $\mathbb{Z}[i]$, is the set of complex numbers of the form $a + bi$, where $a, b \in \mathbb{Z}$. The *norm* of a Gaussian integer $z = a + bi$ is defined as

$$\text{norm}(z) \equiv a^2 + b^2 = |z|^2.$$

Definition 1.3.3 A *valid base* is a pair (b, D) where b is a Gaussian integer and $D \subset \mathbb{Z}[i]$, such that $0 \in D$ and every Gaussian integer z can be represented uniquely as a sum of

powers of b, with coefficients in D. More precisely, each $z \in \mathbb{Z}[i]$ can be written uniquely as

$$z = \sum_{j=0}^{t_z-1} a_j b^j, \text{ where } a_j \in D \text{ and } t_z \in \mathbb{N}^+.$$

If z has this form, write $z = (a_{t_z-1} \cdots a_1 a_0)_b$. This is called *base b positional notation*. The integer b is often referred to as the *base*, and the set D is called the *digit set*. When the base is understood, the subscript b in the positional notation is often omitted.

Example 1.3.4 It can be shown that $(-1+i, \{0,1\})$ is a valid base. We can expand the number 3 in $(-1+i, \{0,1\})$:

$$3 = (-1+i)^3 + (-1+i)^2 + (-1+i)^0 = (1101)_{-1+i}.$$

Definition 1.3.5 Let (b, D) be a valid base. If $z \in \mathbb{C}$ has the form

$$z = \sum_{j=-\infty}^{t_z-1} a_j b^j, \text{ where } a_j \in D \text{ and } t_z \in \mathbb{N}^+,$$

then the *radix expansion of z in base (b, D)* is defined as

$$z = (a_{t_z-1} \cdots a_1 a_0 \,.\, a_{-1} a_{-2} \cdots)_b.$$

The point between the digits a_0 and a_{-1} is called the *radix point*. The string of digits to the left of the radix point is called the *integer part* of z, while the string to the right is called the *radix part*. Another name for the radix expansion of a complex number is the *address* of the number.

Example 1.3.6 The complex number $(-1-8i)/5$ can be expressed in base $(-1+i, \{0,1\})$ as

$$(-1-8i)/5 = (111.\overline{10})_{-1+i},$$

where $\overline{10}$ indicates that the string 10 is repeated indefinitely. This can be seen by considering a series in $(-1+i)^{-1}$:

$$\begin{aligned}\sum_{k=0}^{\infty}(-1+i)^{-2k-1} &= \frac{1}{-1+i}\sum_{k=0}^{\infty}\left(\frac{1}{-2i}\right)^k \\ &= \frac{1}{(-1+i)}\frac{1}{1-\frac{1}{-2i}} \\ &= \frac{1}{(-1+i)}\frac{-2i}{-2i-1} \\ &= \frac{-2i}{3+i} \\ &= \frac{-1-3i}{5}.\end{aligned}$$

Hence

$$\begin{aligned}(111.\overline{10})_{-1+i} &= (-1+i)^2+(-1+i)+1+\frac{-1-3i}{5} \\ &= -2i-1+i+1+\frac{-1-3i}{5} \\ &= \frac{-5i}{5}+\frac{-1-3i}{5} \\ &= \frac{-1-8i}{5}.\end{aligned}$$

An important concept in the theory of complex numbers is that of a complete residue system:

Definition 1.3.7 A subset of $\mathbb{Z}[i]$ consisting of one element from each coset of the quotient ring $\mathbb{Z}[i]/b\mathbb{Z}[i]$ is called a *complete residue system* for $\mathbb{Z}[i]$ modulo b.

The cosets of $\mathbb{Z}[i]/b\mathbb{Z}[i]$ are precisely the sets S_k, $k \in \mathbb{Z}[i]$, where $z \in S_k$ if and only if there is some $m \in \mathbb{Z}[i]$ such that $z = bm + k$. One can show that a complete residue system for $\mathbb{Z}[i]/b\mathbb{Z}[i]$ contains $\text{norm}(b)$ elements.

Example 1.3.8 Letting $b = -1 + i$, then one complete residue system for $\mathbb{Z}[i]/b\mathbb{Z}[i]$ is $\{0, 1\}$. This can be seen easily as follows: Let $z = x + iy \in \mathbb{Z}[i]$. If $x + y$ is even, then

$z = bm$ with

$$m = \frac{-x+y}{2} + i\frac{-x-y}{2} \in \mathbb{Z}[i].$$

If $x + y$ is odd, then $z = bm + 1$ with

$$m = \frac{-x+y+1}{2} + i\frac{-x-y+1}{2} \in \mathbb{Z}[i].$$

Proposition 1.3.9 (Gilbert [32]) If (b, D) is a valid base, then D is a *complete residue system* for $\mathbb{Z}[i]$ modulo b and hence contains norm(b) elements.

Proof Suppose $z = \sum_{j=0}^{t} a_j b^j$, $a_j \in D$. Then $z \equiv a_0 \mod b$. Hence D contains a complete residue system for $\mathbb{Z}[i]$ modulo b. Now, suppose $c, d \in D$ are distinct and $c \equiv d \mod b$. Then let $e = \sum_{j=0}^{t} a_j b^j$, $a_j \in D$ such that $c - d = eb$. Hence, $(c)_b$ and $(a_t \ldots a_0 d)_b$ are two distinct addresses of c, which implies that (b, D) is not a valid base. ■

Theorem 1.3.10 (Gilbert [34]) Each $z \in \mathbb{C}$ has an infinite radix expansion in a valid base. However, this expansion may not necessarily be unique.

Proof The proof involves showing that if (b, D) is a valid base for $\mathbb{Z}[b]$, then every element of $\mathbb{Q}(b) \otimes \mathbb{R}$ has an infinite radix expansion in (b, D). This uses a special norm on $\mathbb{Q}(b) \otimes \mathbb{R}$. The fact that $\mathbb{Q}(b) \otimes \mathbb{R}$ is isomorphic to $\mathbb{C}$ yields the result. ■

It therefore makes sense to define the *fundamental* or *principal tile* of a valid base.

Definition 1.3.11 Given a valid base (b, D), define the *fundamental tile* $T(b, D)$ as the set of complex numbers with zero integer part in the base.

By Theorem 1.3.10, $\mathbb{C} = \cup_{z \in \mathbb{Z}[i]}(T(b, D) + z)$. Gilbert realized that the fundamental tiles of complex bases can be generated using IFS. Indeed, the fundamental tile of the complex base (b, D), is the attractor of the IFS $w = \{w_k\}$ given by

$$w_k(z) = \frac{z + d_k}{b}, \quad d_k \in D.$$

Figure 1.7 on the next page presents the fundamental tiles of a number of different bases, though the fact that they are valid bases is not proved (see Theorem 1.3.13 below). Although the scales are different, each tile has two-dimensional Lebesgue measure 1. Notice also that the origin is contained within the interior of these tiles. This is a key property of complex bases.

The following result shows that there are many valid bases:

Theorem 1.3.12 (Davio, Deschamps and Gossart [21]) Given any $b \in \mathbb{Z}[i]$ with modulus larger than one, except 2 and $1 \pm i$, there exists a complete residue system D such that (b, D) is a valid base for $\mathbb{C}$.

The following result gives an entire class of complex bases:

Theorem 1.3.13 (Kátai and Szabó [51]) If $b \in \mathbb{Z}[i]$, with norm N and $D = \{0, 1, 2, \ldots, N-1\}$, then (b, D) is a valid base for $\mathbb{C}$ iff $b = -n \pm i$, for some positive integer n.

Further generalizations can be found in [30, 50].

Corollary 1.3.14 The pair $(-1+i, \{0, 1\})$ is a valid base.

It is possible to generalize the notion of a valid base to any number system. Thus, an integer β will be called a *valid base* for a number system, using the set of integer digits D, if $0 \in D$ and if every integer z in the number system can be represented uniquely in the form

$$z = \sum_{j=0}^{t_z - 1} d_j \beta^j, \text{ where } d_j \in D \text{ and } t_z \in \mathbb{N}^+.$$

The notation for the expansion is the same as that given in Definition 1.3.5. The following unpublished result of Gilbert characterizes valid bases where β is an algebraic integer. In Section 2.3, we will demonstrate the validity of this result for number systems in $\mathbb{Z}^n$.

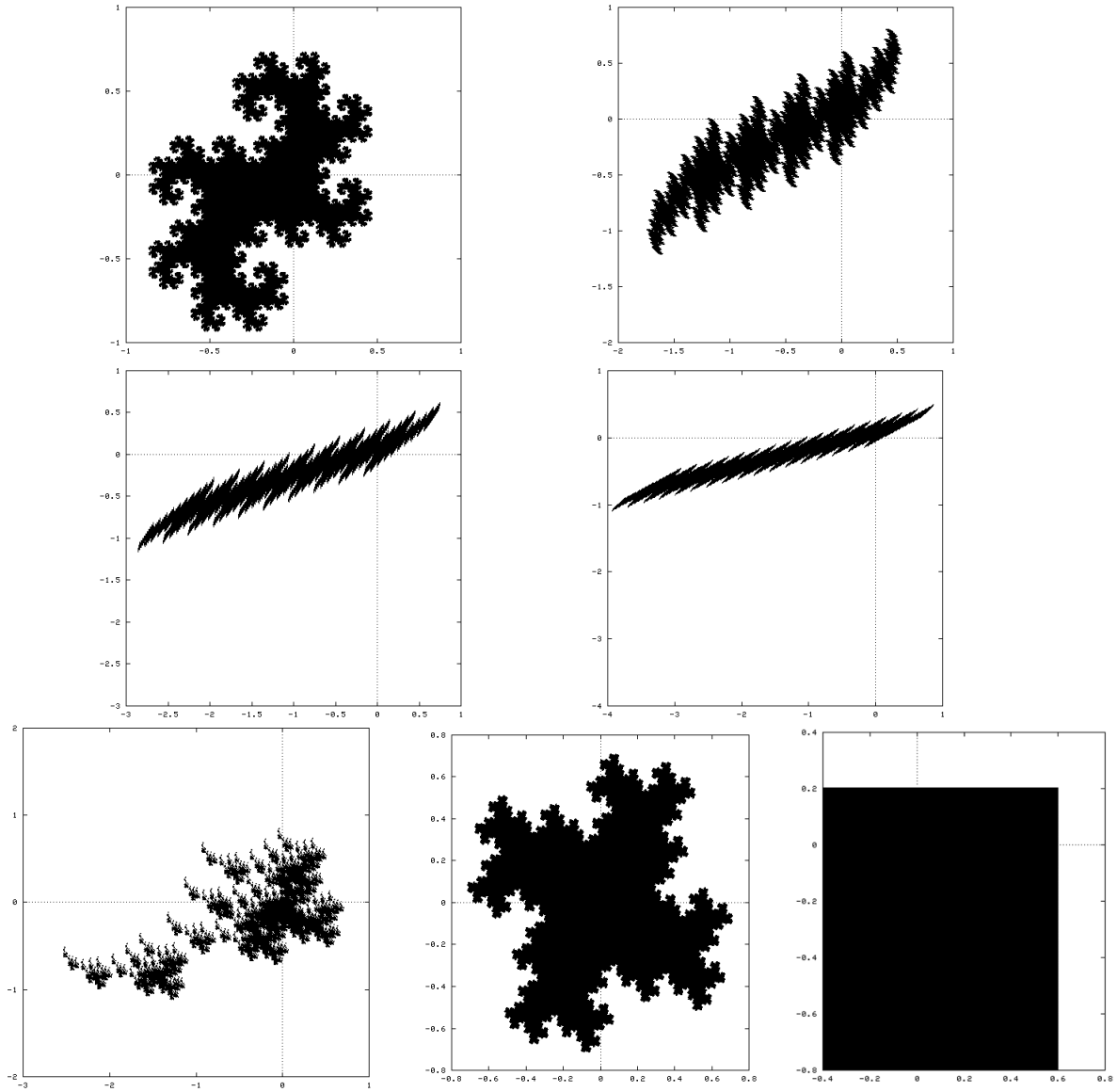


Figure 1.7: Fundamental tiles of the complex bases (from left to right, top to bottom) $(-1+i, \{0,1\})$, $(-2+i, \{0,\ldots,4\})$, $(-3+i, \{0,\ldots,9\})$, $(-4+i, \{0,\ldots,16\})$, $(2+i, \{0,1,i,-i,-2-3i\})$, $(2+i, \{0,1,-1,i,-i\})$ and $(2i, \{0,1,i,1+i\})$. Each of these tiles has Lebesgue measure 1.

Theorem 1.3.15 (Gilbert [34]) If β is an algebraic integer and D is a complete residue system for $\mathbb{Z}[\beta]$ modulo β, that contains 0, then the following are equivalent:

i) The integer β is a valid base using the digit set D;

ii) For every $z \in \mathbb{Z}[\beta]$ there exists a positive integer r such that $\Phi^r(z) = 0$, where the function $\Phi : \mathbb{Z}[\beta] \to \mathbb{Z}[\beta]$ is defined by $\Phi(z) = (z - d)/\beta$, $d \in D$ and $d \equiv z \mod \beta$;[5]

iii) The integer β and all its conjugates have moduli greater than one and there is no positive integer t for which

$$d_{t-1}\beta^{t-1} + \ldots + d_1\beta + d_0 \equiv 0 \mod (\beta^t - 1) \text{ with } d_{t-1}, \ldots, d_0 \in D$$

and not all d_i are equal to zero.

1.3.2 Representation in a Complex Base

There are various algorithms for determining the representation of Gaussian integers in a valid base. These are due to Gilbert [30, 33, 34, 35].

Algorithm 1.3.16 (Base Conversion Algorithm) (Gilbert [34]) Let (b, D) be a valid base. Since D is a complete residue system for $\mathbb{Z}[i]$ modulo b then, given $z \in \mathbb{Z}[i]$, there exist unique integers $q_j \in \mathbb{Z}[i]$ and $a_j \in D$, $j = 1 \ldots t$, $t \in \mathbb{N}^+$ such that

$$\begin{aligned} z &= q_1 b + a_0 \\ q_1 &= q_2 b + a_1 \\ &\vdots \\ q_t &= 0b + a_t. \end{aligned}$$

Hence, $z = (a_t \ldots a_1 a_0)_b$.

[5] Here $\Phi^r(z)$ denotes the r-fold composition of Φ at z.

Example 1.3.17 Let $z = 5 + 12i$. Find the address of z in $(-2 + i, \{0, 1, 2, 3, 4\})$. Using the Base Conversion Algorithm

$$\begin{aligned} 5 + 12i = &\quad (2 - 5i)(-2 + i) + 4 \\ 2 - 5i = &(-1 + 2i)(-2 + i) + 2 \\ -1 + 2i = &\quad (2)(-2 + i) + 3 \\ 2 = &\quad (0)(-2 + i) + 2 \end{aligned}$$

Hence the address of $5 + 12i$ is (2324).

Given a rational number, there is the Long Division Algorithm for finding the representation [35].

Algorithm 1.3.18 (Long Division Algorithm) Let (b, D) be a valid base for $\mathbb{C}$. Given Gaussian integers v and $w \neq 0$, there exists $A \in \mathbb{Z}[i]$, and digits $a_j \in D$, such that

$$\begin{aligned} v &= Aw + r_0 \\ br_0 &= a_1 w + r_1 \\ br_1 &= a_2 w + r_2 \\ &\vdots \end{aligned}$$

where the *remainders* $r_j \in \text{RemSet}(b, D, w) = wT(b, D) \cap \mathbb{Z}[i]$.

There may be choices in the algorithm, but for each choice,

$$\frac{v}{w} = A + (0.a_1 a_2 \ldots).$$

Since the integer part A can be represented in base (b, D), we obtain a representation of v/w in this base. In addition, any expansion can be obtained in this way.

Proof The proof uses IFS. It also uses the Escape Time Algorithm to show that, with the given remainder set, the sequence of remainders stays bounded. ■

There is also an algorithm for finding the remainder set [35]. This set can be found by considering a directed graph over $\mathbb{Z}[i]$. However, this involves determining which integers inside the ball of radius $M|w|/(|b|-1)$, where M is the largest modulus of the elements of D, are remainders. Since b is fixed, this ball grows rapidly with $|w|$. For example, given the base $(-1+i, \{0,1\})$ and $w = 1000$, the ball contains approximately 2×10^6 integers.

An enormous amount of calculation is required to build this graph. Indeed, our first attempt to link wavelets and complex bases utilized this approach. For 512×512 images, $w = 512$. Using a DEC Alpha, estimates indicated unmanageable calculation times would be required to perform the remainder set algorithm, let alone to perform long division.

A fast algorithm, called the Clearing Algorithm, also exists for finding the expansion of integers in bases of the form (b, D) where $D \subset \mathbb{Z}$. For simplicity, we consider only bases $(-n+i, \{0, 1, \ldots, n^2\})$, $n \in \mathbb{N}^+$. The reader is referred to [30] for the general result.

Algorithm 1.3.19 (Clearing Algorithm) (Gilbert [30]) Consider the valid base $(-n+i, \{0, 1, \ldots, n^2\})$ and let $p(x)$ be the minimal polynomial of $b = -n+i$. Thus

$$p(x) = x^2 + 2nx + n^2 + 1.$$

Then the representation of any integer $z \in \mathbb{Z}[i]$ in the base can be obtained as follows:

Begin with $z = m(b) = a_k b^k + \cdots + a_1 b + a_0$, an expansion of z in powers of b with integer coefficients. For instance, any Gaussian integer $c + id$ can be expanded in powers of $b = -n + i$ as $c + id = db + (c + nd)$. Consider this expansion as an element $m(b)$ of the polynomial ring $\mathbb{Z}[b]$. Let r be the least integer such that $a_r \notin D$. If no such r exists, then $m(b)$ is the unique expansion of z in the base. We call such a polynomial *clear*.

If r exists, then let s be the integer such that $0 \leq a_r + s(n^2+1) \leq n^2$. Add sb^r times $p(b)$ to $m(b)$. Remember that we perform this operation in the polynomial ring $\mathbb{Z}[b]$. Call this new polynomial $m_1(b)$.

However, $p(-n+i) = 0$, thus $m(b)$ and $m_1(b)$ are equal in $\mathbb{C}$. Hence, $m_1(b)$ is an

expansion of z in $\mathbb{Z}[b]$. In addition, the coefficient of the r-th power of b in $m_1(b)$ is a digit. We say that a_r has been *cleared*.

Repeat the process of clearing coefficients, by induction on r, until a clear polynomial is obtained. The resulting polynomial provides the expansion of z in (b, D). This process must terminate after a finite number of steps.

Example 1.3.20 Determine the expansion of $5 + 12i$ in base $(-2 + i, \{0, 1, 2, 3, 4\})$.

The minimal polynomial of $b = -2 + i$ is $x^2 + 4x + 5$. Hence, $b^2 + 4b + 5 = 0$ and, by abuse of notation, we can write this as $(1\ 4\ 5)_b = 0$. (Note that 5 is not in the digit set.) Begin with the expansion $5 + 12i = 12b + 29 = (12\ 29)_b$. Then, we clear the polynomial in $\mathbb{Z}[b]$ as follows:

		12	29
	-5	-20	-25
	-5	-8	4
2	8	10	
2	3	2	4

Hence, the expansion of $5+12i$ in base $(-2+i, \{0, 1, 2, 3, 4\})$ is (2324). A quick calculation verifies that this is indeed correct.

Given its generality, we have used the Base Conversion Algorithm to calculate the addresses in our applications, which will be discussed in Chapter 3.

1.4 Wavelet Analysis

In this section, we summarize certain fundamental results of wavelet analysis with Haar wavelets. The reader is referred to the paper of Gröchenig and Madych [37] for a more in-depth discussion. At the end of the section, we present the results of IFS on wavelet coefficient trees.

1.4.1 Fundamentals of Wavelet Analysis

For simplicity, the discussion is restricted to $\mathbb{Z}^n$, although general lattices can be considered.

Definition 1.4.1 (Acceptable Dilation) A matrix A on $\mathbb{R}^n$ is an *acceptable dilation for* $\mathbb{Z}^n$ if $A\mathbb{Z}^n \subset \mathbb{Z}^n$ and if $|\lambda| > 1$ for each eigenvalue λ of A.

Example 1.4.2 The matrix

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

is an acceptable dilation since A has integer entries, and $\lambda = 1 \pm i$, thus $|\lambda| = \sqrt{2} > 1$. In this case, $|\det A| = 2$.

Throughout this section we will let $q = |\det A|$.

Proposition 1.4.3 The properties of an acceptable dilation imply that q is an integer greater than or equal to two.

Proof For $j \in \{1, 2, \ldots, n\}$, let $e_j \in \mathbb{Z}^n$ be the vector with the entry 1 in the j-th component and 0 elsewhere. By definition of A, $Ae_j \in \mathbb{Z}^n$. For $i \in \{1, 2, \ldots, n\}$, the i-th component of Ae_j is the (i, j)-th component of A, which must therefore be an integer. Hence, A is an integer matrix. Thus, the determinant of A is an integer. By a result of linear algebra, $\det A$ is the product of the eigenvalues of A. However, for each eigenvalue λ, $|\lambda| > 1$. Hence, $q = |\det A|$ is an integer greater than 1 and therefore $q \geq 2$. ■

Let A be an acceptable dilation on $L^2(\mathbb{R}^n)$, $f \in L^2(\mathbb{R}^n)$ and $x, y \in \mathbb{R}^n$. Define the unitary dilation operator U_A by

$$U_A f(x) = |\det A|^{-1/2} f(A^{-1}x)$$

and the translation operator τ_y by

$$\tau_y f(x) = f(x - y).$$

Then for each $i \in \mathbb{Z}$ and $j \in \mathbb{Z}^n$ let $f_{i,j} \equiv U_A^{-i}\tau_j f$. Hence,

$$f_{i,j}(x) = q^{i/2} f(A^i x - j).$$

Using this notation, $f_{0,0} = f$. For simplicity also, we will write f_i to mean $f_{0,i}$.

We are interested in wavelet bases given by translation by integers.

Definition 1.4.4 A *wavelet basis* B, associated with an acceptable dilation A, is a basis of $L^2(\mathbb{R}^n)$ whose members are A dilates and $\mathbb{Z}^n$ translates of a finite orthonormal set $S = \{\psi^1, \ldots, \psi^m\} \subset L^2(\mathbb{R}^n)$, where $m \in \mathbb{N}^+$. More precisely,

$$B = \{\psi^l_{i,j} \ : \ l = 1, \ldots, m; \ i \in \mathbb{Z}; \ j \in \mathbb{Z}^n\},$$

where $\psi^l_{i,j} = q^{i/2}\psi^l(A^i x - j)$. The elements of S are called the *basic (mother) wavelets*.

Example 1.4.5 (Haar Wavelets) Let $n = 1$ and $Ax = 2x$, $x \in \mathbb{R}$. Then, the *Haar basis* satisfies Definition 1.4.4, with $m = 1$ and

$$\psi(x) = \begin{cases} 1 & 0 \le x < 1/2, \\ -1 & 1/2 \le x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

The Haar mother wavelet is sketched in Figure 1.8 on page 49.

A direct proof of this fact can be found in [87, p. 9]. However, a general proof that this function generates a wavelet basis follows from the theory of *multiresolution analysis*. This is one of the fundamental concepts of wavelet theory. For more details on MRA, see [20, 61, 69, 70].

Consider the definition of a multiresolution analysis as given in [37] where the lattice $\Gamma = \mathbb{Z}^n$.[6]

Definition 1.4.6 Let A be an acceptable dilation for $\mathbb{Z}^n$. A *multiresolution analysis (MRA)* associated with A is a sequence of closed subspaces $(V_i)_{i\in\mathbb{Z}}$ of $L^2(\mathbb{R}^n)$, satisfying:

i) $V_i \subset V_{i+1}$, $\forall i \in \mathbb{Z}$;

ii) $\overline{\cup_{i\in\mathbb{Z}} V_i} = L^2(\mathbb{R}^n)$;

iii) $V_i = U_A^{-i} V_0$, $\forall i \in \mathbb{Z}$;

iv) $\tau_j V_0 = V_0$, $\forall j \in \mathbb{Z}^n$; and

v) there is a function $\phi \in V_0$, called the *scaling function*, such that $\{\tau_j \phi \ : \ j \in \mathbb{Z}^n\}$ is a basis for V_0.

Properties iii) and iv) imply that $\{\phi_{i,j} : j \in \mathbb{Z}^n\}$ is a basis for V_i for each $i \in \mathbb{Z}$. Since $\phi \in V_0 \subset V_1$, we obtain the *dilation equation*:

$$\begin{aligned}\phi(x) &= \sum_{j\in\mathbb{Z}^n} h_j \phi_{1,j}(x) \\ &= \sum_{j\in\mathbb{Z}^n} h_j U_A^{-1} \tau_j \phi(x) \\ &= \sum_{j\in\mathbb{Z}^n} h_j |\det A|^{1/2} \phi(Ax - j), \ \forall x \in \mathbb{R}^n,\end{aligned}$$

where $h_j = \langle \phi, \phi_{1,j} \rangle, \forall j \in \mathbb{Z}^n$.

Given a multiresolution analysis we define, for each $i \in \mathbb{Z}$, the space W_i as the orthogonal complement of V_i in V_{i+1}: $W_i = V_{i+1} \ominus V_i$. Thus, it follows that $W_i = U_A^{-i} W_0$ and that

[6] In [37], a lattice means an image of $\mathbb{Z}^n$ under some nonsingular linear transformation.

$L^2(\mathbb{R}^n) = \bigoplus_{i\in\mathbb{Z}} W_i$.[7]

Recalling the result of Meyer [68] that there exist $q-1$ functions $\psi^1, \ldots, \psi^{q-1}$ such that $\{\tau_j\psi^l \ : \ j \in \mathbb{Z}^n;\ l = 1, \ldots, q-1\}$ is a basis for W_0, then the set

$$\{\psi^l_{i,j} \ : \ l = 1, \ldots, q-1 \ ; i \in \mathbb{Z} \ ; j \in \mathbb{Z}^n\}$$

is a wavelet basis for $L^2(\mathbb{R}^n)$. Since $W_0 \subset V_1$, we get a dilation equation for each of the ψ^l, $l = 1, \ldots, q-1$:

$$\psi^l = \sum_{j\in\mathbb{Z}^n} g^l_j \phi_{1,j},$$

where $g^l_j = \langle \psi^l, \phi_{1,j}\rangle, \forall j \in \mathbb{Z}^n$. The coefficients h_j and g^l_j are called the *filter coefficients* of the scaling function and wavelet functions respectively.

Example 1.4.7 With the theory of MRA, it follows that the function ψ in Example 1.4.5 generates a wavelet basis, by considering the function $\phi = \chi_{[0,1]}$ in $L^2(\mathbb{R})$, with $A = 2$. This is the most basic scaling function, and is called the *Haar scaling function.* Its dilation equation is

$$\phi(t) = \phi(2t) + \phi(2t-1).$$

The mother wavelet already presented is given by

$$\psi(t) = \phi(2t) - \phi(2t-1).$$

Therefore, the filter coefficients are given by

$$h_0 = g_0 = \frac{1}{\sqrt{2}} \tag{1.5}$$

$$h_1 = -g_1 = \frac{1}{\sqrt{2}}. \tag{1.6}$$

The two functions are sketched in Figure 1.8.

[7]The symbol $\bigoplus$ denotes the orthogonal direct sum.

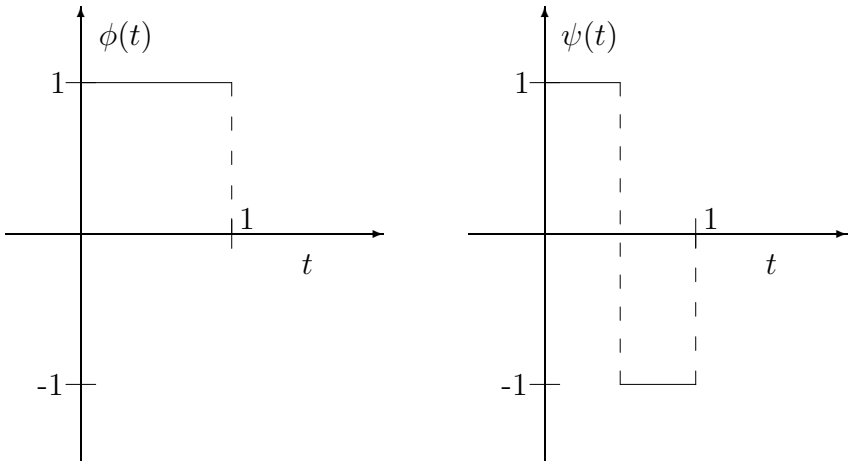


Figure 1.8: The Haar scaling function ϕ (left) and a mother wavelet ψ (right).

Example 1.4.8 The Shannon scaling function is shown in Figure 1.9 on page 51. It is given by the equation

$$\phi(t) = \frac{\sin(\pi t)}{\pi t}.$$

Thus, by the Shannon sampling theorem,[8] it satisfies the scaling relation

$$\phi(t) = \sum_{n\in\mathbb{Z}} \frac{\sin(\pi n/2)}{\pi n/2}\phi(2t-n).$$

In this example, there are infinitely many filter coefficients. In general, there are many choices of the mother wavelet function ψ. For the Shannon system, two such functions are

$$\begin{aligned}\psi(t) &= 2\phi(2t) - \phi(t)\\ &= \frac{\sin(2\pi t) - \sin(\pi t)}{\pi t}\end{aligned}$$

[8]This theorem is fundamental in engineering. It is also referred to as the Nyquist theorem, but is originally a result of Whittaker. For further details, see [15, p. 44].

and

$$\psi(t) = \frac{\sin \pi(t - \frac{1}{2}) - \sin 2\pi(t - \frac{1}{2})}{\pi(t - \frac{1}{2})}.$$

For more details, see [15].

Finding the scaling function ϕ and the functions ψ^l may be extremely difficult in general. However, in the case where the scaling functions are *characteristic functions on self-similar lattice tilings*, such basic wavelets can always be found.

1.4.2 Self-Similar Lattice Tilings

We assume that the reader is familiar with the notions of measurability. A good reference is [77]. Let $Q, R \subset \mathbb{R}^n$ be Lebesgue measurable.

Denote by χ_Q the *characteristic function* of Q. That is

$$\chi_Q(x) = \begin{cases} 1 & x \in Q, \\ 0 & x \notin Q. \end{cases}$$

Example 1.4.9 The Haar function $\chi_{[0,1]}$ is a characteristic function on $\mathbb{R}$.

Let $|Q|$ denote the Lebesgue measure of Q. Write $Q \simeq R$ if $|Q \backslash R| = |R \backslash Q| = 0$.

Example 1.4.10 We have that $[0, 1] \simeq (0, 1)$ since points have Lebesgue measure zero.

We also state the following definitions.

Definition 1.4.11 A set Q is said to *tile* $\mathbb{R}^n$ *by integer translates* if:

i) $\bigcup_{k \in \mathbb{Z}^n} (Q + k) \simeq \mathbb{R}^n$; and

ii) $Q \cap (Q + k) \simeq \emptyset$, $\forall k \in \mathbb{Z}^n \backslash \{0\}$.

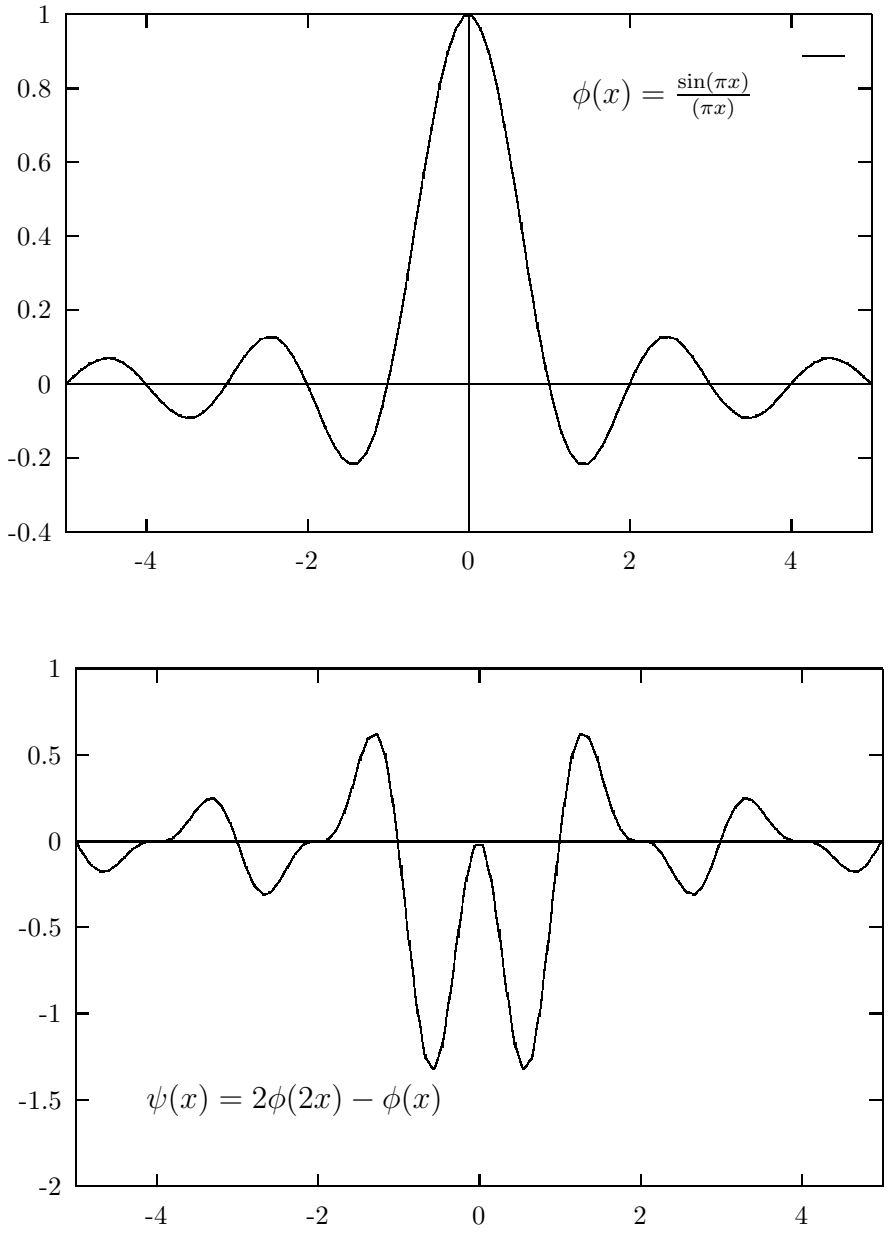


Figure 1.9: The Shannon scaling function ϕ and a mother wavelet ψ.

Example 1.4.12 The set $Q = [0, 1]$ tiles $\mathbb{R}$ by integer translates since $\mathbb{R} = \cup_{n\in\mathbb{Z}}[n, n+1]$.

Proposition 1.4.13 Let Q be a measurable subset of $\mathbb{R}^n$ satisfying $\bigcup_{k\in\mathbb{Z}^n}(Q+k) \simeq \mathbb{R}^n$. Then the following are equivalent:

i) $Q \cap (Q+k) \simeq \emptyset$, $\forall k \in \mathbb{Z}^n \backslash \{0\}$;

ii) $|Q| = 1$.

Proof Let $f(x) = \sum_{k\in\mathbb{Z}^n} \chi_Q(x-k)$.

i) $\Longrightarrow$ ii) Since Q tiles $\mathbb{R}^n$ by integer translates, then $f \geq 1$ almost everywhere (a.e.). However, i) implies that $f \leq 1$ a.e. Thus $f \equiv 1$ a.e., and hence

$$|Q| = \int \chi_Q = \int_{I^n} f = |I^n| = 1,$$

where $I = [0, 1]$.

ii) $\Longrightarrow$ i) Again, $f \geq 1$ a.e. However, by ii), $\int_{I^n} f = |Q| = 1$. Therefore $f \equiv 1$ a.e., which implies i). ■

Definition 1.4.14 We say that a scaling function ϕ for an MRA of $L^2(\mathbb{R}^n)$ is a *Haar scaling function* if $\phi = \chi_Q$ for some measurable subset Q of $\mathbb{R}^n$.

In [37], Gröchenig and Madych completely characterized these scaling functions and their supports. Before presenting their results, we first give a short definition, which is analogous to Definition 1.3.7.

Definition 1.4.15 A subset of $\mathbb{Z}^n$ consisting of one element from each coset of the quotient ring $\mathbb{Z}^n/A\mathbb{Z}^n$ is called a *complete residue system* for $\mathbb{Z}^n$ modulo A.

The cosets of $\mathbb{Z}^n/A\mathbb{Z}^n$ are precisely the sets S_k, $k \in \mathbb{Z}^n$, where $z \in S_k$ if and only if there is some $m \in \mathbb{Z}^n$ such that $z = Am + k$. One can show that a complete residue system for $\mathbb{Z}^n/A\mathbb{Z}^n$ contains $q = |\det A|$ elements.

Example 1.4.16 Let A be the matrix

$$A = \begin{pmatrix} 0 & 2 \\ -1 & 0 \end{pmatrix}.$$

Then $\{(0,-1),(1,-1)\}$ is a complete residue system for $\mathbb{Z}^2/A\mathbb{Z}^2$. This follows since if $z=(x,y)\in\mathbb{Z}^2$ with x even, then $z = Am + k_1$, with $m = (-y-1, x/2)$ and $k_1 = (0,-1)$. If x is odd, then $z = Am + k_2$, with $m = (-y-1,(x-1)/2)$ and $k_2 = (1,-1)$. In both cases, $m \in \mathbb{Z}^2$.

Theorem 1.4.17 Suppose A is an acceptable dilation for $\mathbb{Z}^n$ and let Q be a measurable subset of $\mathbb{R}^n$. The function $\phi = |Q|^{-1/2}\chi_Q$ is the scaling function for a multiresolution analysis associated with A if and only if the following conditions are satisfied:

i) Q tiles $\mathbb{R}^n$ by integer translates;

ii) $AQ \simeq \bigcup_{k\in K}(Q+k)$ for some complete residue system K of $\mathbb{Z}^n/A\mathbb{Z}^n$; and

iii) $Q \simeq C$ for some compact subset C of $\mathbb{R}^n$.

A set satisfying Property ii) is called *self-similar in the affine sense.* By Proposition 1.4.13, we must have $|Q| = 1$. Hence,

$$\phi(x) = \sum_{k\in K} \phi(Ax-k),\ \forall x \in \mathbb{R}^n.$$

Therefore, the filter coefficients h_j of ϕ are identically $q^{-1/2}$ for $j \in K$ and zero otherwise.

Theorem 1.4.18 Let K be a complete residue system of $\mathbb{Z}^n/A\mathbb{Z}^n$. Then there exists a unique solution of $\phi = \sum_{k\in K} q^{-1/2}\phi_{1,k}$ in $L^1(\mathbb{R}^n)$, up to multiplication by a constant. Furthermore, this solution has support in the compact set

$$Q = \left\{ \sum_{i=1}^{\infty} A^{-i}k_i \ :\ k_i \in K \right\}.$$

Theorem 1.4.19 Let A be an acceptable dilation for $\mathbb{Z}^n$ and let $Q \subset \mathbb{R}^n$. Then the function $\phi = \chi_Q$ is the scaling function of an MRA associated with A if and only if $|Q| = 1$ and Q is of the form given in Theorem 1.4.18 for some complete residue system K of $\mathbb{Z}^n/A\mathbb{Z}^n$.

Now, let $(V_i)_{i\in\mathbb{Z}}$ be a multiresolution analysis associated with A, with $\phi = \chi_Q$. By Theorem 1.4.19, let K be the complete residue system generating Q. We write (A, K) to denote an MRA.

Lemma 1.4.20 The space W_0 consists of the functions $f \in L^2(\mathbb{R}^n)$ such that

$$f(x) = \sum_{k\in\mathbb{Z}^n} s_k q^{1/2} \chi_Q(Ax - k),$$

where $(s_k) \in \ell^2(\mathbb{Z}^n)$, satisfying $\sum_{k\in K} s_{k+Am} = 0, \forall m \in \mathbb{Z}^n$.[9]

Proof If $f \in W_0$, then $f = \sum_{k\in\mathbb{Z}^n} s_k \phi_{1,k}$, for some $(s_k) \in \ell^2(\mathbb{Z}^n)$. Given $m \in \mathbb{Z}^n$, $\tau_{-m} f \in W_0$ and

$$\begin{aligned} \tau_{-m} f &= \sum_{k\in\mathbb{Z}^n} s_k \phi_{1,k-Am} \\ &= \sum_{l\in\mathbb{Z}^n} s_{l+Am} \phi_{1,l}. \end{aligned}$$

Thus

$$\begin{aligned} 0 &= <\tau_{-m} f, \phi> \\ &= \sum_{l\in\mathbb{Z}^n} s_{l+Am} <\phi_{1,l}, \phi> \\ &= q^{-1/2} \sum_{l\in K} s_{l+Am}, \end{aligned}$$

by Theorem 1.4.17.

[9]The symbol $\ell^2(\mathbb{Z}^n)$ denotes the space of all square summable sequences over $\mathbb{Z}^n$.

Conversely, if f is of the prescribed form, then $f \in V_1$. Given $m \in \mathbb{Z}^n$,

$$\begin{aligned} < f, \phi_{0,m} > &= \sum_{k \in \mathbb{Z}^n} s_k < \phi_{1,k}, \phi_{0,m} > \\ &= \sum_{k \in \mathbb{Z}^n} s_k < \phi_{1,k-Am}, \phi > \\ &= \sum_{l \in \mathbb{Z}^n} s_{l+Am} < \phi_{1,l}, \phi > \\ &= q^{-1/2} \sum_{l \in K} s_{l+Am} \\ &= 0. \end{aligned}$$

Hence $f \in W_0$. ■

Using this result, Gröchenig and Madych characterized all the wavelets for such MRA.

Lemma 1.4.21 Let $\phi = \chi_Q$ be the scaling function for an MRA of $L^2(\mathbb{R}^n)$ associated with A and let $K = \{k_1, \ldots, k_q\}$ be the complete residue system generating Q. Let $U = (u_{ij})$ be a unitary $q \times q$ matrix, with $u_{1j} = q^{-1/2}$, $j = 1, \ldots, q$. For $i = 1, \ldots, q-1$, define

$$\psi^i = \sum_{j=1}^{q} u_{i+1j} \phi_{1,k_j}.$$

Then $\{\tau_j \psi^i \ : \ i = 1, \ldots, q-1 \ ; j \in \mathbb{Z}^n\}$ is a basis for W_0. Conversely, any set of basic wavelets for an MRA associated with $\phi = \chi_Q$ must arise in such a way.

The last fact is due to ϕ being a characteristic function, and since $\{\tau_k \phi \ : \ k \in \mathbb{Z}^n\}$ is an orthonormal set. We therefore obtain the following important result:

Theorem 1.4.22 Let $\{\psi^1, \ldots, \psi^{q-1}\}$ be as given in Equation (1.4.21). Then $\{\psi^l_{i,j} \ : \ l = 1, \ldots, q-1 \ ; \ i \in \mathbb{Z}; \ j \in \mathbb{Z}^n\}$ is a basis for $L^2(\mathbb{R}^n)$.

Example 1.4.23 For $q = 2$, there are only two possible matrices U:

$$U = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ \pm 1/\sqrt{2} & \mp 1/\sqrt{2} \end{pmatrix}.$$

Another example, slightly different than the one in [37] is the following.[10]

Example 1.4.24 For $q \geq 3$, define the unitary matrix $U = (u_{ij})$, by letting $u_{1j} = q^{-1/2}$ and

$$u_{ij} = \sqrt{\frac{2}{q}} \cos \frac{(i-1)(2j-1)\pi}{2q},$$

for $i = 2, \ldots, q$ and $j = 1, \ldots, q$.

We now show two examples of wavelets constructed using this method.

Example 1.4.25 Let

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

The matrix has determinant 4. A complete residue system is $K = \{(0,0), (1,0), (0,1), (1,1)\}$. Hence, its basic tile is simply I^2.

This is the standard matrix for the *separable wavelet transform*. By separable, we mean that the scaling function associated with this matrix is simply the tensor product of two copies of the one-dimensional Haar scaling function given in Example 1.4.7. However, other complete residue systems can be used and they generate fractal tilings [37].

Separable wavelets are well studied and are behind most of the wavelet compression applications used to date. A discussion of separable wavelets can be found in [20, p. 313]. Our interest resides mainly with non-separable wavelets.

Example 1.4.26 Let

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

[10] The example given in [37] is not a unitary matrix.

Here, $q = |\det A| = 2$. Therefore, any complete residue system K contains 2 elements. Since we can assume $k_1 = (0,0) \in K$, we must find some $k_2 = (k,l) \in \mathbb{Z}^2$, such that there does not exist an $m \in \mathbb{Z}^2$ satisfying

$$k_2 = k_1 + Am.$$

That is, there must not exist a solution $(m_1, m_2) \in \mathbb{Z}^2$ to the system

$$\begin{aligned} k &= m_1 - m_2 \\ l &= m_1 + m_2. \end{aligned}$$

We try $k_2 = (1,0)$, and attempt to solve

$$1 = m_1 - m_2 \tag{1.7}$$
$$0 = m_1 + m_2. \tag{1.8}$$

Given such a solution, Equation (1.8) implies that $m_1 = -m_2$. Thus, by Equation (1.7), $-2m_2 = 1$. Since no integer m_2 satisfies this condition, we see that $K = \{(0,0),(1,0)\}$ is a complete residue system for $\mathbb{Z}^2/A\mathbb{Z}^2$. Let Q be the tile generated by A and K. This tile, known as the *twin dragon*, is shown in Figure 1.10 on the following page.

Recalling Example 1.4.23, we now choose

$$\psi(x) = \chi_Q(Ax) - \chi_Q(Ax - (1,0)).$$

This function is identically 1 on $A^{-1}Q$, -1 on $A^{-1}(Q + (1,0))$, and 0 elsewhere. This basic wavelet is shown in Figure 1.10 on the next page. For clarity, the positive and negative parts are shown in black, and the zero values are not shown.

1.4.3 Mallat Algorithm

The strength of the multiresolution analysis method lies in the reconstruction and decomposition algorithms, discovered initially by Mallat [61]. These algorithms are a fundamen-

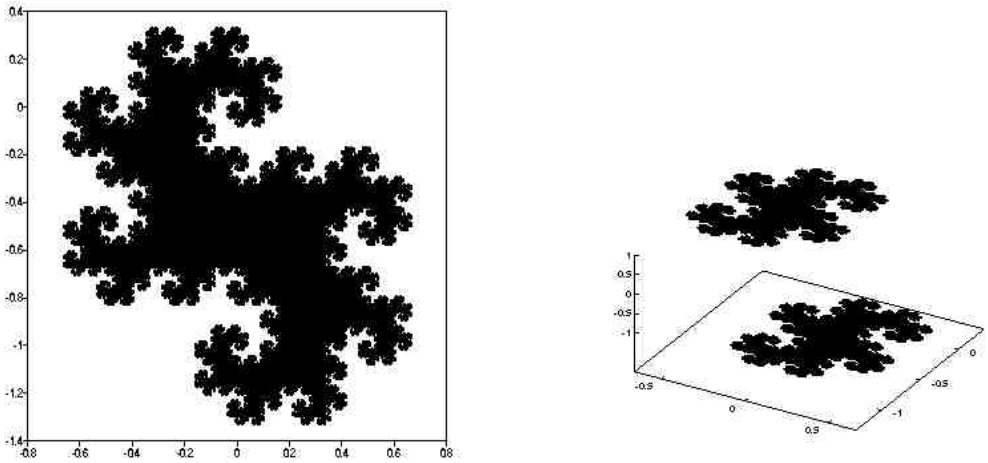


Figure 1.10: The support of the scaling function of Example 1.4.26 (left) and its basic wavelet (right).

tal component of wavelet analysis applied to signal and image processing. They will be reviewed here.

Suppose $(V_i)_{i\in\mathbb{Z}} \subset L^2(\mathbb{R}^n)$ is an MRA and that $f \in V_{i+1} = V_i \oplus W_i$. We have two bases: one for V_{i+1} and one for $V_i \oplus W_i$. Therefore,

$$\begin{aligned} f &= \sum_{z\in\mathbb{Z}^n} s_{i+1,z}\phi_{i+1,z} \\ &= \sum_{j\in\mathbb{Z}^n} s_{i,j}\phi_{i,j} + \sum_{l=1}^{q-1}\sum_{j\in\mathbb{Z}^n} w_{i,j}^l \psi_{i,j}^l, \end{aligned}$$

where the *scaling coefficients* $s_{m,n} = \langle f, \phi_{m,n}\rangle$ and the *wavelet coefficients* $w_{m,n}^l = \langle f, \psi_{m,n}^l\rangle$. By rewriting the dilation equations for ϕ and for the functions ψ^l, one can arrive at dilation equations for $\phi_{i,j}$ and the $\psi_{i,j}^l$. Since $\phi = \sum_{z\in\mathbb{Z}^n} h_z\phi_{1,z}$, then

$$\begin{aligned}\phi_{i,j} &= U_A^{-i}\tau_j\phi \\ &= \sum_{z\in\mathbb{Z}^n} h_z U_A^{-i}\tau_j\phi_{1,z} \\ &= \sum_{z\in\mathbb{Z}^n} h_z U_A^{-i}\phi_{1,z+Aj} \\ &= \sum_{z\in\mathbb{Z}^n} h_z \phi_{i+1,z+Aj} \\ &= \sum_{z\in\mathbb{Z}^n} h_{z-Aj}\phi_{i+1,z}.\end{aligned}$$

Similarly,

$$\psi^l_{i,j} = \sum_{z\in\mathbb{Z}^n} g^l_{z-Aj}\phi_{i+1,z}.$$

Thus we find that

$$\begin{aligned}s_{i,j} &= \langle f, \phi_{i,j}\rangle \\ &= \left\langle f, \sum_{z\in\mathbb{Z}^n} h_{z-Aj}\phi_{i+1,z}\right\rangle \\ &= \sum_{z\in\mathbb{Z}^n} h_{z-Aj}s_{i+1,z}, \qquad (1.9)\end{aligned}$$

and similarly

$$w^l_{i,j} = \sum_{z\in\mathbb{Z}^n} g^l_{z-Aj}s_{i+1,z}. \qquad (1.10)$$

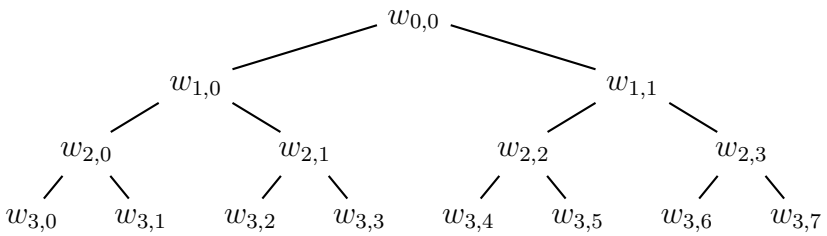


Figure 1.11: Base of the wavelet tree for an $L^2(\mathbb{R})$ function, with $A = 2$.

This is the decomposition of higher resolution scaling coefficients into lower resolution scaling coefficients and wavelet coefficients.

By comparing the two series for f, and since the expansion in each basis is unique, we have

$$s_{i+1,z} = \sum_{j \in \mathbb{Z}^n} h_{z-Aj} s_{i,j} + \sum_{l=1}^{q-1} \sum_{j \in \mathbb{Z}^n} g^l_{z-Aj} w^l_{i,j}. \tag{1.11}$$

This is the reconstruction of the higher resolution scaling coefficients from the lower resolution scaling coefficients and wavelet coefficients.

A multiresolution analysis therefore lets you study a function at various scales, or resolutions. It also gives a well-defined tree structure for the coefficients. An example is shown in Figure 1.11.

Example 1.4.27 Let ϕ be the one-dimensional Haar scaling function, with filter coefficients $h_0 = h_1 = g_0 = -g_1 = \frac{1}{\sqrt{2}}$. Perform the wavelet decomposition of the function

$$f = 5\phi_{2,0} + 3\phi_{2,1} - 2\phi_{2,2} + \phi_{2,3}.$$

Given the filter coefficients, we can rewrite the decomposition equations as

$$s_{i,j} = \frac{1}{\sqrt{2}}(s_{i+1,2j} + s_{i+1,2j+1})$$
$$w_{i,j} = \frac{1}{\sqrt{2}}(s_{i+1,2j} - s_{i+1,2j+1}).$$

Hence we obtain the decomposition given in Table 1.4.

$s_{2,0} = 5$	$s_{2,1} = 3$	$s_{2,2} = -2$	$s_{2,3} = 1$
$s_{1,0} = 8/\sqrt{2}$	$s_{1,1} = -1/\sqrt{2}$	$w_{1,0} = \sqrt{2}$	$w_{1,1} = -3/\sqrt{2}$
$s_{0,0} = 7/2$	$w_{0,0} = 9/2.$		

Table 1.4: Mallat decomposition of a simple function.

The wavelet decomposition is therefore the tree of wavelet coefficients and the remaining scaling coefficient $s_{0,0}$. These can be written in tree form as shown in Figure 1.12.

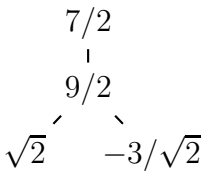


Figure 1.12: Decomposition tree of a simple function.

Example 2.1.7 will demonstrate that this is not always possible for an MRA.

In addition to the decomposition, MRA have a second important property. The scaling functions act like Dirac delta functions, as the following result of Daubechies shows [20].

Theorem 1.4.28 Given a continuous function $f \in L^2(\mathbb{R}^n)$,

$$\lim_{i\to\infty} q^i \int_{\mathbb{R}^n} f(x+y)\phi(A^i y)dy = f(x).$$

Therefore, if we let $s_{i,j} = \langle f, \phi_{i,j}\rangle$, then

$$s_{i,j} \approx q^{-i/2} f(A^{-i} j),$$

for large enough values of i.

In practice, Theorem 1.4.28 is used to sample a function to be studied. Then, the decomposition algorithm is performed on the resulting scaling coefficients. Subsequently, the wavelet coefficients are studied, or acted upon. Then, the reconstruction algorithm is performed on the modified wavelet coefficients to build a new function.

One problem in performing the decomposition algorithm is that at each level it is necessary to compute for each $s_{i,z}$ and each $k \in K$, all the values of j such that $j = A^{-1}(z-k)$. If we start with wavelet coefficients and perform the reconstruction algorithm to reconstruct the original function, it is necessary to find for each j the value of k and z such that $j = A^{-1}(z-k)$. In Chapter 2 we present a case where these relationships between the j and z can be calculated beforehand and stored, hence reducing computation time of the algorithm.

For further theory on the algorithm, and wavelets in general, see Daubechies' book *Ten Lectures on Wavelets* [20, p. 202]. Note that one never needs to compute actual Fourier-type integrals to determine the wavelet expansion coefficients, unlike the situation for non-wavelet bases.

1.4.4 Fundamentals of IFS on Wavelet Trees

The final components of the background material are the topics of IFS on wavelet trees and the associated inverse problem. This theory has been developed in major part by Forte, Mendivil, van de Walle and Vrscay, among others. Some references include [27, 65, 66, 67, 86].

Given the natural relation between $L^2(X)$ and $\ell^2(\mathbb{Z})$, an IFSM operator T can be associated with an operator M on $\ell^2(\mathbb{Z})$. Let (q_i) be an orthonormal basis of $L^2(X)$. Then

a function $u \in L^2(\mathbb{R})$ can be written as

$$u = \sum_i u_i q_i, \tag{1.12}$$

where $u_i =< u, q_i >$. The association between T and M is illustrated by the following commutative diagram:

$$\begin{array}{ccc} L^2(X) & \xrightarrow[\cong]{F} & \ell^2(\mathbb{Z}) \\ \Big\downarrow T & & \Big\downarrow M \\ L^2(X) & \xrightarrow[\cong]{F} & \ell^2(\mathbb{Z}) \end{array}$$

where F is the isomorphism between $L^2(X)$ and $\ell^2(\mathbb{Z})$ given by $F(\sum_n c_n q_n) = (c_n)$.

The function F is an isomorphism of Hilbert spaces. Hence, if T is affine and linear, then so will M be. Also, if T is contractive, then M is contractive. We will consider this relation in the case when (q_i) is a wavelet basis.

Theorem 1.4.29 Let (q_n) be an orthonormal basis of $L^2(\mathbb{R})$ and F be its associated transform as given above. Then T is a contractive operator on $L^2(\mathbb{R})$ with fixed point $\bar{u}_T$ if and only if the operator $M = F \circ T \circ F^{-1}$ is contractive on $\ell^2(\mathbb{Z})$ with fixed point $\bar{\mathbf{u}}_M$, where $\bar{\mathbf{u}}_M = F\bar{u}_T$.

Consider the case when T is the associated IFSM operator of an N-map affine IFSM on $X = L^2(\mathbb{R})$. Then given $u \in L^2(\mathbb{R})$, if $v = Tu$,

$$v = \sum_m v_m q_m,$$

where

$$\begin{aligned} v_m &= < v, q_m > \\ &= < Tu, q_m > \\ &= \sum_{k=1}^{N} \alpha_k < u \circ w_k^{-1}, q_m > + \sum_{k=1}^{N} \beta_k < \chi_{w_k(X)}, q_m > . \end{aligned}$$

Therefore, by Equation (1.12) on page 63,

$$\begin{aligned} v_m &= \sum_{k=1}^{N} \alpha_k \sum_n u_n < q_n \circ w_k^{-1}, q_m > + \sum_{k=1}^{N} \beta_k < \chi_{w_k(X)}, q_m > \\ &= \sum_n a_{mn} u_n + e_m, \end{aligned} \tag{1.13}$$

where $a_{mn} = \sum_{k=1}^{N} \alpha_k < q_n \circ w_k^{-1}, q_m >$ and $e_m = \sum_{k=1}^{N} \beta_k < \chi_{w_k(X)}, q_m >$.

By Equation (1.13), we get the following result [27]:

Proposition 1.4.30 Let (q_n) be an orthonormal basis of $L^2(\mathbb{R})$ with associated transform F. If T is an affine IFSM on $L^2(\mathbb{R})$, then $M = F \circ T \circ F^{-1}$ is an *affine IFS on coefficients* (*IFSC*) on $\ell^2(\mathbb{Z})$ and has the form $M\mathbf{u} = A\mathbf{u} + \mathbf{e}$, where $A = (a_{mn})$ and $\mathbf{e} = (e_m)$, $m, n \in \mathbb{Z}$ are given above.[11]

In general, the matrix A is not sparse, for example with the Discrete Cosine Transform [67]. However, due to the localization properties of wavelets, many of the entries of A will vanish.

Example 1.4.31 Let $X = [0, 1]$ and T be the operator defined by

$$Tu(x) = \frac{1}{2} u(2x) + \frac{1}{2} u(2x - 1) \quad x \in [0, 1].$$

This is the IFSM operator of the affine IFSM given by $w_1(x) = \frac{1}{2}x$, $w_2(x) = \frac{1}{2}(x + 1)$ and

[11] By an IFSC, we mean an operator which acts on sequences in some "IFS"-type manner.

$\phi_1(t) = \phi_2(t) = \frac{1}{2}t$. The fixed point of T is $\bar{u}_T \equiv 0$. Consider the operator M given when (q_n) is chosen to be the Haar basis on $L^2[0,1]$. We assign the following ordering to the basis elements:

$$\begin{aligned} q_{-1} &= \phi; \\ q_{2^i+j-1} &= \psi_{i,j}, \quad i \in \mathbb{N}, 0 \leq j \leq 2^i - 1. \end{aligned}$$

Then the operator A is

$$A = \frac{1}{2\sqrt{2}} \begin{pmatrix} \sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 0 & \ddots \end{pmatrix}$$

The above proposition only works one way. A general question to resolve is, "Given an affine IFSC M on $\ell^2(\mathbb{Z})$, is the operator T an affine IFSM?" The question is therefore:

Question 1.4.32 Given an affine IFSC M on $\ell^2(\mathbb{Z})$, defined by $M\mathbf{u} = A\mathbf{u} + \mathbf{e}$ with $A = (a_{mn})$ and $\mathbf{e} = (e_k)$, does there exist an N-map IFSM $(\mathbf{w}, \mathbf{\Phi})$, for some orthonormal

basis (q_n) of $L^2(\mathbb{R})$, such that

$$a_{mn} = \sum_{k=1}^{N} \alpha_k < q_n \circ w_k^{-1}, q_m >$$

and

$$e_m = \sum_{k=1}^{N} \beta_k < \chi_{w_k(X)}, q_m >$$

for $m, n \in \mathbb{Z}$?

A case where this question has been solved is for LIFSW.

1.4.5 Local IFS on Wavelet Coefficients

In this subsection we present briefly the ideas of local IFS on wavelet coefficients (LIFSW) in the 1-dimensional case as presented in [67]. The 2-dimensional extension can be found in [86]. In Section 2.5, we define a generalized LIFSW in higher dimensions.

Let ϕ be a scaling function with MRA (V_m) and ψ be the associated mother wavelet. For simplicity, we focus our attention on functions $f \in L^2(\mathbb{R})$ that have expansions

$$f = a_{00}\phi + \sum_{i=0}^{\infty} \sum_{j=0}^{2^i-1} b_{i,j}\psi_{i,j}$$

where $a_{00} =< f, \phi >$ and $b_{i,j} =< f, \psi_{i,j} >$. Assume ψ has compact support on $\mathbb{R}$. The expansion coefficients can be written in a meaningful way as shown in Figure 1.13 on the facing page, where $B_{i,j}$ represents the subtree of coefficients with node $b_{i,j}$, and is called the *block* $B_{i,j}$. We say that the coefficients $b_{i,j}$ are on level i, and $a_{0,0}$ is at level -1. The above diagram is called the wavelet (coefficient) tree of f and is denoted by B^f.

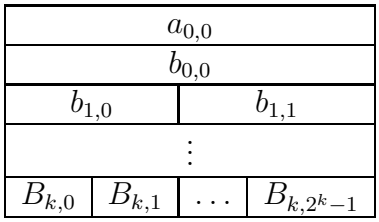


Figure 1.13: Block representation of a wavelet tree.

With the idea of IFSM in mind, one considers an operator on this wavelet tree.

Definition 1.4.33 Given an MRA with scaling function ϕ, mother wavelet ψ and dilation factor 2, define an operator W as follows: Let $k, k^* \in \mathbb{Z}$, with $k \geq 0, k^* > k$, and $\alpha_j \in \mathbb{R}, 0 \leq j \leq 2^{k^*} - 1$. Define the action of the operator W on a wavelet tree $B_{0,0}$ as $W(B_{0,0}) = B^*_{0,0}$, where the nodes of $B^*_{0,0}$ are given by

$$\begin{aligned} a^*_{0,0} &= a_{0,0}, & \\ b^*_{i,j} &= b_{i,j} & 0 \leq i \leq k^* - 1, 0 \leq j \leq 2^i - 1, \\ b^*_{k^*+m,2^m j+n} &= \alpha_j b_{k+m,2^m l(j)+n} & m \geq 0, 0 \leq n \leq 2^m - 1, \\ & & 0 \leq j \leq 2^{k^*} - 1, l(j) \in \{0, 1, \ldots, 2^k - 1\}. \end{aligned}$$

In other words, W leaves the nodes of $B_{0,0}$ untouched at levels less than k^*, and replaces the subtree $B_{k^*,j}$ by $\alpha_j B_{k,l(j)}$. The operator W is called a *local IFS on wavelet coefficients*, or *LIFSW*. The blocks $B_{k,l}, 0 \leq l \leq 2^k - 1$ are called the *parent blocks* (or *domain blocks*), and the blocks $B^*_{k^*,j}, 0 \leq j \leq 2^{k^*} - 1$ are called the *child blocks* (or *range blocks*), since W maps the parent blocks to the child blocks. The parameters α_j are called the *scaling factors* of W.

Since W is an IFSC, it can be associated with an IFSM operator T_W.

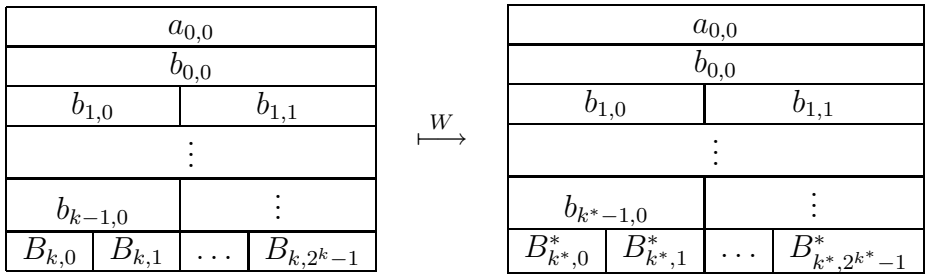


Figure 1.14: Action of W on a wavelet tree.

Definition 1.4.34 Given $f \in L^2(\mathbb{R})$, with wavelet expansion

$$f = a_{0,0}\phi_{0,0} + \sum_{m=0}^{\infty}\sum_{n=0}^{2^m-1} b_{m,n}\psi_{m,n},$$

define the function $f_{k,p}$, for $p \geq 0$, by

$$f_{k,p} = \sum_{m=0}^{\infty}\sum_{n=0}^{2^m-1} b_{k+m,2^m p+n}\psi_{k+m,2^m p+n}.$$

The action of W on the tree B^f is given in Figure 1.14. Now, consider the functions $v_j^* = (T_W f)_{k^*,j}$. All its wavelet coefficients are equal to 0 except that $B^{v_j^*}_{k^*,j} = \alpha_j B^f_{k,l(j)}$. Therefore,

$$v_j^* = \sum_{m=0}^{\infty}\sum_{n=0}^{2^m-1} b^*_{k^*+m,2^m j+n}\psi_{k^*+m,2^m j+n}.$$

However, notice that

$$f_{k,l(j)} = \sum_{m=0}^{\infty}\sum_{n=0}^{2^m-1} b_{k+m,2^m l(j)+n}\psi_{k+m,2^m l(j)+n},$$

and that

$$b^*_{k^*+m,2^m j+n} = \alpha_j b_{k+m,2^m l(j)+n}, \quad m \geq 0, 0 \leq n \leq 2^m - 1. \tag{1.14}$$

One can thus use the scaling and dilation relations between the $\psi_{m,n}$ to write v^*_j as a multiple of $f_{k,l(j)} \circ w_j^{-1}$ for some appropriate function w_j. The function w_j can be calculated as follows:

$$\begin{aligned}
\psi_{k,l(j)}(x) &= 2^{k/2}\psi(2^k x - l(j)), \\
\psi_{k^*,j}(x) &= 2^{k^*/2}\psi(2^{k^*} x - j) \\
&= 2^{(k^*-k)/2}2^{k/2}\psi(2^k w_j^{-1}(x) - l(j)).
\end{aligned}$$

Equating the arguments of ψ yields

$$w_j^{-1}(x) = 2^{k^*-k}x + \frac{l(j) - j}{2^k}, \tag{1.15}$$

and hence by Equations (1.14) and (1.15),

$$v^*_j(x) = 2^{(k^*-k)/2}\alpha_j f_{k,l(j)}\left(2^{k^*-k}x + \frac{l(j) - j}{2^k}\right).$$

The function

$$w_j(x) = 2^{k-k^*}x + \frac{j - l(j)}{2^{k^*}}$$

has contractivity factor 2^{k-k^*}. The resulting operator T_W is called a *recurrent (vector) IFSM with condensation* (c.f. [16, 67]). By this is meant that T_W acts between orthogonal components of the wavelet tree and has condensation function

$$\sum_{m=0}^{k^*-1}\sum_{n=0}^{2^m-1} b_{m,n}\psi_{m,n}.$$

A useful space, over which T_W is contractive, was constructed in [27]. Let $u \in L^2(\mathbb{R})$ and let W be as above. Let

$$C_w(u,k^*) = \{a_{0,0}, b_{m,n}, m \geq 0, 0 \leq n \leq 2^m - 1 : \sum |b_{m,n}|^2 < \infty$$
$$\text{with } b_{m,n} =< u, \psi_{m,n} >, 0 \leq m \leq k^* - 1, 0 \leq n \leq 2^m - 1\}.$$

Consider the metric d_w on C_w by

$$d_w(\mathbf{c},\mathbf{d}) = \max_{0\leq j\leq 2^{k^*}-1} \Delta_j^2,$$

where

$$\Delta_j^2 = \sum_{m=0}^{\infty}\sum_{n=0}^{2^m-1} (b^{\mathbf{c}}_{k^*+m,2^m j+n} - b^{\mathbf{d}}_{k^*+m,2^m j+n})^2,$$

where $b^{\mathbf{c}}$ and $b^{\mathbf{d}}$ refer to the wavelet coefficients of $\mathbf{c}$ and $\mathbf{d}$ respectively. By the completeness of $\ell^2(\mathbb{Z})$ it follows that

Proposition 1.4.35 The metric space $(C_w(u,k^*), d_w)$ is complete.

In addition, the following result holds:

Proposition 1.4.36 For $\mathbf{c}, \mathbf{d} \in C_w(u,k^*)$,

$$d_w(W\mathbf{c}, W\mathbf{d}) \leq c_w d_w(\mathbf{c},\mathbf{d}), \quad c_w = \max_{0\leq j\leq 2^{k^*}-1} |\alpha_j|.$$

Therefore, the BCMP yields the following result:

Corollary 1.4.37 If $c_w < 1$, there exists a unique $\bar{\mathbf{u}} \in C_w(u, k^*)$ such that $W\bar{\mathbf{u}} = \bar{\mathbf{u}}$.

Corollary 1.4.38 Let $\epsilon > 0$ and $\mathbf{c} \in C_w(u, k^*)$. Suppose that W is an LIFSW such that $d_w(\mathbf{c}, W\mathbf{c}) < \epsilon$. Then

$$d_w(\mathbf{c}, \bar{\mathbf{u}}) < \frac{\epsilon}{1 - c_w},$$

where $W\bar{\mathbf{u}} = \bar{\mathbf{u}}$.

Proof The result follows directly from Proposition 1.2.17 on page 13. ■

The following results also hold. They are stated here since they have not been found in the literature.

Theorem 1.4.39 Let W be as above. Letting $W^t(B_{0,0}) = B^t_{0,0}$, $t \geq 0$ then

$$\begin{aligned} a^{t+1}_{0,0} &= a_{0,0}, \\ b^{t+1}_{i,j} &= b_{i,j} \qquad 0 \leq i \leq k^* - 1, 0 \leq j \leq 2^i - 1, \\ b^{t+1}_{i,j} &= b^t_{i,j} \qquad 0 \leq i \leq (t+1)k^* - tk - 1, 0 \leq j \leq 2^i - 1. \end{aligned}$$

That is, the nodes of $B^t_{0,0}$ at level $i < (t+1)k^* - tk$ are invariant under W.

This shows that given a function with finitely many coefficients, the fixed point of the IFSW is constructed after finitely many iterations. Also, in practice, the attractor can therefore be constructed one level at a time.

Theorem 1.4.40 Letting $k = 0$ above (i.e. W is a strict IFSW), then the squared norm of the nodes of $B^t_{0,0}$ from levels tk^* to $(t+1)k^* - 1$ is

$$\sum_{m=0}^{k^*-1} \sum_{n=0}^{2^m-1} |b_{m,n}|^2 \left(\sum_{j=0}^{2^{k^*}-1} |\alpha_j|^2 \right)^t.$$

Theorem 1.4.41 Let $k = 0$ above (i.e. W is a strict IFSW). Then the attractor $B_{0,0}^{\infty} = \lim_{t\to\infty} W^t(B_{0,0})$ of W exists if and only if

$$\sum_{j=0}^{2^{k^*}-1} |\alpha_j|^2 < 1.$$

In that case, the norm of the attractor is

$$\sum_{m=0}^{k^*-1} \sum_{n=0}^{2^m-1} |b_{m,n}|^2 \sum_{t=0}^{\infty} \left(\sum_{j=0}^{2^{k^*}-1} |\alpha_j|^2 \right)^t.$$

Example 1.4.42 Consider the following operator

$$W : B_{00} \mapsto \begin{array}{|c|c|} \hline \multicolumn{2}{|c|}{b_{00}} \\ \hline \alpha_0 B_{00} & \alpha_1 B_{00} \\ \hline \end{array}.$$

We have $k = 0, k^* = 1, l(0) = l(1) = 0$. Therefore

$$w_0^{-1}(x) = 2x + \frac{0-0}{2^0} = 2x$$
$$w_1^{-1}(x) = 2x + \frac{0-1}{2^0} = 2x - 1.$$

Also, $f = f_{0,0}$, hence

$$\begin{aligned} T_W f(x) &= b_{0,0}\psi_{0,0}(x) + v_0^*(x) + v_1^*(x) \\ &= b_{0,0}\psi_{0,0}(x) + \sqrt{2}\alpha_0 f_{0,l(0)} \circ w_0^{-1}(x) + \sqrt{2}\alpha_1 f_{0,l(1)} \circ w_1^{-1}(x) \\ &= b_{0,0}\psi_{0,0}(x) + \sqrt{2}\alpha_0 f_{0,0} \circ w_0^{-1}(x) + \sqrt{2}\alpha_1 f_{0,0} \circ w_1^{-1}(x) \\ &= b_{0,0}\psi_{0,0}(x) + \sqrt{2}\alpha_0 f(2x) + \sqrt{2}\alpha_1 f(2x-1). \end{aligned}$$

It is important to note that the IFSM operator T_W, associated with an LIFSW operator W, depends upon the particular wavelet basis chosen. We illustrate this with some pictures. The attractors of T_W, using the Coifman-6, Daubechies-2 and Haar wavelets, are shown in Figure 1.15 on the next page, where $a_{0,0} = 0$, $b_{0,0} = 1$, $\alpha_1 = 0.2$ and $\alpha_2 = 0.3$. Note

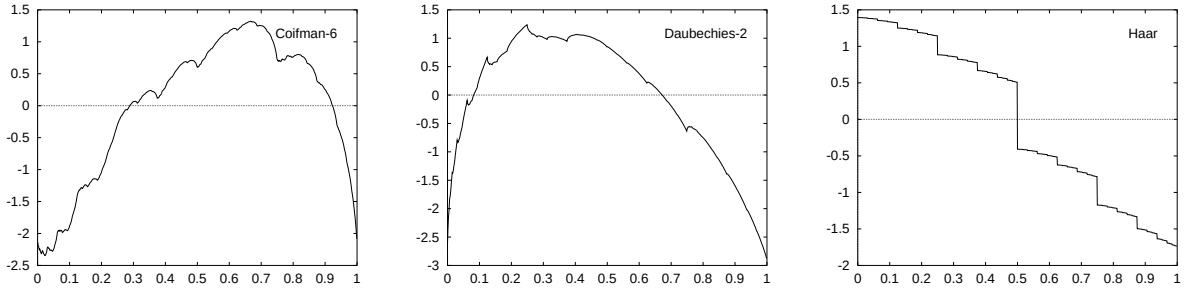


Figure 1.15: The attractors of T_W in Example 1.4.42 using Coifman-6, Daubechies-2 and Haar wavelets.

the dependence of the attractor on the basis chosen. However, the attractor of W is basis independent.

Given a target function $f \in L^2(\mathbb{R})$, we can use Corollary 1.4.38 to construct an LIFSW on its coefficient tree [27]. The squared L^2 distance associated with each range block $B^*_{k^*,j}$ and domain block $B_{k,l}$ is given by

$$\Delta^2_{j,l} = \sum_{m=0}^{\infty} \sum_{n=0}^{2^m-1} (c_{k^*+m,2^m j+n} - \alpha_j c_{k+m,2^m l+n})^2.$$

The optimal scaling factor $\bar{\alpha}_{j,l}$ given by the least squares minimization is [12]

$$\bar{\alpha}_{j,l} = \frac{S_{k^*,j,k,l}}{S_{k,l,k,l}}$$

where

$$S_{a,b,c,d} = \sum_{m=0}^{\infty} \sum_{n=0}^{2^m-1} c_{a+m,2^m b+n} c_{c+m,2^m d+n}.$$

[12]Given that each node has the same number of children, this summation can be thought of naturally as an inner product of coefficient trees at different levels (see [86]).

The minimized collage distance is then

$$\Delta_{j,l}^{\min} = [S_{k^*,j,k^*,j} - \bar{\alpha}_{j,l} S_{k^*,j,k,l}]^{1/2}.$$

Thus, as with LIFSM, for each range block $B_{k^*,j}^*$, choose the domain block $B_{k,l(j)}$ for which $\Delta_{j,l(j)}^{\min}$ is minimized. Then, iterate the associated operator W on any initial $\mathbf{c} \in C_w(f, k^*)$. For simplicity, one can let $\mathbf{c}$ be the sequence with $c_{m,n} = 0$ for all $m \geq k^*$. The function $\bar{u}$ associated with the fixed point $\bar{\mathbf{u}}$ of W is then given by

$$\bar{u} = \bar{a}_{0,0}\phi + \sum_{m=0}^{\infty} \sum_{n=0}^{2^m-1} \bar{b}_{m,n} \psi_{m,n}.$$

Section 2.6 generalizes this result to the case of LIFSW associated with number systems.

Chapter 2

Linking Complex Bases, Number Systems, Wavelets and Fractal-Wavelet Transforms

2.1 Linking Complex Bases to Wavelets

In the following sections, it is assumed that the scaling and wavelet functions arise from multidimensional Haar bases and filters. It is possible that certain results can be generalized to other wavelet bases.

The results in this chapter stem from the realization that the theory of MRA and complex bases are related through a natural association. In order to be able to use this terminology later, we define it here:

Definition 2.1.1 The *natural association* between $\mathbb{C}$ and $\mathbb{R}^2$ is given by $s + it \leftrightarrow (s,t)^T$.

Through the natural association, multiplication by the base $b = c+id$ in $\mathbb{C}$ is equivalent

to multiplication in $\mathbb{R}^2$ by the matrix

$$A = \begin{pmatrix} c & -d \\ d & c \end{pmatrix},$$

that is $(c+id)(s+it) \leftrightarrow A(s,t)^T$. For simplicity of reading, we will generally write $k = (s,t)$ when referring to elements of $\mathbb{R}^2$, choosing to omit the transposition symbol. Thus it shall be understood that Ak will mean Ak^T.

Lemma 2.1.2 Let $b = c + id \in \mathbb{Z}[i]$ and $D = \{d_1, d_2, \ldots, d_{|b|^2}\} \subset \mathbb{Z}[i]$ such that (b, D) is a valid base. Then the matrix A, arising from the natural association of multiplication by b, is an acceptable dilation for $\mathbb{Z}^2$, and the set $K = \{k_1, k_2, \ldots, k_q\}$, derived from D through the natural association with $d_l \leftrightarrow k_l$, is a complete residue system of $\mathbb{Z}^2/A\mathbb{Z}^2$, with $q = |b|^2$.

Proof Since the entries of A are integers, then $A\mathbb{Z}^2 \subset \mathbb{Z}^2$. Since (b, D) is a valid base, $\det A = c^2 + d^2 = |b|^2 \geq 2$. This follows from the fact that D is a complete residue system of $\mathbb{Z}[i]/b\mathbb{Z}[i]$. The eigenvalues of A are given by λ as follows:

$$\begin{aligned} 0 = \det(A - \lambda I) &= (c - \lambda)^2 + d^2 \\ &= \lambda^2 + c^2 - 2c\lambda + d^2 \\ &= (\lambda - (c + id))(\lambda - (c - id)) \end{aligned}$$

and hence $\lambda = c \pm id$. Since $c^2 + d^2 \geq 2$, then $|\lambda| = |c \pm id| > 1$. Therefore, A is an acceptable dilation for $\mathbb{Z}^2$.

To show that K is a complete residue system of $\mathbb{Z}^2/A\mathbb{Z}^2$, let $v = (x, y)^T \in \mathbb{Z}^2$. Then let $z = x + iy \in \mathbb{Z}[i]$. Since (b, D) is a valid base, D is a complete residue system of $\mathbb{Z}[i]/b\mathbb{Z}[i]$ by Proposition 1.3.9, hence there is a unique $d_l \in D$ such that $z = mb + d_l$ for some $m = s + it \in \mathbb{Z}[i]$. Therefore, through the natural association, $v = A(s,t)^T + k_l$ and k_l is the unique element of K satisfying such a relation for an $(s,t) \in \mathbb{Z}^2$. If not, then reversing the previous argument, there would be some $d_p \neq d_l$ in D satisfying $z = nb + d_p$ for some $n \in \mathbb{Z}[i]$, which is a contradiction to the uniqueness of d_l. ■

Notation 2.1.3 The terminology (A, K) *is associated with a valid base* (b, D) will mean associated in the sense of Lemma 2.1.2.

Lemma 2.1.4 Let b, D, A, K be as defined in Lemma 2.1.2. Then the set

$$Q = \left\{ \sum_{i=1}^{\infty} A^{-i} k_i \ : \ k_i \in K \right\}.$$

has two-dimensional Lebesgue measure 1.

Proof Since (b, D) is a valid base, the set $T = T(b, D) = \left\{ \sum_{j=-\infty}^{-1} a_j b^j \ : \ a_j \in D \right\}$ tiles the plane by integer translates by Theorem 1.3.10. Therefore, $|T| \geq 1$. To show $|T| \leq 1$ consider its associated IFS $w = \{w_l : l = 1, \ldots, |b|^2\}$, with

$$w_l(z) = \frac{(z + d_l)}{b}.$$

For each $w_l \in w$ and any set S in $\mathbb{R}^2$, $|w_l(S)| = |S|/|b|$, hence $|w(S)| \leq |S|$. Therefore, starting the IFS on the unit square with vertices $\{0, 1, i, 1 + i\}$, we have $|T| \leq 1$. Hence, $|T| = 1$. Through the natural association

$$T(b, D) = \left\{ \sum_{j=-\infty}^{-1} a_j b^j \ : \ a_j \in D \right\} \equiv \left\{ \sum_{i=1}^{\infty} A^{-i} k_i \ : \ k_i \in K \right\} = Q.$$

Hence, Q has measure 1. ■

Therefore, the following result holds:

Theorem 2.1.5 Let b, D, A, K, Q be as defined in Lemma 2.1.4. Then $\phi = \chi_Q$ is a scaling function for an MRA associated with $(\mathbb{Z}^2, A)$.

Proof By Lemma 2.1.2, A is an acceptable dilation and K is a complete residue system of $\mathbb{Z}^2/A\mathbb{Z}^2$. By Lemma 2.1.4, Q has measure 1 and satisfies the form of Theorem 1.4.19. Therefore, by Theorem 1.4.19, the result holds. ■

Notation 2.1.6 An *MRA associated with* or *arising from a valid base* will mean an MRA of the form of Theorem 2.1.5.

To see why these types of MRA are of interest, consider the following example of the decomposition algorithm.

Example 2.1.7 Recall Example 1.4.26:

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

with $K = \{(0,0),(1,0)\}$. Note that this is not associated with a valid base, a point which will become clear shortly (see Theorem 1.3.12).

Consider the simple function $f = \phi_{(0,0)} + \phi_{(1,0)} + \phi_{(0,1)} + \phi_{(1,1)}$. Recall from page 46 the notation $g_j = g_{0,j}$. Its scaling coefficients are $s_{0,(0,0)} = s_{0,(1,0)} = s_{0,(0,1)} = s_{0,(1,1)} = 1$. We will now perform the decomposition algorithm on these scaling coefficients.

To perform the decomposition algorithm from level $i+1$ to level i, recall the decomposition equation (Equation (1.9) on page 59):

$$s_{i,j} = \sum_{z \in \mathbb{Z}^n} h_{z-Aj} s_{i+1,z}, \quad j \in \mathbb{Z}^2.$$

In this example, the scaling filters are

$$h_{z-Aj} = \begin{cases} 1/\sqrt{2} & \text{for } z - Aj \in K, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, to find the non-zero $s_{i,j}$, we need to solve the equations $z - Aj = (0,0) = k_0$ and $z - Aj = (1,0) = k_1$ for $z, j \in \mathbb{Z}^2$, or

$$j = A^{-1}z \text{ and } j = A^{-1}\left(z - \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right).$$

The inverse of A is

$$A^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Now, define the following sets:

$$G_{0,0} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\},$$
$$G_{0,1} = G_{0,0} - k_1$$

and for all negative $i \in \mathbb{Z}$, define

$$F_i = (A^{-1}G_{i+1,0} \cup A^{-1}G_{i+1,1}) \cap \mathbb{Z}^2,$$
$$G_{i,0} = F_i$$

and

$$G_{i,1} = F_i - k_1.$$

By the definition of the decomposition algorithm, these sets F_i precisely give the non-zero scaling coefficients at level $i < 0$. Then, the sets are as follows:

$$G_{0,0} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}, \quad G_{0,1} = \left\{ \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\},$$
$$F_{-1} = \left(\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \bigcup \left\{ \begin{pmatrix} -1/2 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \right\} \right) \bigcap \mathbb{Z}^2$$
$$= \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\},$$
$$G_{-1,0} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}, \quad G_{-1,1} = \left\{ \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\},$$
$$F_{-2} = \left(\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix} \right\} \bigcup \left\{ \begin{pmatrix} -1/2 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \right) \bigcap \mathbb{Z}^2$$
$$= \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\},$$
$$G_{-2,0} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \quad G_{-2,1} = \left\{ \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\},$$
$$F_{-3} = \left(\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \right\} \bigcup \left\{ \begin{pmatrix} -1/2 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \right) \bigcap \mathbb{Z}^2$$

$$= \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

$$\implies F_i = F_{i+1}, \quad G_{i,0} = G_{i+1,0}, \quad G_{i,1} = G_{i+1,1}, \quad \forall i < -2.$$

Therefore, we have the tree of scaling coefficients created from the coefficients on level 0, as illustrated in Figure 2.1. Since this MRA has two digits, each $s_{i,j}$ is technically constructed from two coefficients below it. However, some coefficients are originally zero and hence many coefficients are essentially not used. These coefficients are designated by the branches with no nodes at their end. This situation arises from the fact that the support of the function is contained in, but does not cover an expanded copy of the principal tile. This point is of particular interest and is discussed in Section 2.4.

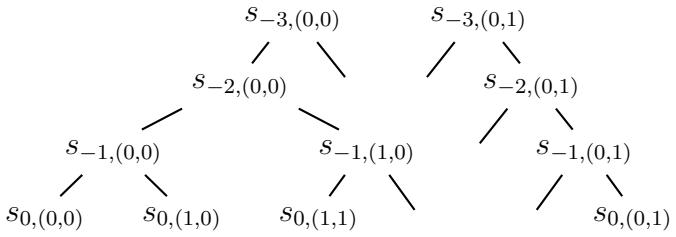


Figure 2.1: Mallat decomposition with non-intersecting branches.

We see from Figure 2.1 that using this MRA, the decomposition creates two non-intersecting trees in the sense that as of level $i = -2$, there are always exactly two non-zero scaling coefficients: $s_{i,(0,0)}$ and $s_{i,(0,1)}$.

The above situation cannot happen with an MRA associated with a valid base. However, before investigating this further, it will be useful to first translate the Mallat algorithm for MRA associated with valid bases.

2.2 Mallat Algorithm Revisited

Consider the reconstruction and decomposition algorithms for an MRA associated with an acceptable dilation A for $\mathbb{Z}^2$, with complete residue system K. Assume (A, K) is associated with a complex base (b, D). Rewriting Equation (1.9) for the Mallat decomposition, we have

$$s_{i,j} = \sum_{z \in \mathbb{Z}^2} h_{z-Aj} s_{i+1,z} \tag{2.1}$$

$$= \sum_{l \in \mathbb{Z}^2} h_l s_{i+1,l+Aj},$$

$$= \sum_{k \in K} h_k s_{i+1,Aj+k}, \tag{2.2}$$

since $h_k = 0$ when $k \notin K$. We can translate this identity to the language of complex bases as

$$s_{i,j} = \sum_{d \in D} h_d s_{i+1,bj+d}, \tag{2.3}$$

where the scaling filters and coefficients are re-indexed via the natural association. The reconstruction algorithm can be converted in a similar fashion. Following the same reasoning as above, we rewrite Equation (1.11) for the Mallat reconstruction:

$$\begin{aligned} s_{i+1,z} &= \sum_{j\in\mathbb{Z}^2} h_{z-Aj} s_{i,j} + \sum_{l=1}^{q-1} \sum_{j\in\mathbb{Z}^2} g^l_{z-Aj} w^l_{i,j} \\ &= \sum_{p\in z-A\mathbb{Z}^2} h_p s_{i,A^{-1}(z-p)} + \sum_{l=1}^{q-1} \sum_{p\in z-A\mathbb{Z}^2} g^l_p w^l_{i,A^{-1}(z-p)} \\ &= \sum_{k\in K} h_k s_{i,A^{-1}(z-k)} + \sum_{l=1}^{q-1} \sum_{k\in K} g^l_k w^l_{i,A^{-1}(z-k)} \end{aligned} \tag{2.4}$$

since $h_k = 0$ when $k \notin K$. We can translate this identity to the language of complex bases as

$$s_{i+1,z} = \sum_{d\in D} h_d s_{i,b^{-1}(z-d)} + \sum_{l=1}^{q-1} \sum_{d\in D} g^l_d w^l_{i,b^{-1}(z-d)}, \tag{2.5}$$

where $b^{-1}(z-d)$ must be a Gaussian integer, by definition of $s_{i,b^{-1}(z-d)}$ and each $w^l_{i,b^{-1}(z-d)}$.

For the moment, let us examine the decomposition algorithm in the language of complex bases. Here, we can consider $bj + d$ as the address of j shifted to the left by one place, with d added as the zero-th order digit. We can then think of the index i as the length of the address of j. The scaling coefficient associated with the point j is thus a linear combination of the coefficients of the points whose addresses are one digit longer than that of j, and start with the address of j. In other words, letting j have address $(a_t \dots a_1)$ in base b, one could write the following equation:

$$s_{-t,(a_t \dots a_1)} = \sum_{d\in D} h_d s_{-t+1,(a_t \dots a_1 d)}.$$

However, this simplistic conversion to address notation may lead to confusion. For example, in base $(-1+i, \{0,1\})$, how does one denote the scaling coefficient derived from $s_{0,(0)}$ and

$s_{0,(1)}$? Also, one needs to put these coefficients in a tree, and *a priori* it is not clear where to put $s_{0,j}$ if each j has a different address length. One of the nice properties of the addressing in practice is the ability to read in a point from a file and directly place it in the appropriate location in the scaling tree. Therefore, this question must be addressed – literally.

In practice, one considers a set of scaling coefficients. These are often samples of a given function on a set of lattice points using Theorem 1.4.28 of Daubechies. For simplicity, we can relabel the scale so that $i = 0$. Therefore, the goal is to decompose the set $\{s_{0,z}\}$ for some finite set X of integer indices $z \in \mathbb{Z}[i]$ into scaling and wavelet coefficients at levels $i < 0$.

In order to simplify our discussion, it will be helpful to think of the addresses in the subscripts as codes. We think of D^t as the code space of words with t letters taken from the alphabet D (see Notation 2.2.6). We generally write these words as concatenations of the digits rather than comma-separated sequences. However, when giving specific examples, we will place a "$\cdot$" between the digits to improve legibility. In this way, for example, we will be able to use concatenation of sequences as we perform the Mallat algorithm. To do this, we will introduce the following notation.

Notation 2.2.1 Let (b, D) be a valid base. If $z = \sum_{k=0}^{u-1} d_k b^k$, $d_k \in D$ then define, for each $t \in \mathbb{N}$, the *address* $a^t_{z,(b,D)}$ in D^t as

$$a^t_{z,(b,D)} = (d_{t-1} \dots d_1 d_0)_{(b,D)},$$

where $d_k = 0$ for $k \geq u$, and with $a^0_{z,(b,D)}$ defined as the empty sequence. When the base is

understood, we will simply write a_z^t. Then, define the *minimal length* of $z \in \mathbb{Z}[i]$ as

$$t_z = \min\left\{t \in \mathbb{N} : z = \sum_{k=0}^{t-1} d_k b^k, d_k \in D\right\}$$

and define

$$a_z = a_{z,(b,D)}^{t_z}.$$

The address a_z is called the *minimal address of* z *in* (b, D).

The number t is the length of the address a_z^t. Also, since (b, D) is a valid base, $t_z > 0$ is well defined and hence a_z is unique and is simply the positional notation given in Definition 1.3.3. Also, t_z is the largest value of t such that $d_{t-1} \neq 0$.

Example 2.2.2 In the base $(-1+i, \{0, 1\})$, the point $-1+2i = b^4 + b^3 + 1$, hence the addresses $a_{-1+2i,(b,D)}^t$ are as follows:

$$a_{-1+2i,(b,D)}^t = \begin{cases} \emptyset & t = 0, \\ (1) & t = 1, \\ (01) & t = 2, \\ (001) & t = 3, \\ (1001) & t = 4, \\ (11001) & t = 5, \\ (0\ldots011001) & t > 5. \end{cases}$$

Therefore, $t_{-1+2i} = 5$ and $a_{-1+2i} = (11001)$. We define $\emptyset$ as the address of length zero. As will be shown later, this is an acceptable minimal address for $z = 0$.

Notation 2.2.3 Let $X \subset \mathbb{Z}[i]$ be a finite set of integers and (b, D) be a valid base. Then the *length of* X *in* (b, D) is defined as

$$L_{X,(b,D)} = \max\{t_z : z \in X\}.$$

Again for simplicity of notation, if the base is understood we simply write L. Finally, let $A = A_X = \{a_z^L : z \in X\}$.

Example 2.2.4 Consider the set $X = \{-3, -2, \ldots, 3\} \times \{-2i, -i, \ldots, 3i\}$. To determine the length of X in base $(-1+i, \{0,1\})$, we need to generate the minimal address of each point of X in this base. Figure 2.2 presents them.

$3i$	11010	11011	1110110	1110111	1010	1011	1100110
$2i$	1110100101	11000	11001	1110100	1110101	1000	1001
i	11110	11111	10	11	1110	1111	111010010
0	10001	11100	11101	0	1	1100	1101
$-i$	11101010	11101011	110	111	111010	111011	111010110
$-2i$	10101	11101000	11101001	100	101	111000	111001
	-3	-2	-1	0	1	2	3

Figure 2.2: Minimal addresses of points in a set.

Therefore, $L = 10$ since this is the minimal length of $-3 + 2i$.

We can now properly define scaling and wavelet coefficients associated with a valid base:

Definition 2.2.5 Let $X \subset \mathbb{Z}[i]$ be a finite set of indices. Let ϕ be a scaling function of an MRA with filter coefficients $\{h_d\}$ and $\{g_d^k\}$, $d \in D$, $k = 1, \ldots, q-1$ associated with a valid base (b, D), with $|b|^2 = q$. Suppose $L = L_{X,(b,D)}$. Then if $f = \sum_{z \in X} s_{0,z}\phi_{0,z}$, define, for $z \in X$,

$$s_{0,(d_{L-1}\ldots d_0)} = s_{0,z}.$$

For $l \in \mathbb{N}$, $0 < l < L$, define

$$\begin{aligned} s_{-l,(d_{L-1}\ldots d_l)} &= \sum_{d \in D} h_d s_{-l+1,(d_{L-1}\ldots d_l d)}, \\ w^k_{-l,(d_{L-1}\ldots d_l)} &= \sum_{d \in D} g_d^k s_{-l+1,(d_{L-1}\ldots d_l d)}, \quad k = 1, \ldots, q-1. \end{aligned}$$

Finally, define

$$\begin{aligned} s_{-L,\emptyset} &= \sum_{d\in D} h_d s_{-L+1,(d)}, \\ w^k_{-L,\emptyset} &= \sum_{d\in D} g^k_d s_{-L+1,(d)}, \qquad\qquad k = 1,\dots,q-1. \end{aligned}$$

Since it is cumbersome to write the sequences in full, it is practical to define a more compact notation for the subscripts of these coefficients. We introduce it here for convenience:

Notation 2.2.6 Let $D^t = \{(d_{t-1}\dots d_0) : d_l \in D\}$ with the convention that $D^0 = \{\emptyset\}$. Then, given $\sigma = (d_{t-1}\dots d_0) \in D^t$, define the l-th prefix of σ as

$$_l\sigma = (d_{t-1}\dots d_l),$$

where we define $_t\sigma = \emptyset$.

We can simplify the notation as follows: Let

$$s_{0,a^L_z} = s_{0,z}$$

and for $l \in \mathbb{N}$, $0 < l \le L$, define

$$\begin{aligned} s_{-l,{}_la^L_z} &= \sum_{d\in D} h_d s_{-l+1,{}_la^L_z d}, \\ w^k_{-l,{}_la^L_z} &= \sum_{d\in D} g^k_d s_{-l+1,{}_la^L_z d}, \quad k = 1,\dots,q-1, \end{aligned}$$

where $\sigma\rho = (s_0\dots s_m t_0\dots t_n)$ denotes the concatenation of the two sequences $\sigma = (s_0\dots s_m)$ and $\rho = (t_0\dots t_n)$.[1]

[1]We write $a^L_z d$ instead of $a^L_z(d)$ as a simplification.

Theorem 2.2.7 The definition of $s_{-l,{}_l a_z^L}$ and $w^k_{-l,{}_l a_z^L}$ is consistent with the Mallat algorithm.

Proof Let X, ϕ, $\{h_d\}$, $\{g_d^k\}$, b, D, L and f be as in the statement of Definition 2.2.5. For $z = \sum_{k=0}^{L-1} d_k b^k, d_k \in D$, $0 < l < L$ let ${}_l z = \sum_{k=l}^{L-1} d_k b^{k-l}$ and let ${}_L z = 0$.[2] For $l = 0$, $s_{0,a_z^L} = s_{0,z}$. Now, let $0 < l \leq L$. By the decomposition equation, we have

$$\begin{aligned} s_{-l,{}_l a_z^L} &= s_{-l,{}_l z} \\ &= \sum_{d\in D} h_d s_{-l+1,{}_l z b+d} \\ &= \sum_{d\in D} h_d s_{-l+1,(d_{L-1}\ldots d_l d)} \\ &= \sum_{d\in D} h_d s_{-l+1,{}_l a_z^L d}. \end{aligned}$$

Replacing $s_{-l,{}_l z}$ by $w^k_{-l,{}_l z}$ and h_d by g_d^k in the above equations yields the consistency for the wavelet coefficients.

Now, consider the reconstruction algorithm from level $0 < l \leq L$. First, notice the following relation:

$$\begin{aligned} {}_{l-1} z &= \sum_{k=l-1}^{t-l} d_k b^{k-(l-1)} = \sum_{k=l-1}^{t-1} d_k b^{k-l} b \\ &= \left(\sum_{k=l}^{t-l} d_k b^{k-l} \right) b + d_{l-1} \\ &= {}_l z b + d_{l-1}. \end{aligned} \tag{2.6}$$

By Equation (2.5) on page 82, the reconstruction algorithm is given as

$$s_{-l+1,{}_{l-1} z} = \sum_{d\in D} h_d s_{-l,b^{-1}({}_{l-1} z - d)} + \sum_{k=1}^{q-1} \sum_{d\in D} g_d^k w^k_{-l,b^{-1}({}_{l-1} z - d)},$$

where each $b^{-1}({}_{l-1} z - d)$ must be a Gaussian integer, by definition of $s_{-l,b^{-1}({}_{l-1} z - d)}$ and

[2] The reason that ${}_L z$ is well defined, will be pointed out in the corollary below.

$w^k_{-l,b^{-1}({}_{l-1}z-d)}$. Given $z \in \mathbb{Z}[i]$, this will only occur if ${}_{l-1}z = d + bm$ for some $m \in \mathbb{Z}[i]$. To find such a d, it is necessary to determine which coset of $d + b\mathbb{Z}[i]$ that ${}_{l-1}z$ is in. However, (b, D) is a valid base, hence there is a unique such d when z is expressed in base b, that is $d = d_{l-1}$.[3] Therefore, by Equation (2.6) we have

$$s_{-l+1,{}_{l-1}z} = h_{d_{l-1}} s_{-l,{}_l z} + \sum_{k=1}^{q-1} g^k_{d_{l-1}} w^k_{-l,{}_l z},$$

or

$$s_{-l+1,{}_{l-1}a^L_z} = h_{d_{l-1}} s_{-l,{}_l a^L_z} + \sum_{k=1}^{q-1} g^k_{d_{l-1}} w^k_{-l,{}_l a^L_z}. \qquad \blacksquare$$

We now arrive at the following result that shows that the situation of Example 2.1.7 cannot occur for MRA associated with a valid base.

Corollary 2.2.8 Consider an MRA associated with a valid base (b, D). Suppose $\{s_{0,z} : z \in \mathbb{Z}[i]\}$ is a set of scaling coefficients such that $s_{0,z} = 0$ for all $z \in \mathbb{Z}[i]\backslash J$, for some finite set $J \subset \mathbb{Z}[i]$. Then, there exists a number $t \in \mathbb{N}$ such that $s_{-t,j} = 0 \ \forall j \in \mathbb{Z}[i]\backslash\{0\}$.

Proof Let $j \in J$. Since (b, D) is a valid base, let $a = a_j^{t_j}$ be the unique finite expansion of j in (b, D). By Theorem 2.2.7, the only scaling coefficients arising from j are

$$s_{0,a}, s_{-1,{}_1 a}, s_{-2,{}_2 a}, \ldots s_{-t_j+1,{}_{t_j-1}a} \text{ and } s_{-t_j,\emptyset}.$$

This also clarifies that ${}_L z = 0$, ${}_L a^L_z = \emptyset$ and $\emptyset$ as an address of zero, are well-defined notations. Since J is finite, let $t = \max\{t_j : j \in J\}$. Therefore, $s_{-t,j} = 0 \ \forall j \in \mathbb{Z}[i]\backslash\{0\}$. $\blacksquare$

This result is reminiscent of Theorem 1.3.15. In the next section, we will characterize the MRA in $\mathbb{Z}^n$ such that this also happens. Before moving on, we demonstrate the use of complex bases in the Mallat algorithm.

[3] One could replace $b^{-1}({}_{l-1}z - d)$ by $\Phi({}_{l-1}z)$ (see Theorem 1.3.15).

Example 2.2.9 Consider the function f given in the MRA associated with the base $(-1+i, \{0,1\})$ by $f = 0\phi_0 + \phi_1 + \phi_i + 2\phi_{1+i}$. We are interested in performing the decomposition of the function f in this MRA.

The set $X = \{0, 1, i, 1+i\}$ and the addresses associated with the scaling coefficients of f are

Integer	Address	Value
0	0	0
1	1	1
i	11	1
1+i	1110	2

Therefore, $L = 4$ and hence $A_X = \{(0000), (0001), (0011), (1110)\}$.

We choose the standard Haar wavelet basis as given in Example 1.4.23. We then obtain the scaling and wavelet trees using the decomposition algorithm. The lowest level on the scaling tree represents the initial values of the function. Going down the left branch in a tree represents the digit 0. The right represents the digit 1. For example, the value of the function on the tile at i, which is at address 0011, is placed in the node down the scaling tree left, left, right, right. The entire tree is given in Figure 2.3. In this example, empty nodes are treated as coefficients with value zero. We will discuss extensions in Section 2.4, where the tree is filled in using the Extension Algorithm.

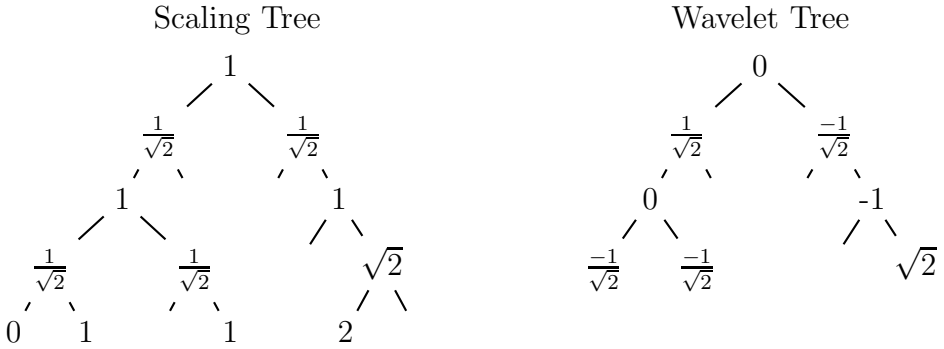


Figure 2.3: Scaling and wavelet trees for a small set of points.

Notice the number of empty nodes in the trees. This demonstrates the asymmetrical nature of the trees generated by this process. This occurs since the support of the original function sits inside the principal tile, but is not the entire tile.

In practice, only a finite number of scaling and wavelet coefficients are ever computed at a given instance. Therefore, given these coefficient indices of the original function, one can calculate in advance, the base expansions of the relevant integers in the subscripts. Hence, the Mallat algorithm can be performed quickly, without the need to calculate the expansions while performing the algorithm. This generates a precise and *easily implementable* tree structure for the coefficients.

2.3 Number Systems

As was shown in the previous section, MRA related to valid bases in $\mathbb{C}$ are well behaved when it comes to performing the decomposition algorithm. We find that this theory applies in higher dimensions as well. We first extend the definition of a valid base to $\mathbb{R}^n$.

Definition 2.3.1 A *valid base* for $\mathbb{Z}^n$ is a pair (B, D) where $B \in M_n(\mathbb{Z})$ and $D \subset \mathbb{Z}^n$, such that $0 \in D$ and every element $z \in \mathbb{Z}^n$ can be represented uniquely as a sum of powers of B, with coefficients in D.[4] The set D is called the *digit set*. More precisely, each $z \in \mathbb{Z}^n$ can be written uniquely as

$$z = \sum_{j=0}^{t_z - 1} B^j d_j, \text{ where } d_j \in D \text{ and } t_z \in \mathbb{N}^+.$$

If z has this form, write $z = (d_{t_z - 1} \cdots d_1 d_0)_B$.

[4]The set $M_n(\mathbb{Z})$ denotes the set of all $n \times n$ matrices over $\mathbb{Z}$.

The analogous terminology to complex bases will be used. Therefore, the radix expansion of a $z \in \mathbb{R}^n$ is an expression of the form

$$z = \sum_{j=-\infty}^{t_z-1} B^j d_j, \text{ where } d_j \in D \text{ and } t_z \in \mathbb{N}^+.$$

The results of complex bases translate to higher dimensions.

Proposition 2.3.2 If (B, D) is a valid base, then D is a complete residue system for $\mathbb{Z}^n/B\mathbb{Z}^n$ and hence contains $|\det(B)|$ elements.

Proof Suppose $z = \sum_{j=0}^{t} B^j d_j$, $d_j \in D$. Then $z \equiv d_0 \mod B$. Hence D contains a complete residue system for $\mathbb{Z}^n/B\mathbb{Z}^n$. Now, suppose $c, d \in D$ are distinct and $c \equiv d \mod B$. Then let $e = \sum_{j=0}^{t} B^j d_j$, $d_j \in D$ such that $c - d = Be$ for some $e \in \mathbb{Z}^n$. Hence, $(c)_B$ and $(d_t \dots d_0 d)_B$ are two distinct addresses of c, which is a contradiction to the assumption that (B, D) is a valid base. ■

Proposition 2.3.3 If $B \in M_n(\mathbb{Z})$ and D is a complete residue system for $\mathbb{Z}^n/B\mathbb{Z}^n$, that contains 0, then the following are equivalent:

i) The matrix B is a valid base using the digit set D;

ii) The matrix B is invertible and for every $z \in \mathbb{Z}^n$ there exists a positive integer t such that $\Phi^t(z) = 0$, where the function $\Phi : \mathbb{Z}^n \to \mathbb{Z}^n$ is defined by $\Phi(z) = B^{-1}(z - d)$, $d \in D$ and $d \equiv z \mod B$.

Proof i) $\implies$ ii) Suppose B is not invertible. Then, choose $z \in \mathbb{Z}^n \backslash \{0\}$ such that $Bz = 0$. Since B is a valid base, let $z = \sum_{k=0}^{t} B^k d_k$, $d_k \in D$, $d_t \neq 0$. Then 0 has two representations in base (B, D), that is (0) and $(d_t \dots d_0 0)$, a contradiction. Hence B is invertible.

For the second part of the implication, let $z \in \mathbb{Z}^n$ and write z in its expansion in base

(B, D). Generalizing the base conversion algorithm to $\mathbb{Z}^n$, we have

$$\begin{aligned} z &= Bz_1 + d_0 \\ z_1 &= Bz_2 + d_1 \\ &\vdots \\ z_t &= B0 + d_t. \end{aligned}$$

Since we assume that (B, D) is a valid base, then

$$\begin{aligned} \Phi(z) &= z_1 \\ \Phi^2(z) &= z_2 \\ &\vdots \\ \Phi^t(z) &= 0. \end{aligned}$$

ii) $\implies$ i) Let $z \in \mathbb{Z}^n$ and consider the sequence of integers $z_0 = z$ and $z_{k+1} = \Phi^{k+1}(z) = B^{-1}(z_k - d_k)$ for $k = 0, \ldots, t$ with $\Phi^t(z) = 0$ such that $z_{t-1} \neq 0$. Then $z = \sum_{k=0}^{t} B^k d_k$. Now, suppose that $z = \sum_{k=0}^{s} B^k e_k$, $e_k \in D$. Then, $d_k \equiv e_k \mod B$ for each k. Since D is a complete residue system of $\mathbb{Z}^n/B\mathbb{Z}^n$, we must have $d_k = e_k$ for each k. Hence the representation of z is unique. ■

Definition 2.3.4 Let $w = \{w_k\}$ be an IFS. We say that z is a *cyclic point* of the IFS if there exists a finite sequence of indices $k_1, k_2, \ldots, k_n \in \mathbb{Z}$ such that $z = w_{k_1} \circ w_{k_2} \circ \cdots \circ w_{k_n}(z)$.

Consider now the generalization of Theorem 1.3.15 when B is an acceptable dilation:

Theorem 2.3.5 If $B \in M_n(\mathbb{Z})$ is an acceptable dilation and D is a complete residue system for $\mathbb{Z}^n/B\mathbb{Z}^n$, that contains 0, then the following are equivalent:

i) The matrix B is a valid base using the digit set D;

ii) For every $z \in \mathbb{Z}^n$ there exists a positive integer t such that $\Phi^t(z) = 0$, where the function $\Phi : \mathbb{Z}^n \to \mathbb{Z}^n$ is defined by $\Phi(z) = B^{-1}(z - d)$, $d \in D$ and $d \equiv z \mod B$;

iii) There is no positive integer t for which

$$B^{t-1}d_{t-1} + \ldots + Bd_1 + d_0 \equiv 0 \mod (B^t - I) \text{ with } d_{t-1}, \ldots, d_0 \in D$$

and not all d_i are equal to zero;

iv) The IFS w has no nonzero cyclic points in $\mathbb{Z}^n$, where $w = \{w_k\}$ with $w_k(z) = B^{-1}(z - d_k)$ and $D = \{d_k : k = 1, \ldots, |\det(B)|\}$.

Proof i) $\iff$ ii) Proved by Proposition 2.3.3 since an acceptable dilation is invertible. iii) $\implies$ iv) Suppose $z \in \mathbb{Z}^n \backslash \{0\}$ is a cyclic point of the IFS w. Then, let $\sigma = (\sigma_0 \ldots \sigma_{t-1})$, $t > 0$ such that $z = w_{\sigma_{t-1}} \circ \ldots \circ w_{\sigma_0}(z)$. Define the sequence $z_k = B^{-1}(z_{k-1} - d_{\sigma_{k-1}})$, where $z_0 = z$. Then

$$\begin{aligned} z = z_0 &= Bz_1 + d_{\sigma_0} \\ &= B(Bz_2 + d_{\sigma_1}) + d_{\sigma_0} = B^2 z_2 + Bd_{\sigma_1} + d_{\sigma_0} \\ &\vdots \\ &= B^t z_t + \sum_{k=0}^{t-1} B^k d_{\sigma_k} \\ &= B^t z + \sum_{k=0}^{t-1} B^k d_{\sigma_k}. \end{aligned}$$

Therefore, $\sum_{k=0}^{t-1} B^k d_{\sigma_k} \equiv 0 \mod (B^t - I)$, a contradiction to the hypothesis. Hence, no nonzero integer cyclic point of the IFS w exists.

iv) $\implies$ iii) Suppose $\sum_{k=0}^{t-1} B^k d_{\sigma_k} \equiv 0 \mod (B^t - I)$ for some positive integer t. Then, let $z \in \mathbb{Z}^n$ such that

$$\sum_{k=0}^{t-1} B^k d_{\sigma_k} = B^t z - z.$$

Define the sequence (z_k) as above. Then

$$\begin{aligned} z_1 = B^{-1}(z_0 - d_{\sigma_0}) \qquad &= w_{\sigma_0}(z) \\ z_2 = B^{-1}(z_1 - d_{\sigma_1}) \qquad &= w_{\sigma_1}(z_1) = w_{\sigma_1} \circ w_{\sigma_0}(z) \\ \vdots & \\ z = z_t = B^{-1}(z_{t-1} - d_{\sigma_{t-1}}) &= w_{\sigma_{t-1}} \circ \ldots \circ w_{\sigma_0}(z). \end{aligned}$$

Therefore, z is a non-zero cyclic integer point of the IFS $w = \{w_k\}$, a contradiction to the hypothesis. Hence, no such t exists, and thus iii) is proved.

iv) $\implies$ ii) Let $z \in \mathbb{Z}^n \backslash \{0\}$. Consider the sequence $z_k = \Phi^k(z) = B^{-1}(z_{k-1} - d_{\sigma_{k-1}})$, $d_{\sigma_{k-1}} \in D$, where $z_0 = z$. Then,

$$\begin{aligned} z_1 &= B^{-1}z - B^{-1}d_{\sigma_0} \\ z_2 &= B^{-2}z - B^{-2}d_{\sigma_0} - B^{-1}d_{\sigma_1} \\ &\vdots \\ z_t &= B^{-t}z - \sum_{k=1}^{t} B^{-k}d_{\sigma_{t-k}}, \end{aligned}$$

hence

$$||z_t|| \leq \left|\left|B^{-t}z\right|\right| + \sum_{k=1}^{t} \left|\left|B^{-k}d_{\sigma_{t-k}}\right|\right|,$$

where $||\cdot||$ is the Euclidean norm. Since B is an acceptable dilation, using the Jordan normal form, let C, s, λ be positive real constants, $\lambda < 1$ and $s \in \mathbb{N}$, such that

$$\left|\left|B^{-j}x\right|\right| \leq C\lambda^j j^s \, ||x|| \quad \forall x \in \mathbb{R}^n.$$

Then

$$\begin{aligned} ||z_t|| &\leq C\,||z||\,\lambda^t t^s + Ca\sum_{k=1}^{t}\lambda^k k^s, \qquad \text{where } a = \max\{||d|| : d \in D\}. \\ &\leq C\,||z||\,\lambda^t t^s + Ca\sum_{k=1}^{\infty}\lambda^k k^s, \qquad \text{since } \lambda > 0. \end{aligned}$$

Consider the limit of the positive sequence $(\lambda^k k^s)$. Using l'Hôpital's rule we get the following result:

$$\begin{aligned} \lim_{k\to\infty}\lambda^k k^s &= \lim_{k\to\infty}\frac{k^s}{\lambda^{-k}} \\ &= \frac{-s}{\ln\lambda}\lim_{k\to\infty}\lambda^k k^{s-1} \\ &\;\;\vdots \\ &= \frac{\prod_{n=0}^{s-1}(s-n)}{(\ln\lambda)^{s-1}}\lim_{k\to\infty}\lambda^k \\ &= 0. \end{aligned}$$

Hence, let $N \in \mathbb{R}^+$ such that $C\lambda^k k^s\,||z|| \leq N$ for all $k \in \mathbb{N}$. Furthermore,

$$\lim_{k\to\infty}\frac{\lambda^{k+1}(k+1)^s}{\lambda^k k^s} = \lambda\lim_{k\to\infty}\left(1+\frac{1}{k}\right)^s = \lambda < 1.$$

Hence, by the ratio test, $\sum_{k=1}^{\infty}\lambda^k k^s = M$ for some $M \in \mathbb{R}^+$. Therefore

$$||z_t|| \leq N + CaM \quad \forall t \in \mathbb{N}^+.$$

This means the set of iterates of $\{\Phi^t(z)\}_{t=1}^{\infty}$ is bounded and so takes on only finitely many values. Therefore, if there is no positive integer r such that $\Phi^r(z) = 0$, then let $0 \leq s < t \in \mathbb{N}^+$ such that $0 \neq \Phi^s(z) = \Phi^t(z)$. Hence $\Phi^{t-s}(z) = z$, which means that z is a nonzero cyclic point of the IFS w, which is a contradiction to the hypothesis. Therefore, there is a positive integer r such that $\Phi^r(z) = 0$.

ii) $\implies$ iv) Suppose z is a nonzero cyclic point of the IFS w. Then let $\sigma = (\sigma_0 \ldots \sigma_{t-1})$,

$t > 0$ such that $z = w_{\sigma_{t-1}} \circ \ldots \circ w_{\sigma_0}(z)$. Then, none of the points $z_n = w_{\sigma_{n-1}} \circ \ldots \circ w_{\sigma_0}(z)$ can be zero. For, if a given $z_n = 0$, then $d_{\sigma_n} = 0$, since D is a complete residue system of $\mathbb{Z}^n/B\mathbb{Z}^n$. This implies that $z_k = 0\ \forall k \geq n$, thereby contradicting the hypothesis that $z \neq 0$. Therefore, $\Phi^t(z) \equiv z_t = z$ and there is no positive r such that $\Phi^r(z) = 0$, which is a contradiction to the hypothesis of ii). Hence, there is no nonzero integer cyclic point of the IFS. ■

Corollary 2.3.6 Theorem 2.3.5 holds for complex bases.

Proof A complex base is equivalent to an acceptable dilation matrix over $\mathbb{Z}^2$. ■

Corollary 2.3.7 $\pm 1, \pm i, 1 \pm i$ and 2 are never valid bases for $\mathbb{Z}[i]$.

Proof By Theorem 2.3.5, a base must have norm greater than 1, thus ± 1 and $\pm i$ cannot be valid bases. For $b = 2$ or $b = 1 \pm i$, let $0 \neq k \in D$, where D is a complete residue system of $\mathbb{Z}[i]$ modulo b, which contains zero. Then, let $w(z) = b^{-1}(z - k)$. If $b = 2$, then $z = -k$ is a fixed point of w and if $b = 1 \pm i$, then $z = \pm ik$ is a fixed point of w. ■

The concept of MRA associated to complex bases is easily generalized to obtain MRA associated to number systems. Consequently, the results demonstrated in the previous two sections can be generalized to $\mathbb{Z}^n$. This fact will be used in subsequent discussions and we therefore obtain the following result:

Corollary 2.3.8 If (B, D) is a valid base of $\mathbb{Z}^n$, and if $\{s_{0,j} : j \in J\}$, $J \subset \mathbb{Z}^n$ is a finite set of scaling coefficients, then there is a positive integer r such that $s_{-r,z} = 0$ for all $z \in \mathbb{Z}^n \backslash \{0\}$.

And from iv) $\implies$ ii) of the proof of the above theorem, we see that

Corollary 2.3.9 Let (A, D) be an MRA in $\mathbb{Z}^n$ and suppose that $\{s_{0,z} : z \in \mathbb{Z}^n\}$ is a set of scaling coefficients such that all but finitely many are zero. Then, there is a positive integer r such that $s_{-r,z} = 0$ for all $z \in \mathbb{Z}^n$ with z not a cyclic point of the IFS w.

Therefore, the situation of Example 2.1.7 cannot occur for a valid base. It is potentially feasible to still perform the Mallat algorithm in the case where cyclic points are generated (i.e. where the indices of the coefficients are cyclic). In these instances, one could maintain a list of the cyclic points. If a cyclic point is generated, the decomposition of that branch would stop. This could give an algorithm in practice for doing a terminating Mallat for all MRA. This has not been investigated and there may be additional issues to resolve.

The following example was contributed by William Gilbert.

Example 2.3.10 For an example of a representation in $\mathbb{Z}^3$, consider the matrix of determinant -2:

$$B = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

which is the companion matrix of the polynomial $m(x) = x^3 + x^2 + x + 2$. If s is a root of $m(x)$, so that $m(x)$ is the minimal polynomial of s, then $\mathbb{Q}(s)$ is a cubic number field and all elements can be represented as $a_0 + a_1 s + a_2 s^2$ with a_i in $\mathbb{Q}$. We can also represent this element as (a_0, a_1, a_2) in $\mathbb{Q}^3$ and multiplication by s corresponds to matrix multiplication by B. Now s is a valid base for the integers $\mathbb{Z}[s]$ using the binary digit set $\{0, 1\}$. This is pictured in Figure 5 of [32]. This base extends to a representation of all of $\mathbb{Q}(s)$ using digits to the right of the radix point.

Since $\mathbb{Z}[s]$ is in 1-1 correspondence with $\mathbb{Z}^3$, the matrix base B will give a unique representation of the elements of $\mathbb{Z}^3$ using the binary digit set $\{(0,0,0), (1,0,0)\}$.

It would be desirable to remove the assumption that B is an acceptable dilation from the statement of Theorem 2.3.5. Indeed, the equivalence of i) and ii) does not depend on this assumption at all, but simply the need for B to be invertible. It is a sufficient condition in iii) and iv) that B be an acceptable dilation to demonstrate iii) or iv) $\implies$ ii). However, it is not clear that this is a necessary condition. We will motivate the possibility that this

is the case ending this section with a conjecture.

The following is a result of functional analysis, which holds also for infinite dimensional spaces:

Theorem 2.3.11 If L is a diagonalizable linear operator on a finite dimensional vector space V, then there exists a norm $||\cdot||_L$ on V such that the operator norm of L with respect to this norm is equal to its spectral radius, $\rho(L)$.

With this result, one can demonstrate the following proposition, which is essentially a result of Gilbert [30]. Indeed, the key requirement of the matrix A is that there is a consistent matrix norm $||\cdot||$ such that $||A|| = \rho(A) \geq 1$. The proof is omitted here.

Proposition 2.3.12 Let $x_k = Ax_{k-1} + f(x_{k-1})$ be a real non-linear difference equation with $x_{k-1}, x_k \in \mathbb{R}^n$. If the matrix A is diagonalizable and has an eigenvalue of modulus greater than or equal to one, and $f(x)$ is bounded for each vector $x \in \mathbb{Z}^n$, then there is some initial vector $x_0 \in \mathbb{Z}^n$ for which the solution of the difference equation is bounded away from zero.

In the immediate circumstances, the difference equation is translated to the division algorithm equation:

$$z_{t+1} = B^{-1}(z_t - d_t) = B^{-1}z_t - B^{-1}d_t.$$

If B is diagonalizable, then so is B^{-1}. Hence, if such a B has an eigenvalue of modulus less than or equal to one in modulus, it can never be a valid base for any digit set. In other words, any diagonalizable matrix, that is a valid basis for some digit set, must be an acceptable dilation.

This result may lead one to conjecture that the assumptions of Theorem 2.3.5 might be weakened as follows:

Conjecture 2.3.13 If $B \in M_n(\mathbb{Z})$ and D is a complete residue system for $\mathbb{Z}^n/B\mathbb{Z}^n$, that contains 0, then the following are equivalent:

i) The matrix B is a valid base using the digit set D;

ii) The matrix B is invertible and for every $z \in \mathbb{Z}^n$ there exists a positive integer t such that $\Phi^t(z) = 0$, where the function $\Phi : \mathbb{Z}^n \to \mathbb{Z}^n$ is defined by $\Phi(z) = B^{-1}(z - d)$, $d \in D$ and $d \equiv z \mod B$;

iii) The matrix B is an acceptable dilation and there is no positive integer t for which

$$B^{t-1}d_{t-1} + \ldots + Bd_1 + d_0 \equiv 0 \mod (B^t - I) \text{ with } d_{t-1}, \ldots, d_0 \in D$$

and not all d_i are equal to zero;

iv) The matrix B is an acceptable dilation and the IFS w has no nonzero cyclic points in $\mathbb{Z}^n$, where $w = \{w_k\}$ with $w_k(z) = B^{-1}(z - d_k)$ and $D = \{d_k : k = 1, \ldots, |\det(B)|\}$.

The obstacle is proving that i) or ii) $\implies$ iii) or iv). The difficulty is showing that if B is not diagonalizable, and is a valid base for some digit set D, then it must be an acceptable dilation. There are results of operator theory which provide the existence of a norm arbitrarily close to the spectral radius of such a matrix: $||B|| \leq \rho(B) + \epsilon$. The problem in proving an equivalent result of Proposition 2.3.12 is the need to carry the ϵ in the equations, which prevents the same inequalities from being constructed. Indeed, this ϵ might be used to construct a counter example.

One final note on number systems is that it may be possible to extend the above results to cases where one has a unique integer fixed point. In such a case, one could write each integer as $z = \bar{z} + \sum_{k=0}^{t} B^k a_k$. In this case, as may be possible for MRA with cyclic integer points, the decomposition algorithm could be halted at $\bar{z}$. This has not been investigated.

2.4 Extensions

We now return to address the situation that arose when decomposing a function whose support was not an expanded copy of the principal (fundamental) tile of the MRA (see Example 2.2.9). In this case (which is essentially every case other than the separable wavelet case), the trees have many "missing" nodes. When using these MRA in applications, one generally obtains artefacts in the image.[5] For example, in the case of pruning or thresholding, the edges of the compressed image are often black. In the case of LIFSW, one may find such artefacts in the interior of the image.

The reason these artefacts are generated is because we assume that outside the boundaries of the image, the function is identically zero. That is, the square image *is* the support of the function. However, the image actually sits inside an expanded copy of the principal tile of the MRA. It is from the points outside the image, those that reside in this expanded tile, that we have holes in the trees.

Given these unnatural-seeming artefacts, we are motivated to extend the function to an expanded copy of the principal tile in a way that is consistent with the Mallat algorithm. To begin, we will need to define some notation.

Let (B, D) be a number system for $\mathbb{Z}^n$. Consider the MRA associated with (B, D) with scaling function $\phi = \chi_Q$ where $Q = T_{(B,D)}$ is the principal tile of (B, D). Let $X \subset \mathbb{Z}^n$ be finite. The MRA and X will be fixed for the remainder of this discussion.

Definition 2.4.1 Let $t \in \mathbb{N}$. Define A_t to be the set of all addresses of length t. That is $A_t = \{(d_{t-1} \dots d_0) : d_k \in D\}$, and define A_0 to be the set containing the address of length zero. Then, let X_t be the set of all points in $\mathbb{Z}^n$ with addresses of length t. Formally, $X_t = \{z \in \mathbb{Z}^n : a_z^t \in A_t\}$.

[5] Chapter 3 discusses applications to image analysis.

Definition 2.4.2 Let $L = L_{X,(B,D)}$. Define $\tilde{A} = A_L$ and $\tilde{X} = X_L$.

For the remainder of this section, let $L = L_{X,(B,D)}$ and $A = A_X$.

Definition 2.4.3 Given $z \in \mathbb{Z}^n$, define the *z-translate* of Q as $Q_z = Q + z = \{q + z : q \in Q\}$. Then define the *tile of* X as $Q_X = \cup_{x \in X} Q_x$. Finally, define the *supertile* of X (or Q_X) as $\tilde{Q} = Q_{\tilde{X}}$.

Let us summarize the notation given above. First, L is the smallest natural number such that every $x \in X$ can be represented as $\sum_{k=0}^{L-1} B^k d_k$. The set A is the set of all addresses of length L of the points in X expressed in base (B, D). The set $\tilde{A}$ is the set of all addresses of length L in base (B, D). The set $\tilde{X}$ is all the points of $\mathbb{Z}^n$ that can be written as $\sum_{k=0}^{L-1} B^k d_k$ and therefore... $Q_{\tilde{X}}$ is the smallest expanded copy of the principal tile that completely contains the points of X!

Example 2.4.4 Let $X = \{0, 1, i, 1 + i\}$ with base $(-1 + i, \{0, 1\})$. From Example 2.2.9, we have $L = 4$ and hence the set A_L is the set of all sequences of length four with digits 0 and 1. The set $\tilde{X} = X_L$ is the set of integers represented by the addresses of A_L. These two sets are illustrated in Figure 2.4.

The set $Q_{\tilde{X}}$ is simply the union of the translates over $\tilde{X}$ of the twin dragon tile containing zero. Envisioning $Q_{\tilde{X}}$ is simple if one imagines replacing each square by the twin dragon tile Q_0.

We now move to the discussion of the situation in question, that is the extension of functions. Let $f = \sum_{x \in X} s_{0,x} \phi_{0,x}$, where $Q_X \neq Q_{\tilde{X}}$. For instance, consider Example 2.2.9. In this case then, the support of f, which is by definition Q_X, is contained in, but not equal to $Q_{\tilde{X}}$. Step 1 in Example 2.4.13 illustrates this situation (see Figure 2.5 on page 110). This is represented in the drawing of the scaling and wavelet trees as empty nodes, the nodes coming from points outside the support of f. In this sense, we can think of the tree

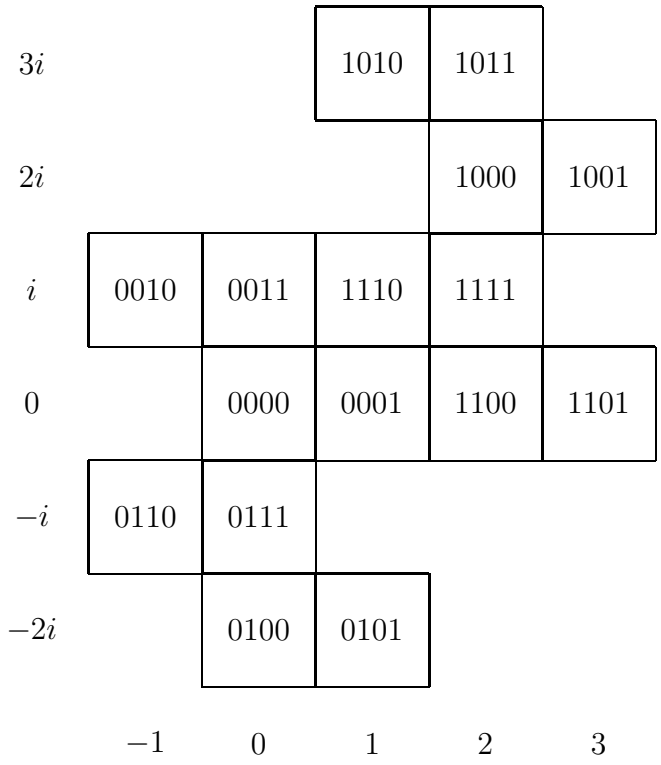


Figure 2.4: Addresses and points of a set for the twin dragon.

as not being full. As we move up the tree, these external scaling coefficients interact with the ones from within Q_X and influence the corresponding scaling and wavelet coefficients at lower levels in the tree. This generally causes points on the boundary of the support to become edges. When performing an IFSW, if a domain is taken such that a part of it is outside Q_X, this will introduce uniform black artefacts into the image, which will be propagated by the IFSW into the lower detail levels.

It should be noted that in the case where digits are elements in $\mathbb{Z}^n$, trees are generated by simply indexing the digits in some (arbitrary) way and then creating the branches based on the chosen order. The only important thing is maintaining consistency of the chosen

ordering between levels.

If $X = \tilde{X}$, then the tree will be full. What we mean by full here is that every scaling coefficient on level zero is derived from a point x in the support of f. Therefore, every scaling coefficient above level zero has all of its children. Thus, when performing wavelet or IFSW compression, or any other manipulation on the trees, no external (non-image related) information is introduced, the trees are self-contained and exactly define the original function.

To resolve the problem when $Q_X \neq Q_{\tilde{X}}$ then, we would need to extend the image in a "natural way", that is one that utilizes the tiling itself to extend the image and preserves the properties of f without introducing edges or other anomalies into the wavelet trees. We will propose such an algorithm after a bit more notation.

We will consider some definitions of addresses and trees:

Definition 2.4.5 If $\rho = (d_{t-1} \cdots d_0) \in D^t$ and $\sigma = (d_{t-1} \cdots d_0 d) \in D^{t+1}$ we say that σ is a *child* of ρ and that ρ is the *parent* of σ. If $\lambda = (d_{t-1} \cdots d_0 d') \in D^{t+1}$, $d' \neq d$, we say that λ is a *sibling* of σ.

Example 2.4.6 In base $(2+i, \{0, 1, i, -i, -2-3i\})$, the address $(-2-3i \cdot 0 \cdot i)$ is the child of $(-2-3i \cdot 0)$ and its siblings are $(-2-3i \cdot 0 \cdot 0), (-2-3i \cdot 0 \cdot 1), (-2-3i \cdot 0 \cdot -i)$ and $(-2-3i \cdot 0 \cdot -2-3i)$.

It is clear from the definition that an address can have at most one parent and has as many children as the number of elements of D. In the case of a valid base, $|D| = |\det B|$ and all nodes have a parent except the address of length zero. This definition allows us to clearly define the meanings of parent, child and sibling for the scaling and wavelet trees:

Definition 2.4.7 Let $0 < l \leq L$. The coefficient $s_{-l,(d_{L-1} \cdots d_l)}$ is the *parent* of $s_{-l+1,(d_{L-1} \cdots d_l d)}$ $\forall d \in D$ and these scaling coefficients are its *children*. For $d \in D$, the

siblings of $s_{-l+1,(d_{L-1}\cdots d_l d)}$ are $s_{-l+1,(d_{L-1}\cdots d_l d')}$, $d \neq d' \in D$. Analogous terminology is assigned to the wavelet coefficients.

Using this definition, we define the scaling tree S_f of f formally as the graph with nodes $\{s_{-l,\sigma} : l \in \{0,\ldots,L\}, \sigma \in D^{L-l}\}$ and edges joining parents and their children. Define the wavelet trees $W_f^k, k = 1,\ldots,q-1$, with $q = |\det B|$, in an equivalent manner.

Notation 2.4.8 Let T be a tree, t a node of T and T' be a set of nodes of T. Then define $\mathrm{sib}(t,T')$ to be the set of all siblings of t in T' and define $\mathrm{child}(t,T')$ to be the set of all children of t in T'.

In our algorithm, we will want to extend f from X to a function $\tilde{f}$ on $\tilde{X}$ such that if $f = \sum_{x\in X} s_x\phi_x$, and $\tilde{f} = \sum_{z\in\tilde{X}} \tilde{s}_z\phi_z$, then $\tilde{s}_z = s_z$ for all $z \in X$.

We are going to use the Mallat algorithm as a foundation for the extension. What this means is that we will want to construct a scaling tree and wavelet trees $\tilde{S}$ and $\tilde{W}^k$, $k = 1,\ldots,q-1$ for $\tilde{f}$ using the scaling and wavelet trees S and W^k, $k = 1,\ldots,q-1$ of f.

The only difference between them at first is that the scaling coefficients of f are s_x, for $x \in X$ and $s_z = 0$, for $z \in \tilde{X}\backslash X$. We will denote the scaling and wavelet coefficients of $\tilde{f}$ by $\tilde{s}_{-l,\sigma}$ and $\tilde{w}^k_{-l,\sigma}$. As we build the trees for $\tilde{f}$, at each level we will have "assigned" and "unassigned" coefficients. By this we mean that the assigned coefficients are those to which the algorithm has explicitly given a value. The unassigned ones are the ones which have not been given a value at a given step in the algorithm. We will therefore define sets of coefficients that will keep track of these distinguished coefficients for us through the algorithm.

Algorithm 2.4.9 (Extension Algorithm) i) Assumptions: Let $X \subset \mathbb{Z}^n$ be finite, $L = L_{X,(B,D)}$ and $X \subsetneq \tilde{X}$. Suppose $f = \sum_{x\in X} s_{0,a_x^L}\phi_{0,x}$. Formally, define $\tilde{f} = \sum_{z\in\tilde{X}} \tilde{s}_{0,a_z^L}\phi_{0,z}$. Define for each $0 \leq l \leq L$, $\tilde{S}_{-l} = \{\tilde{s}_{-l,\sigma} : \sigma \in D^{L-l}\}$;

ii) Setup: Set for each $x \in X$, $\tilde{s}_{0,a_x^L} = s_{0,a_x^L}$. Then set $\tilde{S}_0^A = \{\tilde{s}_{0,a_x^L} : x \in X\}$ and $\tilde{S}_0^U = \tilde{S}_0 \backslash \tilde{S}_0^A$;

iii) Iteration: For each $l = 0, \ldots, D^{L-1}$:

(a) For each $\sigma \in D^{L-l}$, if $\tilde{s}_{-l,\sigma} \in \tilde{S}_{-l}^U$ and if $\text{sib}(\tilde{s}_{-l,\sigma}, \tilde{S}_{-l}^A) = \{\tilde{s}_{-l,\sigma_1}, \ldots, \tilde{s}_{-l,\sigma_k}\} \neq \emptyset$, then for each $n = 0, \ldots, l$ and for each $\lambda \in D^n$, set $\tilde{s}_{-l+n,\sigma\lambda} = \frac{1}{k} \sum_{m=1}^{k} s_{-l+n,\sigma_m\lambda}$;

(b) Set $\tilde{S}_{-l-1}^A = \emptyset$ and $\tilde{S}_{-l-1}^U = \tilde{S}_{-l-1}$;

(c) For each $\sigma \in D^{L-l-1}$, if $\text{child}(\tilde{s}_{-l-1,\sigma}, \tilde{S}_{-l}^A) \neq \emptyset$, set $\tilde{s}_{-l-1,\sigma} = \sum_{d \in D} h_d \tilde{s}_{-l,\sigma d}$. Set $\tilde{S}_{-l-1}^A = \tilde{S}_{-l-1}^A \cup \{\tilde{s}_{-l-1,\sigma}\}$ and $\tilde{S}_{-l-1}^U = \tilde{S}_{-l-1}^U \backslash \{\tilde{s}_{-l-1,\sigma}\}$.

Note that the algorithm neglects the construction of the wavelet coefficients. One can restate the algorithm to construct the wavelet coefficients as the algorithm progresses, using the equivalent assignment formulae as for the scaling coefficients. One might worry however that such an approach would cause an inconsistency with the Mallat algorithm's definition of the wavelet coefficients. However, this is not the case and will be noted following the next results.

Proposition 2.4.10 Algorithm 2.4.9 is well-defined.

Proof Showing the algorithm is well-defined means ensuring that no already assigned coefficient is reassigned. This is clear in the definition of the algorithm since the only time a parent is assigned is through the Mallat algorithm (Step iii(c)), which is well-defined and the only time a child is assigned is when it is a grandchild of an originally unassigned coefficient (Step iii(a)), and hence was undefined beforehand. If this last statement is not intuitively clear for the reader, it can be shown through an induction on n by assuming that if some grandchild is already assigned that the original coefficient must have already been as well. ■

Theorem 2.4.11 Algorithm 2.4.9 is consistent with the Mallat algorithm.

Proof To prove this, we must show that at each level $1 \leq l \leq L$, if $\tilde{s}_{-l,\gamma} \in \tilde{S}_{-l}$ then $\tilde{s}_{-l,\gamma} = \sum_{d\in D} h_d \tilde{s}_{-l+1,\gamma d}$. Note that if a coefficient was assigned in Step iii(c), that is if $\tilde{s}_{-l,\gamma} \in \tilde{S}^A_{-l}$, then there is nothing to show since it was assigned using precisely the Mallat algorithm. Also, $\tilde{s}_{-L,\emptyset}$ is constructed in Step iii(c). Therefore, we only need to prove the case where coefficients are assigned in Step iii(a), that is for coefficients in $\tilde{S}^U_{-l}$. We will prove the result by induction on $l + n$.

Before we begin, notice the following. Suppose that $\tilde{s}_{-l,\sigma} \in \tilde{S}^U_{-l}$. Then it must hold that for each $1 \leq n \leq l$ and each $\lambda \in D^n$ that $\tilde{s}_{-l+n,\sigma\lambda} \in \tilde{S}^U_{-l+n}$. If not, then for some n and λ, $\tilde{s}_{-l+n,\sigma\lambda} \in \tilde{S}^A_{-l+n}$. Thus, one can construct its siblings, and hence its parent is in $\tilde{S}^A_{-l+n-1}$, that is its parent can be assigned in Step iii(c). By induction then, $\tilde{s}_{-l,\sigma} = \tilde{s}_{-l+(n-n),\sigma\emptyset} \in \tilde{S}^A_{-l}$. This is a contradiction to the hypothesis on $\tilde{s}_{-l,\sigma}$. So, let $1 \leq l < L$, $\sigma \in D^{L-l}$, such that $\tilde{s}_{-l,\sigma} \in \tilde{S}^U_{-l}$, with $\text{sib}(\tilde{s}_{-l,\sigma}, \tilde{S}^A_{-l}) = \{\tilde{s}_{-l,\sigma_m}\}_{m=1}^k \neq \emptyset$:

i) If $l = 1$ and $n = 0$, then by definition, each $\tilde{s}_{-l,\sigma_m}$ is in $\tilde{S}^A_{-l}$. Hence

$$\begin{aligned}
\tilde{s}_{-l,\sigma} &= \frac{1}{k}\sum_{m=1}^{k} \tilde{s}_{-l,\sigma_m} \\
&= \frac{1}{k}\sum_{m=1}^{k}\sum_{d\in D} h_d \tilde{s}_{-l+1,\sigma_m d} \quad (\text{since } \tilde{s}_{-l,\sigma_m} \in \tilde{S}^A_{-l}) \\
&= \sum_{d\in D} h_d \frac{1}{k}\sum_{m=1}^{k} \tilde{s}_{-l+1,\sigma_m d} \\
&= \sum_{d\in D} h_d \tilde{s}_{-l+1,\sigma d} \qquad (\text{since } \tilde{s}_{-l,\sigma} \in \tilde{S}^U_{-l});
\end{aligned}$$

ii) Now, suppose that for $1 \leq l < q < L$ and $0 \leq n < p \leq l-1$ that $\tilde{s}_{-l+n,\sigma\lambda} = \sum_{d\in D} h_d \tilde{s}_{-l+n+1,\sigma\lambda d}$, for all $\lambda \in D^n$. Let $l = q$, $n = p$ and $\lambda \in D^n$. Consider a given $\tilde{s}_{-l+n,\sigma_m\lambda}$. Since $\tilde{s}_{-l,\sigma_m} \in \tilde{S}^A_{-l}$, two situations can occur. Either $\tilde{s}_{-l+n,\sigma_m\lambda} \in \tilde{S}^A_{-l+n}$ or $\tilde{s}_{-l+n,\sigma_m\lambda} \in \tilde{S}^U_{-l+n}$. In the second instance, there must exist some integer $1 \leq r \leq n$ such that $\tilde{s}_{-l+r,\sigma_m\rho} \in \tilde{S}^U_{-l+r}$ and $\tilde{s}_{-l+r-1,\sigma_m\theta} \in \tilde{S}^A_{-l+r-1}$ with $\rho \in D^r$, $\theta \in D^{r-1}$, $\rho = \theta d$ for some $d \in D$ and $\lambda = \rho\omega$ for some $\omega \in D^{n-r}$.

Suppose not. Then each $\tilde{s}_{-l+r,\sigma_m\rho} \in \tilde{S}^A_{-l+r}$ for each $1 \leq r \leq n$. By the note made

above, this contradicts the assumption that $\tilde{s}_{-l+n,\sigma_m\lambda} \in \tilde{S}^U_{-l+n}$. Therefore, let $-t = -l + r$ and $k = n - r$. Then

$$\begin{aligned} t + k &= (l - r) + (n - r) \\ &< q + p. \end{aligned}$$

Thus, by the induction hypothesis,

$$\tilde{s}_{-t+k,(\sigma_m\rho)\omega} = \sum_{d\in D} h_d \tilde{s}_{-t+k+1,(\sigma_m\rho)\omega d},$$

that is,

$$\tilde{s}_{-l+n,\sigma_m\lambda} = \sum_{d\in D} h_d \tilde{s}_{-l+n+1,\sigma_m\lambda d}.$$

Therefore,

$$\begin{aligned} \tilde{s}_{-l+n,\sigma\lambda} &= \frac{1}{k}\sum_{m=1}^{k} \tilde{s}_{-l+n,\sigma_m\lambda} && (\text{since } \tilde{s}_{-l,\sigma} \in \tilde{S}^U_{-l}) \\ &= \frac{1}{k}\sum_{m=1}^{k}\sum_{d\in D} h_d \tilde{s}_{-l+n+1,\sigma_m\lambda d} && (\text{by the above}) \\ &= \sum_{d\in D} h_d \frac{1}{k}\sum_{m=1}^{k} \tilde{s}_{-l+n+1,\sigma_m\lambda d} \\ &= \sum_{d\in D} h_d \tilde{s}_{-l+n+1,\sigma\lambda d} && (\text{since } \tilde{s}_{-l,\sigma} \in \tilde{S}^U_{-l}). \quad \blacksquare \end{aligned}$$

Given the above result, we can demonstrate the consistency of defining of the wavelet coefficients directly within the algorithm. The following result shows that using the algorithm as it is currently stated and then constructing the wavelet coefficients of $\tilde{f}$ using the decomposition algorithm yields the same coefficients as using the modified algorithm where the wavelet coefficients would be assigned directly in the same way as the scaling

coefficients:

Corollary 2.4.12 For each $\sigma \in D^{L-l}$, if $\tilde{s}_{-l,\sigma} \in \tilde{S}^U_{-l}$ and if $\text{sib}(\tilde{s}_{-l,\sigma}, \tilde{S}^A_{-l}) = \{\tilde{s}_{-l,\sigma_1}, \ldots, \tilde{s}_{-l,\sigma_k}\} \neq \emptyset$, then for each $n = 0, \ldots, l-1$, $\lambda \in D^n$, and $p = 1, \ldots, |\det B| - 1$, we have $\tilde{w}^p_{-l+n,\sigma\lambda} = \frac{1}{k}\sum_{m=1}^{k} \tilde{w}^p_{-l+n,\sigma_m\lambda}$.

Proof Let $0 < l < L$. Choose $\sigma \in D^{L-l}$, such that $\tilde{s}_{-l,\sigma} \in \tilde{S}^U_{-l}$ and $\text{sib}(\tilde{s}_{-l,\sigma}, \tilde{S}^A_{-l}) = \{\tilde{s}_{-l,\sigma_1}, \ldots, \tilde{s}_{-l,\sigma_k}\} \neq \emptyset$. Let $0 \leq n < l-1$, $\lambda \in D^n$, and $0 \leq p < |\det B|$. Then, we have

$$\begin{aligned}
\tilde{w}^p_{-l+n,\sigma\lambda} &= \sum_{d\in D} g^p_d \tilde{s}_{-l+n+1,\sigma\lambda d} && \text{(by the decomposition algorithm)} \\
&= \sum_{d\in D} g^p_d \frac{1}{k} \sum_{m=1}^{k} \tilde{s}_{-l+n+1,\sigma_m\lambda d} && \text{(by Theorem 2.4.11)} \\
&= \frac{1}{k} \sum_{m=1}^{k} \sum_{d\in D} g^p_d \tilde{s}_{-l+n+1,\sigma_m\lambda d} \\
&= \frac{1}{k} \sum_{m=1}^{k} \tilde{w}^p_{-l+n,\sigma_m\lambda} && \text{(by the decomposition algorithm).} \quad \blacksquare
\end{aligned}$$

By reversing the argument and considering the reconstruction rather than the decomposition algorithm, one can show that by defining the algorithm with the direct assignment of "missing" wavelet coefficients, the function $\tilde{f}$ is the same as that of the algorithm as originally stated. The proof is omitted for the sake of brevity; it is again a consequence of the orthogonality of the wavelet and scaling functions and spaces. This result gives flexibility to the implementor of the algorithm and removes the need of performing the wavelet decomposition only at the end of the construction of $\tilde{f}$. One can simply pass through the extension of the function f and at the same time create the wavelet tree.

One other interesting note is that the use of averaging in the algorithm is a somewhat arbitrary choice. Indeed, when reviewing the proof of the consistency of the algorithm(s) with the Mallat algorithm, one sees that any function M satisfying the following property

will allow a consistent extension:

$$M\left(\sum_d \alpha_d s_{d,1}, \ldots, \sum_d \alpha_d s_{d,k}\right) = \sum_d \alpha_d M(s_{d,1}, \ldots, s_{d,k}). \tag{2.7}$$

Furthermore, given the orthogonality of the wavelets and scaling functions at a given resolution, one is free to modify this map at each Step iii(a). It is possible that adapting the algorithm to fit the given function may be of interest.

Example 2.4.13 Consider the extension algorithm applied to the function f in Example 2.2.9 where M is the averaging operator used in the definition of the algorithm.

We will demonstrate the extension of f to $\tilde{f}$ by looking at the sets Q_σ as an equivalent representation of the levels of the scaling tree. At each step in Figure 2.5 on the following page, the numbers in the boxes represent the values of the scaling coefficients at the given address corresponding to the respective set Q_σ. In this representation, we normalize the filter coefficients to better illustrate how the coefficients are combining. Figure 2.6 on page 111 presents the complete scaling tree with the actual scaling coefficient values.

2.5 LIFSW and Number Systems

In this section we present the generalization of LIFSW to number systems on $\mathbb{R}^n$. As was discussed in Section 2.4, if $X \neq \tilde{X}$, we are in a sense "missing" coefficients in the scaling, and hence wavelet, tree of any function f with support in X. Given the discussion at the beginning of Section 2.4, it may be of interest to consider a definition of an LIFSW that acts solely on X and ignores coefficients in $\tilde{X}\backslash X$. In this way, we can always define a standard LIFSW acting on $\tilde{f}$ generated by the extension function $M = 0$.

In order to generalize LIFSW, we will want to consider it to act solely on *full subtrees* of X. Intuitively, a full subtree is one arising from coefficients on level zero that are all in X, i.e. no grandchild came from a point outside of X. We now discuss this formally.

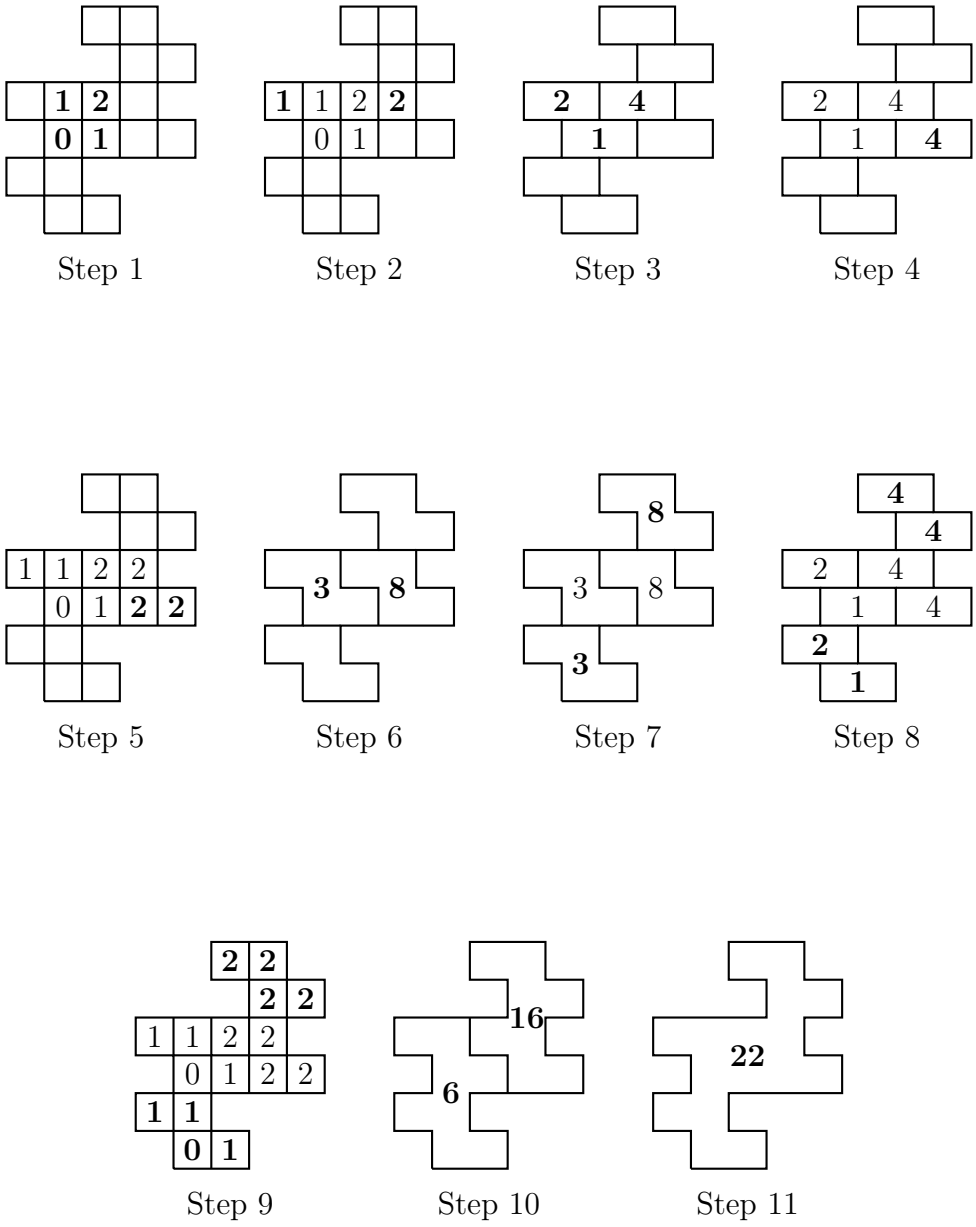


Figure 2.5: Illustration of extending a function.

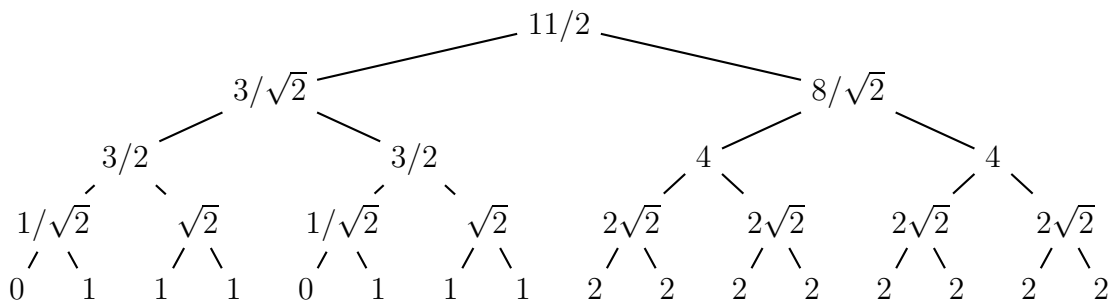


Figure 2.6: Scaling tree of an extended function.

Notation 2.5.1 Define the following sets:

$$Q_\sigma = \begin{cases} Q_z & \sigma = a_z^L \text{ for some } z \in \tilde{X}, \\ \bigcup_{d\in D} Q_{\sigma d} & \sigma \in D^{L-t}, t = 1, \ldots, L. \end{cases}$$

If $\sigma \in D^{L-t}$, we call Q_σ a tile of size t.[6] Now, define the following collections of sets:

$\mathcal{U}_t$ is the collection of the sets Q_σ where $\sigma \in D^{L-t}, t = 0, \ldots, L$;
$\mathcal{V}_t$ is the collection of the sets Q where $Q \in \mathcal{U}_t$ and $Q \cap (\tilde{X} \backslash X) = \emptyset, t = 0, \ldots, L$; and
$\mathcal{W}_t$ is the collection of the sets $Q_{\sigma d} \in \mathcal{V}_t, d \in D$ such that $Q_\sigma \notin \mathcal{V}_{t+1}, t = 0, \ldots, L-1$.

Let us discuss the meaning of these sets and collections. The sets $Q_{a_z^L}$ are simply the $\mathbb{Z}^n$ translates of Q over the set $Q_{\tilde{X}}$. A set Q_σ for $\sigma \in D^{L-t}$, $t = L-1, \ldots, 0$ is simply the union of its $|D| = q = |\det B|$ subtiles, and so $Q_\emptyset = Q_{\tilde{X}}$.

For each $t = 0, \ldots, L$, $\mathcal{U}_t$ is simply the collection of all tiles of size t, and since such tiles are disjoint, its elements form a partition of $\tilde{X}$. The collection $\mathcal{V}_t$ is the collection of all tiles of size t that are completely contained inside Q_X and the collection $\mathcal{W}_t$ ($t < L$) is the collection of all tiles of size t that expand to a tile of size $t+1$ which intersects the

[6] A slightly better terminology might be *height* t, since Q_σ has n-dimensional Lebesgue measure q^t.

exterior of Q_X. It is in this sense that we define the meaning of a *maximal tile* and a *full subtree* in X:

Definition 2.5.2 A *maximal tile with respect to* X is an element of some $\mathcal{W}_t$, $t = 0, \ldots, L-1$. If $f = \sum_{x \in X} s_x \phi_x$, then a *full subtree* T *of the scaling tree of* f *with respect to* X is a subtree of the scaling tree with root $s_{-l,\sigma}$ such that $Q_\sigma \in \mathcal{V}_l$. The subtree T is called a *maximal full subtree* if $Q_\sigma \in \mathcal{W}_l$.

The collections defined build a hierarchy of subtiles of Q_X and in the above sense, the collections $\mathcal{W}_t$ provide a partitioning of Q_X in a maximal way:

$$Q_X = \bigcup_{t=0}^{L-1} \bigcup_{Q_\sigma \in \mathcal{W}_t} Q_\sigma$$

Since we are assuming that $X \neq \tilde{X}$, then $\mathcal{V}_0 \neq \emptyset$ and $\mathcal{V}_L = \emptyset$.[7] Therefore, there is a greatest integer $0 \leq t_X < L$ such that $\mathcal{V}_{t_X} \neq \emptyset$ and $\mathcal{V}_t = \emptyset$, for all $t > t_X$. This also means that $\mathcal{W}_{t_X} \neq \emptyset$ and $\mathcal{W}_t = \emptyset$ for all $t > t_X$. Hence, we have

$$Q_X = \bigcup_{0 \leq t \leq t_X} \bigcup_{Q_\sigma \in \mathcal{W}_t} Q_\sigma.$$

To facilitate the understanding of the sets Q_σ, and to illustrate the construction of the above partitioning, we construct a fairly large manual example.

Example 2.5.3 Consider the set $X = \{-3, -2, \ldots, 3\} \times \{-2, -1, \ldots, 3\}$. We wish to determine the Q_σ decomposition of X in $(-1 + i, \{0, 1\})$. Example 2.2.4 on page 85 presents the minimal addresses of the points of X in this base. They are reproduced in Figure 2.7 on the facing page. The addresses allow us to easily see what the Q_σ sets will be. However, let us go through their construction rigorously by growing the tiles around each point to their maximal size in X.

[7] By $\emptyset$ we mean here the empty collection.

Since the longest address in X is at point $-3+2i$, having length ten, then $L = 10$. Let $\mathcal{W} = \bigcup_{t=0}^{L-1} \mathcal{W}_t$. As an algorithm, to generate each Q_σ in $\mathcal{W}$, start with each point x in X in its own tile. These are the sets $Q_{a_x^L}$. We reconstruct the table of addresses here. However, we leave them unpadded since this makes it easier to see how to combine the sets Q_σ.

3	11010	11011	1110110	1110111	1010	1011	1100110
2	1110100101	11000	11001	1110100	1110101	1000	1001
1	11110	11111	10	11	1110	1111	111010010
0	10001	11100	11101	0	1	1100	1101
-1	11101010	11101011	110	111	111010	111011	111010110
-2	10101	11101000	11101001	100	101	111000	111001
	-3	-2	-1	0	1	2	3

Figure 2.7: Addresses of a set in base $(-1+i, \{0,1\})$.

Next, we combine the sets $Q_{a_x^L}$ to generate the sets of the next level. Looking at the addresses, this means combining all points having addresses with the same prefix of length $L-1$, padding short addresses to length L to start. If the parent tile does not intersect the complement of X in $\mathbb{Z}[i]$, then it is in $\mathcal{V}_{L-1}$. We keep growing tiles in this fashion, combining the sets $Q_{\sigma d}$ into Q_σ at each level. When a tile cannot grow and still be in some $\mathcal{V}_t$, that is, be completely contained in Q_X, then it is a maximal tile and hence in $\mathcal{W}$. Figure 2.8 on the next page demonstrates the growth of these tiles at each step. To see what these Q_σ sets are, prepad the lengths of all the addresses to length 10. This is demonstrated in Figure 2.9 on page 115.

Next, for each partitioned set, we need to determine the value of σ. This value is given by the common prefix of all the addresses in the boxes. For example, looking at the set containing the point $(-2, 2)$, all of the addresses are the same up to the seventh digit from the left. After that, the addresses differ. Therefore, this set is $Q_{(0000011)}$. The entire collection $\mathcal{W}$ is:

$$\mathcal{W} = \{Q_{(000000)}, Q_{(0000011)}, Q_{(00001110)}, Q_{(00011101)}, Q_{(00111010)}, Q_{(0000010101)},$$
$$Q_{(0000010001)}, Q_{(0001100110)}, Q_{(0111010010)}, Q_{(0111010110)}, Q_{(1110100101)}\}. \quad \blacksquare$$

Now, suppose $f = \sum_{x \in X} s_x \phi_x$. For $t = 0, \ldots, t_X$, $Q_\sigma \in \mathcal{W}_t$, define $X_\sigma = X \cap Q_\sigma$ and

	-3	-2	-1	0	1	2	3
3	11010	11011	1110110	1110111	1010	1011	1100110
2	1110100101	11000	11001	1110100	1110101	1000	1001
1	11110	11111	10	11	1110	1111	111010010
0	10001	11100	11101	0	1	1100	1101
-1	11101010	11101011	110	111	111010	111011	111010110
-2	10101	11101000	11101001	100	101	111000	111001

	-3	-2	-1	0	1	2	3
3	11010	11011	1110110	1110111	1010	1011	1100110
2	1110100101	11000	11001	1110100	1110101	1000	1001
1	11110	11111	10	11	1110	1111	111010010
0	10001	11100	11101	0	1	1100	1101
-1	11101010	11101011	110	111	111010	111011	111010110
-2	10101	11101000	11101001	100	101	111000	111001

	-3	-2	-1	0	1	2	3
3	11010	11011	1110110	1110111	1010	1011	1100110
2	1110100101	11000	11001	1110100	1110101	1000	1001
1	11110	11111	10	11	1110	1111	111010010
0	10001	11100	11101	0	1	1100	1101
-1	11101010	11101011	110	111	111010	111011	111010110
-2	10101	11101000	11101001	100	101	111000	111001

	-3	-2	-1	0	1	2	3
3	11010	11011	1110110	1110111	1010	1011	1100110
2	1110100101	11000	11001	1110100	1110101	1000	1001
1	11110	11111	10	11	1110	1111	111010010
0	10001	11100	11101	0	1	1100	1101
-1	11101010	11101011	110	111	111010	111011	111010110
-2	10101	11101000	11101001	100	101	111000	111001

Figure 2.8: Partitioning of a set with maximal subtiles.

3	0000011010	0000011011	0001110110	0001110111	0000001010	0000001011	0001100110
2	1110100101	0000011000	0000011001	0001110100	0001110101	0000001000	0000001001
1	0000011110	0000011111	0000000010	0000000011	0000001110	0000001111	0111010010
0	0000010001	0000011100	0000011101	0000000000	0000000001	0000001100	0000001101
-1	0011101010	0011101011	0000000110	0000000111	0000111010	0000111011	0111010110
-2	0000010101	0011101000	0011101001	0000000100	0000000101	0000111000	0000111001
	-3	-2	-1	0	1	2	3

Figure 2.9: Prepadded addresses of a partitioned set.

$f_\sigma = f \restriction_{Q_\sigma}$, the restriction of f to Q_σ. Then

$$f_\sigma = \sum_{x \in X_\sigma} s_x \phi_x$$

and hence we have an orthogonal decomposition of f as follows:

$$\begin{aligned} f &= \sum_{t=0}^{t_X} \sum_{Q_\sigma \in \mathcal{W}_t} f_\sigma \\ &= \sum_{t=0}^{t_X} \sum_{Q_\sigma \in \mathcal{W}_t} \sum_{x \in X_\sigma} s_x \phi_x \\ &= \sum_{t=0}^{t_X} \sum_{Q_\sigma \in \mathcal{W}_t} \left[s_{-t,\sigma} \phi_{-t,\sigma} + \sum_{p=1}^{q-1} \sum_{n=0}^{t-1} \sum_{\lambda \in D^n} w^p_{-t+n,\sigma\lambda} \psi^p_{-t+n,\sigma\lambda} \right] \end{aligned}$$

where $q = |\det B|$ and by using the convention that any sum of the form $\sum_{n=m_1}^{m_2}$ with $m_2 < m_1$ is taken to be the empty sum, with value zero.

We can now define an LIFSW associated with a number system in $\mathbb{R}^n$.

Definition 2.5.4 A *local IFS on wavelets* or *LIFSW* associated with X is an operator $W = \sum_{t=0}^{t_X} \sum_{Q_\sigma \in \mathcal{W}_t} W_\sigma$ where if $g = \sum_{x \in X} {}^g s_x \phi_x$ then[8]

[8]The notation ${}^g s_x$ is used simply to denote the scaling coefficients of g. Analogous notation is used for its wavelet coefficients.

$$W_\sigma(g) = s_{-t,\sigma}\phi_{-t,\sigma} + \sum_{l=1}^{q-1}\sum_{n=0}^{t-r_\sigma^l-1}\sum_{\lambda\in D^n} w^l_{-t+n,\sigma\lambda}\psi^l_{-t+n,\sigma\lambda}$$
$$+ \sum_{l=1}^{q-1}\sum_{\theta\in D^{t-r_\sigma^l}} \alpha^l_{\sigma\theta} \sum_{n=0}^{\min(-r_\sigma^l+d_\sigma^l-1,r_\sigma^l-1)}\sum_{\rho\in D^n} w^l_{-d_\sigma^l+n,\sigma\omega^l_{\sigma\theta}\rho}\psi^l_{-r_\sigma^l+n,\sigma\theta\rho}$$
$$+ \sum_{l=1}^{q-1}\sum_{\theta\in D^{t-r_\sigma^l}} \alpha^l_{\sigma\theta} \sum_{n=d_\sigma^l-r_\sigma^l}^{r_\sigma^l-1}\sum_{\rho\in D^n} {}^g w^l_{-d_\sigma^l+n,\sigma\omega^l_{\sigma\theta}\rho}\psi^l_{-r_\sigma^l+n,\sigma\theta\rho}$$

where $0 \le r_\sigma^l < d_\sigma^l \le t$ are integers, $\alpha^l_{\sigma\theta} \in \mathbb{R}$, $\omega^l_{\sigma\theta} \in D^{t-d_\sigma^l}$ and the scaling and wavelet coefficients are real numbers.[9]

We can simplify this notation by considering the following:

$$C_\sigma = s_{-t,\sigma}\phi_{-t,\sigma} + \sum_{l=1}^{q-1}\sum_{n=0}^{t-r_\sigma^l-1}\sum_{\lambda\in D^n} w^l_{-t+n,\sigma\lambda}\psi^l_{-t+n,\sigma\lambda}$$
$$D^l_{\sigma\theta} = \sum_{n=0}^{\min(-r_\sigma^l+d_\sigma^l-1,r_\sigma^l-1)}\sum_{\rho\in D^n} w^l_{-d_\sigma^l+n,\sigma\omega^l_{\sigma\theta}\rho}\psi^l_{-r_\sigma^l+n,\sigma\theta\rho}$$
$$E^l_{\sigma\theta}(g) = \sum_{n=d_\sigma^l-r_\sigma^l}^{r_\sigma^l-1}\sum_{\rho\in D^n} {}^g w^l_{-d_\sigma^l+n,\sigma\omega^l_{\sigma\theta}\rho}\psi^l_{-r_\sigma^l+n,\sigma\theta\rho}.$$

Then, we can write $W_\sigma(g)$ more compactly as

$$W_\sigma(g) = C_\sigma + \sum_{l=1}^{q-1}\sum_{\theta\in D^{t-r_\sigma^l}} \alpha^l_{\sigma\theta}\left[D^l_{\sigma\theta} + E^l_{\sigma\theta}(g)\right]$$

[9]Given the Fourier relationship between L^2 and ℓ^2, we will often talk about the action of W on the functions and the wavelet coefficient trees as best suits the discussion. Technically, the action on the functions should be written as T and the action on the coefficient trees written as M. Furthermore, in [27], instead of α the value used is $2^{(r-d)/2}\alpha$. This simply moves the power of 2 from one side of the Fourier relationship to the other. For simplicity, we have chosen to use α here.

and so

$$W(g) = \sum_{n=0}^{t_X} \sum_{Q_\sigma \in \mathcal{W}_t} \left[C_\sigma + \sum_{l=1}^{q-1} \sum_{\theta \in D^{t-r_\sigma^l}} \alpha_{\sigma\theta}^l \left[D_{\sigma\theta}^l + E_{\sigma\theta}^l(g) \right] \right].$$

By letting

$$D_\sigma^l = \sum_{\theta \in D^{t-r_\sigma^l}} \alpha_{\sigma\theta}^l D_{\sigma\theta}^l$$

and

$$E_\sigma^l(g) = \sum_{\theta \in D^{t-r_\sigma^l}} \alpha_{\sigma\theta}^l E_{\sigma\theta}^l(g),$$

then

$$W(g) = \sum_{n=0}^{t_X} \sum_{Q_\sigma \in \mathcal{W}_t} \left[C_\sigma + \sum_{l=1}^{q-1} \left[D_\sigma^l + E_\sigma^l(g) \right] \right].$$

Note that W is generally a non-linear operator.

Maintaining our convention regarding empty sums and sums from a larger to a smaller summand, we make a few notes about W_σ:

i) We call the level d_σ^l the *domain* or *parent* level and the level r_σ^l the *range* or *child* level. We often refer to the trees having their roots at these levels as *domain blocks* and *range blocks* respectively;

ii) When $r_\sigma^l - 1 \leq -r_\sigma^l + d_\sigma^l - 1$, then $r_\sigma^l - 1 < -r_\sigma^l + d_\sigma^l$. Hence, the third line of the definition of W_σ is well defined;

iii) When $t = 0$, then the three groups of summations disappear and $W_\sigma(g) = s_{-t,\sigma}\phi_{-t,\sigma}$. This is reasonable since in this case there is no wavelet decomposition to perform, wavelet coefficients being defined only for levels less than zero;

iv) When $r^l_\sigma = 0$, then the last two summations disappear and W_σ reduces to the constant operator

$$W_\sigma(g) = s_{-t,\sigma}\phi_{-t,\sigma} + \sum_{l=1}^{q-1}\sum_{n=0}^{t-1}\sum_{\lambda\in D^n} w^l_{-t+n,\sigma\lambda}\psi^l_{-t+n,\sigma\lambda};$$

v) When $d^l_\sigma = t$, there is only one domain to choose for any range and W_σ is a basic IFSW. This is demonstrated since $\omega^l_{\sigma\theta} \in D^{t-d^l_\sigma}$, hence we have each $\omega^l_{\sigma\theta}$ being the empty string and so they can be removed from the equations defining W_σ;

vi) For the third summation to be nonzero, we must have $d^l_\sigma \leq 2r^l_\sigma - 1$. Since $0 \leq r^l_\sigma < d^l_\sigma$, then the third summation can be nonzero only if $r^l_\sigma > 2$. That is, W_σ must act on a tree of at least 3 levels, i.e. $t \geq 3$.

To understand the action of W_σ, let us discuss the meaning of each of C_σ, $D^l_{\sigma\theta}$ and $E^l_{\sigma\theta}(g)$ in turn.

The piece C_σ is the *condensation* piece. Essentially, from levels $-t$ to $-t + r^l_\sigma - 1$, W_σ wipes out all of the information of g (or more precisely g_σ) and replaces it with a constant set of values.

The piece $D^l_{\sigma\theta}$ involves a shifting of coefficients. Here, W_σ is replacing the coefficients of g_σ from levels $-r^l_\sigma$ to $-2r^l_\sigma + d^l_\sigma - 1$ (or -1 as appropriate) by scaled versions of the coefficients on levels $-d^l_\sigma$ to (or up to) $-r^l_\sigma - 1$. For each root node of the $l - th$ wavelet tree on level $-r^l_\sigma$ (at address $\sigma\theta$), this is the action of multiplying the tree on level $-d^l_\sigma$ (with root at address $\sigma\omega^l_{\sigma\theta}$ and ending at level $-r^l_\sigma - 1$) by $\alpha^l_{\sigma\theta}$ and copying it over that

tree at level $-r_\sigma^l$.

The piece $E_{\sigma\theta}^l(g)$ is the same action as $D_{\sigma\theta}^l$ except that it copies trees with roots at level $-r_\sigma^l$ to trees with roots at level $-2r_\sigma^l + d_\sigma^l$. Following the convention, $E_{\sigma\theta}^l(g) = 0$ if $d_\sigma^l > 2r_\sigma^l$. Notice that this is the only part of W that depends on g.

One could define an LIFSW on infinite trees with a given initial root by considering radix representations. In this case, it would not be necessary to define the D_σ^l and E_σ^l separately since one would not have to worry about passing the bottom level of the tree. Instead, one could define the action of each $D_{\sigma\theta}^l$ as simply copying the entire, infinite, domain block to the entire, infinite, range block.

Figure 2.10 on the next page illustrates the action space of each component of W.

Example 2.5.5 We will complete the example we have been developing by constructing an LIFSW on the set $X = \{-3, -2, \ldots, 3\} \times \{-2, -1, \ldots, 3\}$. The Q_σ partition of X was shown in Figure 2.9 on page 115.

In any MRA associated with the base $(-1 + i, \{0, 1\})$, a function which is a linear combination of scaling functions on Q_X has the scaling tree shown in Figure 2.11 on page 121.

We can see clearly the Q_σ decomposition of X in this tree since the sets Q_σ have addresses which are the roots of the full subtrees of the scaling tree. For example, $Q_{(0000011)}$ is the fourth maximal full subtree from the left, with root at (0000011).

Since X has 11 maximal full subtrees (given by the sets in $\mathcal{W}$), to construct an LIFSW W, we have 11 W_σ's to construct. Let us look at them individually by value of $t = L-$ length(σ) in the definition of W.

Case $t = 0$**:** The sets Q_σ are $\{Q_{0000010101}, Q_{0000010001}, Q_{0001100110}, Q_{0111010010}, Q_{0111010110}, Q_{1110100101}\}$. As we noted earlier, in this case for any simple function g on Q_X, $W_\sigma(g) = s_{0,\sigma}\phi_{0,\sigma}$.

Case $t = 2$**:** The sets Q_σ are $\{Q_{(00001110)}, Q_{(00011101)}, Q_{(00111010)}\}$. In this case, we can have $0 \le r_\sigma < d_\sigma \le t = 2$ (for this MRA, there is only one mother wavelet, so we drop

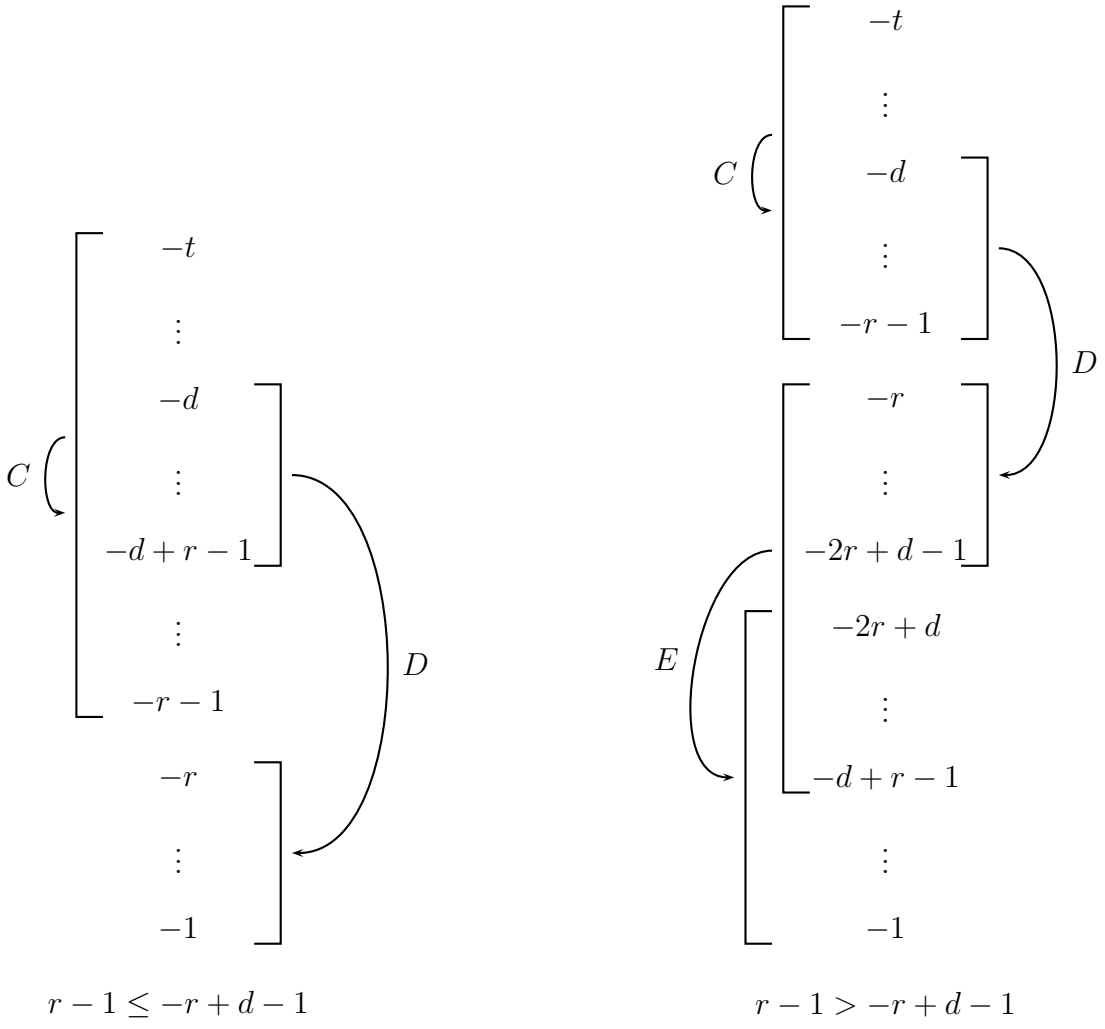


Figure 2.10: Action space of each component of an LIFSW.

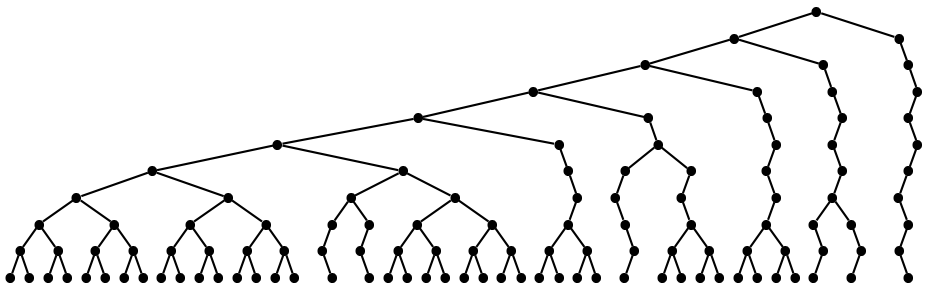


Figure 2.11: Scaling tree of a 42-point set.

the l). So, we can have $r = 0, d = 1$, $r = 0, d = 2$, and $r = 1, d = 2$. Since there are three trees, let us assign these three choices to each of the three sets in the order listed. Therefore,

$$W_{(00001110)}(g) = s_{-2,(00001110)}\phi_{-2,(00001110)} + \sum_{n=0}^{1}\sum_{\lambda\in D^n} w_{-2+n,(00001110)\lambda}\psi_{-2+n,(00001110)\lambda}.$$

$$W_{(00011101)}(g) = s_{-2,(00011101)}\phi_{-2,(00011101)} + \sum_{n=0}^{1}\sum_{\lambda\in D^n} w_{-2+n,(00011101)\lambda}\psi_{-2+n,(00011101)\lambda}$$

$$\begin{aligned} W_{(00111010)}(g) = {} & s_{-2,(00111010)}\phi_{-2,(00111010)} + w_{-2,(00111010)}\psi_{-2,(00111010)} \\ & + \alpha_{(001110100)}w_{-2,(00111010)}\psi_{-1,(001110100)} \\ & + \alpha_{(001110101)}w_{-2,(00111010)}\psi_{-1,(001110101)}. \end{aligned}$$

Case $t = 3$: Here there is only $Q_{(0000011)}$ and we can choose $0 \le r_\sigma < d_\sigma \le 3$. Let us take $d_\sigma = 2$ and $r_\sigma = 1$. Therefore,

$$C_\sigma = s_{-3,\sigma}\phi_{-3,\sigma} + w_{-3,\sigma}\psi_{-3,\sigma} + w_{-2,\sigma0}\psi_{-2,\sigma0} + w_{-2,\sigma1}\psi_{-2,\sigma1}.$$

To determine D_σ, we need to make our choice of domain for each range, that is, assign a value to each $\omega_{\sigma\theta}$. Table 2.1 on the next page shows the choice.

θ	$\omega_{\sigma\theta}$
00	0
01	1
10	0
11	1

Table 2.1: Mappings of D_σ for an LIFSW.

Therefore,

$$D_\sigma = \alpha_{\sigma 00} w_{-2,\sigma 0} \psi_{-1,\sigma 00} + \alpha_{\sigma 01} w_{-2,\sigma 1} \psi_{-1,\sigma 01} + \alpha_{\sigma 10} w_{-2,\sigma 0} \psi_{-1,\sigma 10} + \alpha_{\sigma 11} w_{-2,\sigma 1} \psi_{-1,\sigma 11}.$$

Finally, this choice implies $E_\sigma = 0$.

Case $t = 4$**:** Again, there is only one set: $Q_\sigma = Q_{(000000)}$ and we can choose any $0 \leq r_\sigma < d_\sigma \leq 4$. So far, we have had E_σ=0. So, we will choose $d_\sigma = 4$ and $r_\sigma = 3$. In this case, $C_\sigma = s_{-4,\sigma}\phi_{-4,\sigma} + w_{-4,\sigma}\psi_{-4,\sigma}$. For D_σ, we have $\min(-r_\sigma + d_\sigma - 1, r_\sigma - 1) = 0$. Therefore,

$$D_\sigma = \sum_{d \in D} \alpha_{\sigma d} w_{-4,\sigma} \psi_{-3,\sigma d}.$$

Now, $r_\sigma - 1 = 2$ and $d_\sigma - r_\sigma = 1$, hence E_σ is given by

$$E_\sigma = \sum_{d \in D} \alpha_{\sigma d} \sum_{n=1}^{2} \sum_{\rho \in D^n} {}^{g}w_{-4+n,\sigma\rho} \psi_{-3+n,\sigma d\rho}.$$

Since there is only one domain block, there is no choice in defining E_σ. Table 2.2 on the facing page lists the mappings of E_σ.

The action of some of the W_σ operators are illustrated in Figures 2.12, 2.13 and 2.14 on page 124.

d	n	ρ	$\sigma\rho$	$\sigma d\rho$
0	1	0	$\sigma 0$	$\sigma 00$
		1	$\sigma 1$	$\sigma 01$
	2	00	$\sigma 00$	$\sigma 000$
		01	$\sigma 01$	$\sigma 001$
		10	$\sigma 10$	$\sigma 010$
		11	$\sigma 11$	$\sigma 011$
1	1	0	$\sigma 0$	$\sigma 10$
		1	$\sigma 1$	$\sigma 11$
	2	00	$\sigma 00$	$\sigma 100$
		01	$\sigma 01$	$\sigma 101$
		10	$\sigma 10$	$\sigma 110$
		11	$\sigma 11$	$\sigma 111$

Table 2.2: Mappings of E_σ for an LIFSW.

Figure 2.12: The actions of $W_{(00001110)}$ (left) and $W_{(00111010)}$ (right) on the wavelet tree of $g_{(00001110)}$ and $g_{(00111010)}$ respectively.

Figure 2.13: Action of $W_{(0000011)}$ on the wavelet tree of $g_{(0000011)}$.

Figure 2.14: Action of $W_{(000000)}$ on the wavelet tree of $g_{(000000)}$.

2.6 Inverse Problem Revisited

We now revisit the inverse problem in the light of the generalized LIFSW. It is first important to determine the contractivity of the LIFSW in some complete metric space. When it is contractive, the iterates of an LIFSW applied to any initial function will converge to its unique fixed point, meaning that the LIFSW provides all the information necessary to reconstruct that fixed point.

When iterating the LIFSW as defined in the previous section, notice that after the first iteration, the operators C_σ and D^l_σ no longer contribute any new information. Thus, if all the $E^l_{\sigma\theta}$ are the zero operator, one only needs perform a single iteration to construct the attractor of W. Otherwise, on subsequent iterations, each $E^l_{\sigma\theta}$ component adds $d_\sigma - r_\sigma$ successive new levels towards the fixed point. This is the situation described in Theorem 1.4.39. Thus, one can compute successive levels at each iteration, without computing previous levels that are already fixed by W.

Consider the L^2-norm of the difference of W applied to two functions g and h. Since the C and D components are the same for each function, and using the orthonormality of the wavelets we have

$$\begin{aligned} ||W(g) - W(h)||^2 &= \left|\left| \sum_{t=0}^{t_X} \sum_{Q_\sigma \in W_t} \sum_{l=1}^{q-1} \sum_{\theta \in D^{t-r^l_\sigma}} \alpha^l_{\sigma\theta} [E^l_{\sigma\theta}(g) - E^l_{\sigma\theta}(h)] \right|\right|^2 \\ &= \sum_{t=0}^{t_X} \sum_{Q_\sigma \in W_t} \sum_{l=1}^{q-1} \sum_{\theta \in D^{t-r^l_\sigma}} |\alpha^l_{\sigma\theta}|^2 \left|\left| E^l_{\sigma\theta}(g) - E^l_{\sigma\theta}(h) \right|\right|^2 . \end{aligned}$$

Now consider the difference of the E operators:

$$\left|\left|E^l_{\sigma\theta}(g) - E^l_{\sigma\theta}(h)\right|\right|^2 = \left|\left| \sum_{n=d^l_\sigma - r^l_\sigma}^{r^l_\sigma - 1} \sum_{\rho\in D^n} \left({}^g w^l_{-d^l_\sigma+n,\sigma\omega^l_{\sigma\theta}\rho} - {}^h w^l_{-d^l_\sigma+n,\sigma\omega^l_{\sigma\theta}\rho} \right) \psi^l_{-r^l_\sigma+n,\sigma\theta\rho} \right|\right|^2$$
$$= \sum_{n=d^l_\sigma - r^l_\sigma}^{r^l_\sigma - 1} \sum_{\rho\in D^n} \left| {}^g w^l_{-d^l_\sigma+n,\sigma\omega^l_{\sigma\theta}\rho} - {}^h w^l_{-d^l_\sigma+n,\sigma\omega^l_{\sigma\theta}\rho} \right|^2 .$$

Let $M = \max\{|\alpha^l_{\sigma\theta}|\}$. Then

$$||W(g) - W(h)||^2 \le M^2 \sum_{t=0}^{t_X} \sum_{Q_\sigma\in W_t} \sum_{l=1}^{q-1} \sum_{\theta\in D^{t-r^l_\sigma}} \sum_{n=d^l_\sigma - r^l_\sigma}^{r^l_\sigma - 1} \sum_{\rho\in D^n} \left| {}^g w^l_{-d^l_\sigma+n,\sigma\omega^l_{\sigma\theta}\rho} - {}^h w^l_{-d^l_\sigma+n,\sigma\omega^l_{\sigma\theta}\rho} \right|^2$$
$$\le M^2 \left|\left|g - h\right|\right|^2 ,$$

by letting the summations run over all wavelet coefficients of g and h. Therefore, W is uniformly continuous for any choice of $\alpha^l_{\sigma\theta}$. If $M < 1$, then W is contractive in the L^2-norm. We could also consider the metric presented in [27].

Now, given an LIFSW operator, consider the collage distance between f and $W(f)$, where the wavelet coefficients defining W are precisely the coefficients of f. For simplicity, we can omit the superscript f for the wavelet coefficients. We then have

$$\Delta^2 = ||f - W(f)||^2$$
$$= \left|\left| \sum_{t=0}^{t_X} \sum_{Q_\sigma\in\mathcal{W}_t} \sum_{l=1}^{q-1} \sum_{\theta\in D^{t-r^l_\sigma}} \sum_{n=0}^{r^l_\sigma - 1} \sum_{\rho\in D^n} \left[w^l_{-r^l_\sigma+n,\sigma\theta\rho} - \alpha^l_{\sigma\theta} w^l_{-d^l_\sigma+n,\sigma\omega^l_{\sigma\theta}\rho} \right] \psi^l_{-r^l_\sigma+n,\sigma\theta\rho} \right|\right|^2$$
$$= \sum_{t=0}^{t_X} \sum_{Q_\sigma\in\mathcal{W}_t} \sum_{l=1}^{q-1} \sum_{\theta\in D^{t-r^l_\sigma}} \sum_{n=0}^{r^l_\sigma - 1} \sum_{\rho\in D^n} \left| w^l_{-r^l_\sigma+n,\sigma\theta\rho} - \alpha^l_{\sigma\theta} w^l_{-d^l_\sigma+n,\sigma\omega^l_{\sigma\theta}\rho} \right|^2$$

$$= \sum_{t=0}^{t_X} \sum_{Q_\sigma \in \mathcal{W}_t} \sum_{l=1}^{q-1} \sum_{\theta \in D^{t-r_\sigma^l}} \Delta^2_{\sigma,\theta,\omega_{\sigma\theta}^l}$$

by the orthonormality conditions on the wavelets. Therefore, minimizing the collage distance is equivalent to minimizing each $\Delta^2_{\sigma,\theta,\omega_{\sigma\theta}^l}$. As before, this can be done using *least squares* minimization to find the optimal scaling factors $\bar{\alpha}_{\sigma\theta}^l$:

$$\begin{aligned}
0 &= \frac{\partial \Delta^2_{\sigma,\theta,\omega_{\sigma\theta}^l}}{\partial \bar{\alpha}_{\sigma\theta}^l} \\
&= -2 \sum_{n=0}^{r_\sigma^l - 1} \sum_{\rho \in D^n} (w^l_{-r_\sigma^l + n, \sigma\theta\rho} - \bar{\alpha}_{\sigma\theta}^l w^l_{-d_\sigma^l + n, \sigma\omega_{\sigma\theta}^l \rho}) w^l_{-d_\sigma^l + n, \sigma\omega_{\sigma\theta}^l \rho} \\
\iff 0 &= \bar{\alpha}_{\sigma\theta}^l \sum_{n=0}^{r_\sigma^l - 1} \sum_{\rho \in D^n} (w^l_{-d_\sigma^l + n, \sigma\omega_{\sigma\theta}^l \rho})^2 - \sum_{n=0}^{r_\sigma^l - 1} \sum_{\rho \in D^n} w^l_{-r_\sigma^l + n, \sigma\theta\rho} w^l_{-d_\sigma^l + n, \sigma\omega_{\sigma\theta}^l \rho} \\
\iff \bar{\alpha}_{\sigma\theta}^l &= \frac{\sum_{n=0}^{r_\sigma^l - 1} \sum_{\rho \in D^n} w^l_{-r_\sigma^l + n, \sigma\theta\rho} w^l_{-d_\sigma^l + n, \sigma\omega_{\sigma\theta}^l \rho}}{\sum_{n=0}^{r_\sigma^l - 1} \sum_{\rho \in D^n} (w^l_{-d_\sigma^l + n, \sigma\omega_{\sigma\theta}^l \rho})^2}.
\end{aligned}$$

Formally, define

$$S^l_{\sigma,a,\gamma,b,\omega} = \sum_{n=0}^{r_\sigma^l - 1} \sum_{\rho \in D^n} w^l_{-a+n,\sigma\gamma\rho} w^l_{-b+n,\sigma\omega\rho}.$$

Given that each node has the same number of children, the above summation can be thought of naturally as an inner product of coefficient trees at different levels (c.f. [86]). With this notation,

$$\bar{\alpha}_{\sigma\theta}^l = \frac{S^l_{\sigma, r_\sigma^l, \theta, d_\sigma^l, \omega_{\sigma\theta}^l}}{S^l_{\sigma, d_\sigma^l, \omega_{\sigma\theta}^l, d_\sigma^l, \omega_{\sigma\theta}^l}}$$

and the minimized squared collage distance $\bar{\Delta}^2_{\sigma,\theta,\omega_{\sigma\theta}^l}$ can be found as follows:

$$\bar{\Delta}^2_{\sigma,\theta,\omega_{\sigma\theta}^l} = \sum_{n=0}^{r_\sigma^l - 1} \sum_{\rho \in D^n} \left((\bar{\alpha}_{\sigma\theta}^l)^2 (w^l_{-d_\sigma^l + n, \sigma\omega_{\sigma\theta}^l \rho})^2 - 2\bar{\alpha}_{\sigma\theta}^l w^l_{-d_\sigma^l + n, \sigma\omega_{\sigma\theta}^l \rho} w^l_{-r_\sigma^l + n, \sigma\theta\rho} + (w^l_{-r_\sigma^l + n, \sigma\theta\rho})^2 \right)$$

$$
\begin{aligned}
&= (\bar{\alpha}^l_{\sigma\theta})^2 \sum_{n=0}^{r^l_\sigma-1} \sum_{\rho\in D^n} (w^l_{-d^l_\sigma+n,\sigma\omega^l_{\sigma\theta}\rho})^2 - 2\bar{\alpha}^l_{\sigma\theta} \sum_{n=0}^{r^l_\sigma-1} \sum_{\rho\in D^n} w^l_{-d^l_\sigma+n,\sigma\omega^l_{\sigma\theta}\rho} w^l_{-r^l_\sigma+n,\sigma\theta\rho} \\
&\quad + \sum_{n=0}^{r^l_\sigma-1} \sum_{\rho\in D^n} (w^l_{-r^l_\sigma+n,\sigma\theta\rho})^2 \\
&= \frac{(S^l_{\sigma,r^l_\sigma,\theta,d^l_\sigma,\omega^l_{\sigma\theta}})^2}{(S^l_{\sigma,d^l_\sigma,\omega^l_{\sigma\theta},d^l_\sigma,\omega^l_{\sigma\theta}})^2} S^l_{\sigma,d^l_\sigma,\omega^l_{\sigma\theta},d^l_\sigma,\omega^l_{\sigma\theta}} - 2\frac{S^l_{\sigma,r^l_\sigma,\theta,d^l_\sigma,\omega^l_{\sigma\theta}}}{S^l_{\sigma,d^l_\sigma,\omega^l_{\sigma\theta},d^l_\sigma,\omega^l_{\sigma\theta}}} S^l_{\sigma,d^l_\sigma,\omega^l_{\sigma\theta},r^l_\sigma,\theta} + S^l_{\sigma,r^l_\sigma,\theta,r^l_\sigma,\theta} \\
&= S^l_{\sigma,r^l_\sigma,\theta,r^l_\sigma,\theta} - \bar{\alpha}^l_{\sigma\theta} S^l_{\sigma,r^l_\sigma,\theta,d^l_\sigma,\omega^l_{\sigma\theta}}
\end{aligned}
$$

since $S^l_{\sigma,r^l_\sigma,\theta,d^l_\sigma,\omega^l_{\sigma\theta}} = S^l_{\sigma,d^l_\sigma,\omega^l_{\sigma\theta},r^l_\sigma,\theta}$ (c.f. [75]).

Chapter 3

Application to Image Coding

3.1 Approximation of Images

In this chapter we discuss the application of number systems to image processing. There are two basic concepts that first need to be defined: (1) *What is a discrete image?* and (2) *How do we compare two discrete images?*

Definition 3.1.1 A *discrete image* is defined here as a real-valued function $u : \mathbb{Z}^n \to \mathbb{R}$, $n = 1, 2, 3, \ldots$ with compact support. Elements of this finite subset are called the *pixels* of the image.

In practical applications, the range $\mathbb{R}$ is, in fact, replaced by a finite set $\{r_1, ..., r_m\}$ of grey-scale values, reflecting the way in which images are stored in computers. A finite integer number nb of computer bits are allocated to each pixel. The grey-scale value u associated with each pixel is then an integer between 0 and $2^{nb} - 1$. (Typically, images are stored at $nb = 8$ bits per pixel, implying that $m = 256$ above.) The conversion of a real-valued discrete image function value to an integer value is known as *quantization*. Discussing how "real" images, that is, photos defined over continuous spatial domains, are

spatially digitized to produce a lattice of pixels, is beyond the scope of this work.

In the following presentation, we limit the discussion to two-dimensional images, i.e. $n = 2$, although the methods developed apply to any dimension.

Once one has a digital image (by which we mean a discrete image stored on digital media), one often wants to process it in some way (e.g. compress or transmit). Processing an image creates a modified version of the original. If the processing is supposed to maintain the overall integrity of the original image, as is the case in both compression and transmission, it is important to develop a method of comparing images. This helps determine the effectiveness of the processing. Since images are functions, comparisons can be done using a metric.

There is an entire field of study dedicated to metrics and other forms of comparison on images. Indeed, the holy grail of this field is to define a method of comparison that exactly matches the way human vision distinguishes objects. A discussion of such "visual quality metrics" is beyond the scope of this thesis. The reader is encouraged to refer to the human vision and image processing literature for further information. For an excellent introduction to the area of human vision, see [88].

One metric which is used is simply the normalized L^2 distance, often called the *root-mean-squared error* (or *RMSE*) of the images. For example, if we have two images u and v defined on $i, j = 0, \ldots, 511$, with pixel values $\{u_{ij}\}$ and $\{v_{ij}\}$ respectively, ranging between the values $0, \ldots, 255$, then the RMSE of the two images is defined as

$$d_2(u, v) = \frac{1}{512} \left(\sum_{i,j=0}^{511} |u_{ij} - v_{ij}|^2 \right)^{1/2}.$$

Although this metric is useful, another measure of distance is more often used in the field of image processing when one image is an approximation of the other: the *peak signal-to-noise ratio* (or PSNR). It is based on the RMSE and defined as follows:

$$\mathrm{PSNR}(u,v) = 20\log_{10}\left(\frac{255}{d_2(u,v)}\right).$$

This value is measured in decibels (dB). This is not a metric since its value is zero when the distance between the two images is 255, or the maximal value for the given grey-scale used. Note also that as the distance $d_2(u,v)$ approaches zero, the PSNR approaches infinity.

3.2 Image Compression

We now illustrate applications of the methods presented earlier in this thesis to image compression. We begin by discussing different ways to compress images using these methods and then provide examples.

In essence, compressing an image means representing the image in a way that requires less information to store than the original representation. For computer images, this means storing an image with less computer memory than the original. There are two basic categories of compression: *lossless* and *lossy*. In lossless compression, there is no change in the quality of the image, whereas in lossy compression, there generally is. This distinction is important. For example, when storing a text document, one can't afford to lose a few letters here and there, so lossless compression would be preferred. For images, lossy compression is acceptable, provided the compressed image maintains the key visual characteristics of the original. The methods discussed here are forms of lossy compression.

Almost all methods of wavelet compression involve performing a wavelet decomposition

on an image and removing wavelet coefficients that, in a sense, provide little information to the viewer of the image. These are generally coefficients with absolute values smaller than some threshold value. The general motivation behind this methodology is based on the principle that the wavelet coefficients can provide information on edges in the image. As well, the most significant wavelet coefficients (i.e. those of highest magnitude) correspond to edges in the image. The scaling coefficients provide information on the values of the image at given points, whereas the wavelet coefficients provide information on the frequencies, or differences between the values of the subtiles (hence no wavelet coefficients on level zero).

For example, consider an image tiled with the twin-dragon tiling. Suppose that on level zero of the scaling coefficient tree, two sibling scaling coefficients have the same value. Hence, the function takes the same value on both subtiles of these coefficients' parent. Since the relation that $g_0 + g_1 = 0$ holds for the Haar wavelet filters, the wavelet coefficient of the parent is zero. This illustrates the fact that there is effectively no "edge" between the two subtiles, as far as an observer is concerned. Therefore, throwing out that wavelet coefficient removes no information from the image. Similarly, if the wavelet coefficient is small in absolute value, this relates to an edge between two shades of grey that are close together. On the other hand, if the values on the subtiles are substantially different, then the wavelet coefficient will have a large absolute value.

The discussion above was not intended to be mathematically rigorous but rather to give the reader an idea of the motivation for these basic wavelet compression techniques. Indeed, there are a few mathematical results that express the correlation between smooth regions and edges and wavelet coefficients (for example, the decay of wavelet coefficients according

to the degree of local regularity and irregularity at a point in an image [46, 47]). Some of these correlations have been exploited [24]. However, in most practical applications to date, the more general methods of pruning and thresholding have been used to obtain very good compression behaviour. For more information on this topic, the reader is referred to [82].

Before going further, it should be noted that the emphasis of this thesis was the development of the mathematics behind the methods as opposed to the construction of feasible image compression schemes. For this reason, no effort was made to optimize the implementations of the algorithms, nor has coding been done to store the compressed images. In a proper analysis of a compression scheme, one compares the amount of computer memory required to store the compressed image to the amount required by other schemes. Because the compression aspect of our applications is not optimized (i.e. proper attention to quantization, Huffman coding of quantized coefficients, etc.) no such analyses are performed here.

The general process for implementing our results on discrete images in the framework of image compression is as follows:

i) Consider a discrete image with pixels as Gaussian integers in $\mathbb{C}$. The pixels form the set X. For example, a 512×512 pixel image would be sitting in the region $\{x + iy \ : \ 0 \leq x, y < 512\}$. An eight-bit-per-pixel digitized representation implies that each pixel assumes one of $2^8 = 256$ values, typically 0 to 255. Each of these values represents a *grey-scale* associated with a pixel. Typically 0 represents *black* and 255 represents *white*, with 254 intermediate shades of grey lying between these two extremes;

ii) Choose a valid base (b, D) in which to expand the image. This defines an MRA and a scaling function ϕ, which is the characteristic function of $T(b, D)$;

iii) The image is assumed to be the function $f = \sum_{x \in X} s_x \phi_x$, where s_x is the value of the pixel at position $x \in X$;

iv) Find the address of each pixel in the image. The longest (minimal) address determines $L_{X,(b,D)}$. The pixels with shorter addresses are prepadded with zeros, up to the longest address length. This gives the set A_X;

v) Choose a wavelet basis using Lemma 1.4.21;

vi) Choose a method of extending the image to X_L. Choosing to not extend the function is like choosing to extend it using the function $M = 0$ for the extension algorithm (see Equation (2.7) on page 109);

vii) Perform the wavelet decomposition on the image (or its extension). At each level of the decomposition, the address length decreases by one. The decomposition algorithm ends when the length is zero;

viii) Implement a form of compression on the wavelet tree, such as a zero-tree or *LIFSW* method (see below);

ix) Store the compressed image in some way, possibly utilizing other compression schemes such as Huffman encoding;

x) Perform the wavelet reconstruction algorithm to reconstruct the approximation of the original image.

The following are sample methods of wavelet compression:

Pruning: This method consists of simply choosing a *threshold* level t and zeroing all the wavelet coefficients below that level. This is the simplest form of wavelet compression and, consequently, usually yields the least favorable results. Indeed, in this method, wavelet coefficients of high magnitude may be discarded.

Thresholding: This method is often called the *zero-tree* method, although the following description is a simplification of the algorithm. In this method, one determines a *threshold* value by which to measure the importance of the wavelet coefficients. One then navigates each successive level of the tree (or trees), starting at the root (level L) to determine which nodes to remove. At each level of the tree, starting at the root of a given subtree, one calculates the squared l^2 norm of the subtree. For example, if the root node is $w_{-t,\sigma}$, then the squared l^2 norm of the subtree is $\sum_{n=0}^{t-1} \sum_{\rho \in D^n} |w_{-t+n,\sigma\rho}|^2$. If this norm is less than the threshold, that subtree is considered of little importance to the quality of the image. In the terminology of image compression, the subtree is said to be *zeroed-out*, denoting that all of its wavelet coefficients are set to zero. One then moves to the next subtree on that level. Once all the subtrees on that level have been thresholded, one proceeds to threshold the subtrees of the next level. The algorithm ignores the subtrees whose parents have already been zeroed-out. Variants on this method provide some of the best compression results to date, for example, the method of *set partitioning in hierarchical trees*, or *SPIHT*: for short [80].

LIFSW: This method, referred to as *fractal-wavelet* compression, utilizes the idea of the Inverse Problem (see Sections 1.2 and 2.6). Recall that in its basic form, fractal-wavelet compression seeks to replace lower wavelet subtrees by suitably scaled copies

of higher wavelet subtrees. In particular, the idea is to choose a domain level d^l and a range level r^l in each wavelet tree of the image function. For each range block, one calculates the minimized collage distance of this block with each domain block. The domain block for which the minimized collage distance is the smallest is retained. This gives for each range block (with address γ, say) a unique domain block at address $\omega^{-1}(\gamma)$ and a scaling factor $\bar{\alpha}$. Once these comparisons have been completed, the compressed image is the LIFSW operator defined by the scaling coefficient at level L, the wavelet coefficients on levels L up to level $-r^l - 1$ (which defines C) and the mappings ω and the scaling coefficients $\bar{\alpha}$ (which defines the operators D and E). Iterating this LIFSW generates its fixed point, which by the Collage Theorem is an approximation to the original image.

In the generalized case, one can choose different levels for the domains and ranges for each full subtree of the wavelet tree. After this, the process is the same as described above.

Using the extension algorithm in image compression is somewhat counterproductive. Technically, no additional information is created when extending the image. However, there are many more coefficients than in the non-extended function. Table 3.1 on the next page gives the ratio of the number of points in the extension to the number of points in the original image. However, schemes could be developed to utilize only certain parts of the extension to enhance the compression method, for example, simply allowing the removal of artefacts.

		N ($N \times N$ pixels)	16	32	64	128	256	512	1024
Base	Norm	# points in image	256	1024	4096	16384	65536	262144	1048576
$(-1+i, \{0,1\})$	2	Levels	12	15	17	20	20	23	25
		# points in extension	4096	3.3E4	1.3E5	1.0E6	1.0E6	8.4E6	3.4E7
		Ratio	16	32	32	64	16	32	32
$(-2+i, \{0,\ldots,4\})$	5	Levels	7	8	9	9	11	11	13
		# points in extension	7.8E4	3.9E5	2.0E6	2.0E6	4.9E7	4.9E7	1.2E9
		Ratio	305	381	477	119	745	186	1164
$(-3+i, \{0,\ldots,9\})$	10	Levels	6	6	7	7	8	9	9
		# points in extension	1.0E6	1.0E6	1.0E7	1.0E7	1.0E8	1.0E9	1.0E9
		Ratio	3906	977	2441	610	1526	3815	954
$(-4+i, \{0,\ldots,16\})$	17	Levels	5	5	6	7	7	7	8
		# points in extension	1.4E6	1.4E6	2.4E7	4.1E8	4.1E8	4.1E8	7.0E9
		Ratio	5546	1387	5893	25045	6261	1565	6653
$(2+i, \{0,1,i,-i,-2-3i\})$	5	Levels	7	9	10	11	12	13	14
		# points in extension	7.8E4	2.0E6	9.8E6	4.9E7	2.4E8	1.2E9	6.1E9
		Ratio	305	1907	2384	2980	3725	4657	5821

Table 3.1: Tree depths for various complex bases.

3.3 Examples

This section presents examples of the methods discussed in the previous section. This first example is simply an illustration of converting an image to a function.

Example 3.3.1 (Mallat algorithm) Suppose we have the following discrete image of $2 \times 2 = 4$ pixels that we wish to decompose in the usual Haar MRA associated with $(-1+i, \{0,1\})$:

1	2
0	1

We associate these four grey-scale values with the pixel locations $0, 1, i, 1+i$ in $\mathbb{C}$, the bottom lefthand corner being the integer 0. Therefore, this image is simply the function $f = \phi_1 + \phi_i + 2\phi_{1+i}$, and so the decomposition is the same as the one given in Example 2.2.9.

Example 3.3.2 (Pruning) Consider a 512×512 image of Lena as shown in the top lefthand corner of Figure 3.1 on page 139. Figures 3.1, 3.2 and 3.3 show the result of performing the pruning algorithm on Lena at various levels and using different bases. Table 3.2 on page 142 gives the compression ratio and PSNR values of the pruned images. Here, the compression ratio is defined as the ratio of the number of wavelet coefficients in the original image to the number of wavelet coefficients remaining after pruning. Notice that in some cases, the first few levels of pruning yield almost completely black images.

Upon careful inspection, the tile edges can be seen. This is due to the extension problem where the majority of the points in the supertile are black.

Example 3.3.3 (Thresholding) Figures 3.4, 3.5 and 3.6 on pages 143 to 145 show the result of performing the zero-tree algorithm on Lena at various levels and using different bases. Table 3.3 on page 146 gives the compression ratio and PSNR values of the compressed images. The compression ratio is again defined here as the ratio of the number of wavelet coefficients in the original image to the number of wavelet coefficients remaining after thresholding.

Example 3.3.4 (LIFSW) Figures 3.7, 3.8 and 3.9 on pages 147 to 149 show the result of performing the LIFSW algorithm on Lena with various domain-range level pairs and using different bases. Table 3.4 on page 150 gives the compression ratio and PSNR values of the compressed images. In this instance, the compression ratio is defined in a slightly different fashion: the number of bits needed to store the original image divided by the number of bits needed to store the fractal-wavelet code. It is important to note that this compression ratio cannot be compared with the compression ratio of the other two schemes. This ratio is presented solely to enable a comparison between the different LIFSW images. The reader is cautioned against using it to draw conclusions regarding the usefulness of the LIFSW method.

In our examples, we are considering 512×512 images with 256 grey-scales. It takes 8 bits to store each grey-scale value. Hence, the image takes 2^{21} bits to store. The fractal-wavelet code consists of the wavelet coefficients below the range level (level $-L$ to $-r^l+1$), the map that associates to each range block, the domain block chosen in the algorithm, and the α coefficients. We assume that the wavelet coefficients and α coefficients are stored with 8-bit precision and the maps with 2 bit precision.

Notice that if the range level is taken to be too close to -1, one actually gets an expansion rather than a compression. This is caused by the need to have "more levels than necessary" in the tree since the image is inside a much larger supertile. One could improve the results by placing the image in an optimal location in the supertile, that is, one that minimizes the length of the image in the supertile.

Figure 3.1: Pruning with base $(-1+i, \{0,1\})$: The original 512×512 image of Lena is in the top lefthand corner. The other images, from top to bottom, left to right, result from pruning the wavelet trees at levels -21 to -1 by increments of two.

Figure 3.2: Pruning with base $(-2+i, \{0,1,2,3,4\})$: The original 512×512 image of Lena is in the top lefthand corner. The other images, from top to bottom, left to right, result from pruning the wavelet trees at levels -11 to -1.

Figure 3.3: Pruning with base $(2+i, \{0, 1, i, -i, -2-3i\})$: The original 512×512 image of Lena is in the top lefthand corner. The other images, from top to bottom, left to right, result from pruning the wavelet trees at levels -11 to -1.

Base	Level										
$-1+i$	-21	-19	-17	-15	-13	-11	-9	-7	-5	-3	-1
Compression	8.8E4	3.8E4	1.9E4	9.1E3	3.8E3	1.4E3	4.1E2	1.2E2	3.0E1	7.8	2.0
PSNR (dB)	7.2	8.2	10.2	12.6	14.2	15.9	18.3	20.5	23.8	26.8	33.9

Base	Level										
$-2+i$	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
Compression	6.7E4	2.2E4	1.3E4	6.1E3	2.8E3	1.1E3	3.6E2	9.7E1	2.2E1	4.8	1.0
PSNR (dB)	6.7	7.1	8.6	10.5	12.8	14.8	16.9	19.9	23.7	28.2	49.0

Base	Level										
$2+i$	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
Compression	2.2E4	1.1E4	6.7E3	4.5E3	2.9E3	1.7E3	8.6E2	3.2E2	9.1E1	2.2E1	4.8
PSNR (dB)	6.6	6.7	7.0	8.2	10.1	12.6	14.6	16.7	18.9	21.8	25.8

Table 3.2: Pruning: The above tables give the compression ratio and PSNR of the images in Figures 3.1, 3.2 and 3.3.

Figure 3.4: Thresholding with base $(-1+i, \{0,1\})$: The original 512×512 image of Lena is on the top. The other images, from top to bottom, left to right, result from thresholding the wavelet trees using threshold values 1000, 5000, 10000, 15000, 20000 and 25000.

Figure 3.5: Thresholding with base $(-3+i, \{0, \ldots, 9\})$: The original 512×512 image of Lena is on the top. The other images, from top to bottom, left to right, result from thresholding the wavelet trees using threshold values 1000, 5000, 10000, 15000, 20000 and 25000.

Figure 3.6: Thresholding with base $(2+i, \{0, 1, i, -i, -2-3i\})$: The original 512×512 image of Lena is in the top lefthand corner. The other images, from top to bottom, left to right, result from thresholding the wavelet trees using threshold values 1000, 5000, 10000, 15000, 20000 and 25000.

Base $-1+i$	Threshold 1000	5000	10000	15000	20000	25000
Compression	9.2	19.2	27.6	34.3	40.8	46.7
PSNR (dB)	32.4	28.7	27.1	26.2	25.7	25.3

Base $-2+i$	Threshold 1000	5000	10000	15000	20000	25000
Compression	11.3	24.7	35.4	44.8	52.2	58.4
PSNR (dB)	32.1	27.9	26.3	25.3	24.7	24.3

Base $-3+i$	Threshold 1000	5000	10000	15000	20000	25000
Compression	11.7	26.5	39.5	49.8	59.0	67.4
PSNR (dB)	31.8	27.5	25.8	25.0	24.4	23.9

Base $-4+i$	Threshold 1000	5000	10000	15000	20000	25000
Compression	11.5	26.6	39.4	50.4	60.1	68.6
PSNR (dB)	31.3	27.1	25.4	24.5	23.9	23.5

Base $2+i$	Threshold 1000	5000	10000	15000	20000	25000
Compression	7.0	15.6	23.9	30.6	36.6	42.2
PSNR (dB)	30.0	25.4	23.8	23.0	22.4	22.0

Table 3.3: Thresholding: The above tables give the compression ratio and PSNR of the images in Figures 3.4, 3.5 and 3.6. Bases $(-2+i, \{0,1,2,3,4\})$ and $(-4+i, \{0,\ldots,16\})$ are included for comparison.

Figure 3.7: Fractal-wavelet (LIFSW) coding with base $(-1+i, \{0,1\})$: The original 512×512 image of Lena is in the top lefthand corner. The other images, from top to bottom, left to right, result from choosing the domain level to be -13, with ranges $-11, -9, -7, -5, -3$.

Figure 3.8: Fractal-wavelet (LIFSW) coding with base $(-4+i, \{0, \ldots, 9\})$: The original 512×512 image of Lena is at the top. The other images, from top to bottom, left to right, result from choosing the domain level to be -5, with ranges $-4, -3, -2$ and -1.

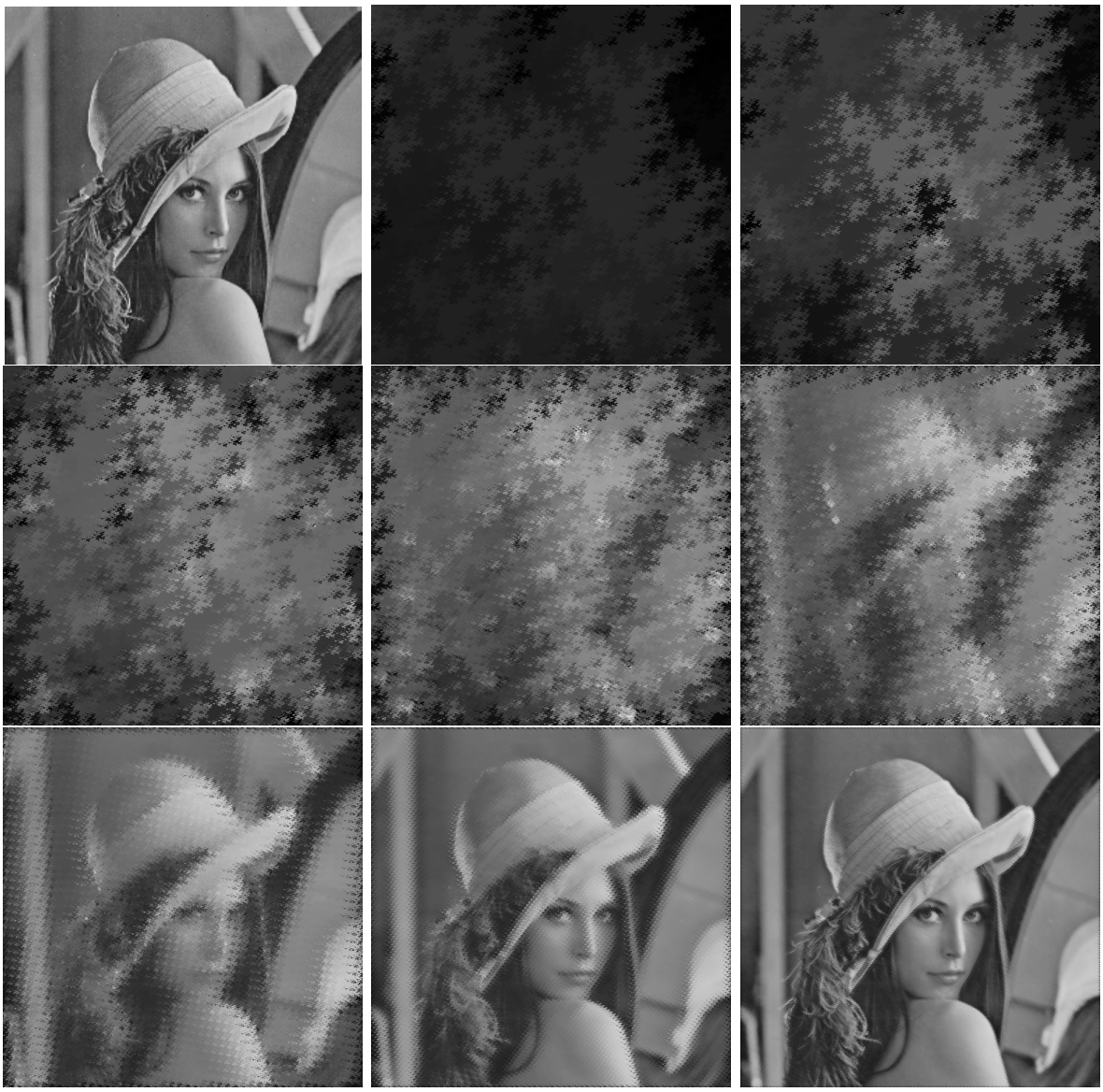


Figure 3.9: Fractal-wavelet (LIFSW) coding with base $(2+i, \{0, 1, i, -i, -2-3i\})$: The original 512×512 image of Lena is in the top lefthand corner. The other images, from top to bottom, left to right, result from choosing the domain level to be -9, with ranges taking all values between -8 and -1.

Base	Levels (d, r)				
$-1+i$	$(-13,-11)$	$(-13,-9)$	$(-13,-7)$	$(-13,-5)$	$(-13,-3)$
Compression	678.7	195.9	52.8	13.7	3.5
PSNR (dB)	19.1	21.8	23.9	27.6	31.5

Base	Levels (d, r)				
$-2+i$	$(-6,-5)$	$(-6,-4)$	$(-6,-3)$	$(-6,-2)$	$(-6,-1)$
Compression	303.4	80.9	18.6	4.0	0.8
PSNR (dB)	18.0	21.3	25.3	30.7	46.1

Base	Levels (d, r)				
$-3+i$	$(-6,-5)$	$(-6,-4)$	$(-6,-3)$	$(-6,-2)$	$(-6,-1)$
Compression	1068.9	301.8	55.9	7.3	0.8
PSNR (dB)	11.1	14.5	17.8	22.3	29.0

Base	Levels (d, r)			
$-4+i$	$(-5,-4)$	$(-5,-3)$	$(-5,-2)$	$(-5,-1)$
Compression	560.1	109.2	11.1	0.8
PSNR (dB)	10.5	14.3	18.8	25.4

Base	Levels (d, r)							
$2+i$	$(-9,-8)$	$(-9,-7)$	$(-9,-6)$	$(-9,-5)$	$(-9,-4)$	$(-9,-3)$	$(-9,-2)$	$(-9,-1)$
Compression	2621.4	1524.1	746.8	269.7	75.5	18.1	3.9	0.8
PSNR (dB)	8.8	11.0	13.3	15.3	17.5	20.0	23.4	27.8

Table 3.4: Fractal-wavelet (LIFSW) coding: The above tables give the compression ratio and PSNR of the images in Figures 3.7, 3.8 and 3.9 using domain-range level pairs $(-d, -r)$. Bases $(-2+i, \{0,1,2,3,4\})$ and $(-3+i, \{0,\ldots,9\})$ are included for comparison.

Example 3.3.5 Figure 3.10 on page 153 shows the result of performing the LIFSW algorithm on a few other standard images in image analysis: mandrill, boat, goldhill. These were compressed using the bases $(-2+i, \{0,1,2,3,4\})$, $(-3+i, 0, \ldots, 9)$ and $(2+i, \{0,1,i,-i,-2-3i\})$ from levels $(d,r) = (-3,-2)$, with compression ratios 4.0, 7.3 and 3.9 respectively. Table 3.5 gives the PSNR of each image. Visually, it seems like the fractal nature of the tiles works well for the mandrill image, better than for the other two images, even though the PSNR is lower. Notice also that as n increases, the bases $(-n+i, \{0, \ldots, n^2\})$ stretch out, causing difficulty for compression. The shape of the tile of $-3+i$ causes the disruption in the lines on the left-hand side of the nose of the mandrill, but less on the right.

Base	Mandrill	Boat	Goldhill
$-2+i$	20.9	24.5	25.6
$-3+i$	18.8	19.6	21.0
$2+i$	20.7	23.6	25.1

Table 3.5: PSNR values for LIFSW on mandrill, boat and goldhill.

Example 3.3.6 Figure 3.11 on page 154 shows the result of applying the LIFSW algorithm to various gradients using base $(-3+i, \{0, \ldots, 9\})$ with different values of (d,r). One notices again that the rotation of the tiles at each level d and r affects the quality of compression. Indeed, for the horizontal and vertical gradients, for levels $(-2,-1)$, the choice of domain block seems quite inappropriate. Further study revealed that even performing the LIFSW on levels $(d,-1)$, $d<-2$, showed proper behaviour in the sense that, for example, even the choice of $(-6,-1)$ gave a smooth looking image for the vertical gradient that looked like the original. Figure 3.12 on page 155 shows the vertical and horizontal gradients pruned at level -2. The edge effects from representing the bottom, and top and bottom of the gradient respectively, seems to explain this interesting behaviour. At level -2, the tiles are vertical, mono-colour strips of 10 pixels in length. The edge effects are generated by the need to take pixels from outside the supertile for the wavelet decomposition. This is

clearly demonstrated by the colour-to-length relationship. The larger the proportion of a tile is in the image, the lighter it is. The more a tile is outside the image, the darker it is. If one were to colour the supertile white, then the opposite effect would occur. Furthermore, the vertical tile would exhibit odd effects on the dark edge instead of the light. This type of behaviour prompts an interest in the extension algorithm.

Figure 3.10: Fractal-wavelet (LIFSW) coding with various bases from Example 3.3.5. The original 512 × 512 images are at the top.

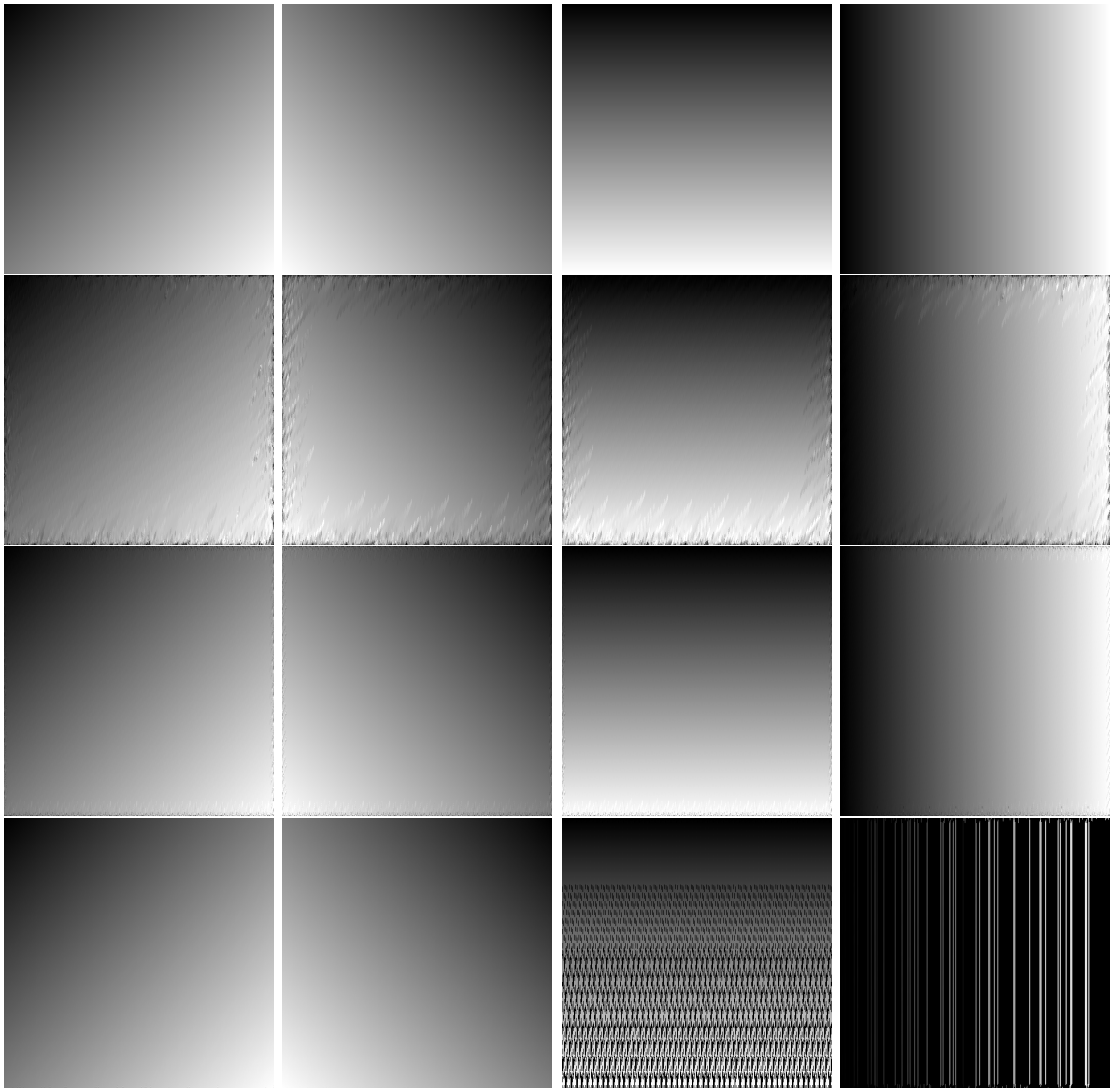


Figure 3.11: Fractal-wavelet (LIFSW) coding using base $(-3+i, \{0, \ldots, 9\})$ from levels $(-4, -3)$, $(-3, -2)$ and $(-2, -1)$ from Example 3.3.6. The original 512×512 gradient images are at the top.

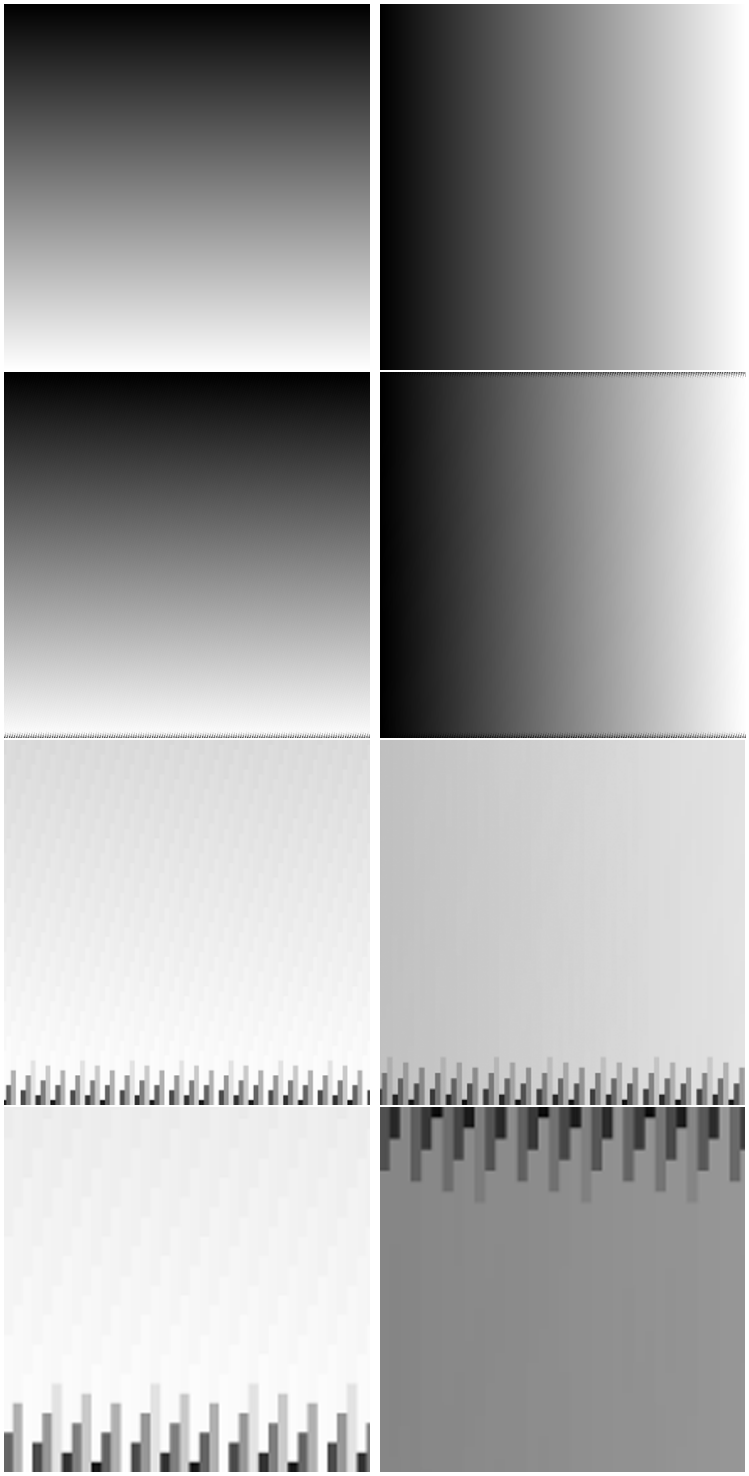


Figure 3.12: Vertical and horizontal gradients (top) from Example 3.3.6 pruned at level -2 in base $(-3+i, \{0, \ldots, 9\})$ (middle). The bottom images show closeups of the lower edge of each image. The edge effects are generated by the supertile.

Conclusion

In summary, a link between two-dimensional non-separable Haar wavelets and complex bases has been established, allowing the construction of an entire class of MRA in two dimensions. The translation of scaling and wavelet coefficients through this link has been shown to be consistent with the Mallat algorithm. Results regarding the termination of the Mallat algorithm for certain MRA have been proved. These results have been generalized to higher dimensional number systems, which retain the properties of complex bases. This includes an important equivalence result to establish whether or not a given matrix and digit set are indeed a valid base for a number system.

The asymmetry of the wavelet tree for generalized number systems was discussed. A class of extensions of functions was developed to address this situation. This algorithm was shown to be well defined and consistent with the Mallat algorithm. The theory of LIFSW was generalized to number systems and a method of constructing LIFSW approximations to functions was presented.

There seems to be a general consensus in the IFS research community that IFS, as presently used, cannot compete with state-of-the-art compression methods involving wavelets. However, this does not mean that fractal-based methods could not be used to enhance such methods (e.g. extrapolation of wavelet coefficients) or provide further insight

into the properties of given images. It is hoped that further research into this area will yield such successful applications.

Given the above results, a number of questions arise. They are listed here as potential problems for future study:

i) Would it be of interest to use the Long Division Algorithm to perform wavelet decompositions, or perhaps for LIFSW "zooming" (p. 42)?

ii) It would be extremely interesting to determine which results of this thesis can be generalized to non-Haar bases (p. 75). For instance, smooth wavelets yield a very rapid decay of the wavelet coefficients of a function (see [75] for a summary of such results). This would significantly improve the applicability of certain methods to image compression;

iii) Would it be useful to study addressing schemes for MRA where integer cyclic points exist (p. 80)? Would such schemes be well-defined?

iv) Can statistical analysis of the addresses of a complex base be used directly to place an image in an optimal location (p. 89)? Would an exhaustive searching scheme be of interest? Optimal means in the sense of minimizing the ratio between the number of points in the image and the number of points in the supertile that are outside the image. This might significantly improve any coding scheme using the methods of this thesis (p. 137);

v) Can an algorithm be constructed to modify a dilation and digit set associated with an MRA in such a way that the result is a valid base for a number system and such that

the resulting tile resembles the tile of the original MRA (p. 90)? Simple translations are not generally enough (see [37]);

vi) It has been shown that the fundamental tile of a valid base has non-empty interior (see [56]). Intuitively, this makes sense since if 0 is on the boundary of the fundamental tile, it would have multiple radix representations. It should be provable that the integer 0 is in the interior of every fundamental tile. In that case, an additional equivalence for a valid base should be that the only integer contained in the interior of the fundamental tile is 0 (p. 92). Such a result would make it easy to visually identify valid bases;

vii) In practice, what is the physical interpretation of scaling coefficients associated with non-zero cyclic points (p. 96)? A single terminating scaling coefficient is typically associated with the intensity of the signal;

viii) What coding schemes could be developed to enhance the LIFSW method for compression (p. 115)?

ix) Is it possible to solve Conjecture 2.3.13, or can one construct a non-diagonalizable matrix, which is a valid base for some digit set, where the matrix in not an acceptable dilation (p. 98)?

x) What could be some practical applications of the extension algorithm (p. 104) and the restricted LIFSW algorithm (p. 115)? One possibility comes from noticing that the fractal tiles of valid bases generate a variety of interesting textures in the compressed images. Such features might be of interest to the graphics design industry. For example, the LIFSW method with certain tiles seems to produce effects similar

to water colour paintings (Figure 3.10). The thresholding method produces effects similar to brush strokes on canvas (Figure 3.6). Could alternate extension methods be developed utilizing a more direct approach (p. 104)?

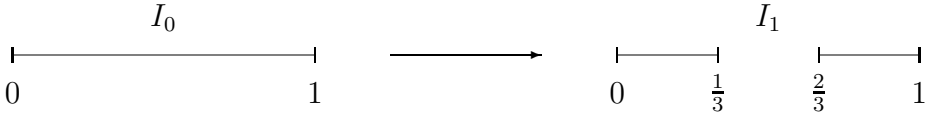
xi) What other operators M may be of interest for the extension algorithm, if indeed such operators exist? What other properties might such operators share (p. 109)?

xii) Is it possible to construct a tighter bound on M (p. 126)?

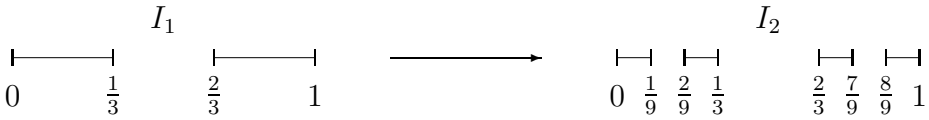
Appendix A

A.1 Motivation of IFS

To motivate the definition of IFS, consider the famous construction of the Cantor "middle-thirds" set [39, pp. 114-116]. Typically, the Cantor set is constructed by induction. Let $I_0 = I = [0,1] \subset \mathbb{R}$. Let $I_1 = I_0 \backslash (\frac{1}{3}, \frac{2}{3})$, that is, the interval $[0,1]$ with the open middle-third removed.



Construct I_2 from I_1 by removing the open middle-thirds from the two remaining closed intervals.



Inductively, construct I_{n+1} by removing the open middle-thirds from the 2^n closed intervals of I_n. The *Cantor set* is then defined to be $\mathcal{C} = \bigcap_{n=0}^{\infty} I_n$.

Suppose that you were asked to describe $\mathcal{C}$. At this point, it might be difficult without giving the argument for its construction. Returning to the ideas presented at the end of Section 1.2.1, we would like to find a function f, on some appropriate space, for which $\mathcal{C}$

is the attractor. This function could then be iterated to find $\mathcal{C}$ and would technically be an exact description of $\mathcal{C}$.

To understand what we wish to do in general, let us look at two characteristics of $\mathcal{C}$. By construction, $\mathcal{C}$ is compact. A second characteristic is the self-similarity we find within it. This is one of the reasons $\mathcal{C}$ is called a fractal. A general definition of self-similarity has been given in [58], but intuitively, an object is considered self-similar if parts of the object resemble others. An object is considered a fractal if this self-similarity is infinite. In essence, if you zoom into the object for a closer look, you cannot tell the difference from viewing the unzoomed object.

To see the self-similarity in $\mathcal{C}$, let $\mathcal{C}_1 = \mathcal{C} \cap [0, \frac{1}{3}]$ and let $\mathcal{C}_2 = \mathcal{C} \cap [\frac{2}{3}, 1]$. Intuitively, if we were to "zoom in" on $\mathcal{C}_1$ or $\mathcal{C}_2$, we could not distinguish either from $\mathcal{C}$. Mathematically, we see that the maps $w_1 : \mathcal{C} \to \mathcal{C}_1$ defined by $x \mapsto \frac{x}{3}$ and $w_2 : \mathcal{C} \to \mathcal{C}_2$ defined by $x \mapsto \frac{x}{3} + \frac{2}{3}$ are metric equivalences under the induced topology of $\mathbb{R}$. Indeed, $\mathcal{C}$ is the disjoint union of two metrically equivalent subsets:

$$\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2.$$

We wish $\mathcal{C}$ to be the fixed point of a certain function. We motivate the following definition by the fact that $\mathcal{C}$ is a subset of I.

Definition A.1.1 Let X and Y be sets and $f : X \to Y$. We define the set mapping $\hat{f}$: $P(X) \to P(Y)$ by

$$\hat{f}(A) = \{f(a) : a \in A\} \quad \forall A \in P(X),$$

where $P(X)$ denotes the power set of X.

We see that $\mathcal{C}_1 = \hat{w_1}(\mathcal{C})$ and $\mathcal{C}_2 = \hat{w_2}(\mathcal{C})$. Therefore,

$$\mathcal{C} = \hat{w}_1(\mathcal{C}) \cup \hat{w}_2(\mathcal{C}). \tag{A.1}$$

Hence, $\mathcal{C}$ can be written as a union of contracted copies of itself. This is the central theme of IFS. Given a set A, try to write A as a union of contracted copies of itself.

Definition A.1.2 Let X and Y be sets and $f_\lambda : X \to Y$, $\lambda \in \mathcal{A}$, where $\mathcal{A}$ is some indexing set. Let $\mathbf{f} = \{f_\lambda\}$. We define $\hat{\mathbf{f}} = \cup_{\lambda\in\mathcal{A}} \hat{f}_\lambda$, that is for $A \subset X$, we have

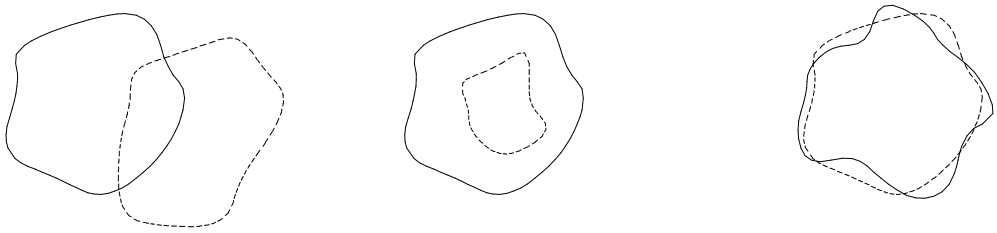
$$\hat{\mathbf{f}}(A) = \bigcup_{\lambda\in\mathcal{A}} \hat{f}_\lambda(A).$$

If we now set $\mathbf{w} = \{w_1, w_2\}$, we see by Equation (A.1), that $\mathcal{C}$ is the fixed point of $\hat{\mathbf{w}}$. Now, for this $\mathbf{w}$ to be useful, we would need $\mathcal{C}$ to be its attractor in some appropriate space. Intuitively, this makes sense since $I_{n+1} = \hat{\mathbf{w}}(I_n)$, for each $n \in \mathbb{N}$. Hence, in a way, $\mathbf{w}$ is an exact description of $\mathcal{C}$. In general, the desired approximations would be obtained by iterating maps of the form given in Definition A.1.2.

A.2 A Complete Space for IFS

We will use the example of $\mathcal{C} \subset \mathbb{R}$ to motivate the search for a space consisting of subsets of a complete metric space. Given a complete space (X, d), the goal is to find a complete space (Y, d_Y) with $Y \subset P(X)$. We will first construct a distance function d_Y on $P(X)$ and use the conditions needed for it to be a metric to help us determine Y. To begin, consider the three pairs of sets in $\mathbb{R}^2$ pictured in Figure A.1 on the next page.

Each case consists of two sets, one bounded by the solid line and one by the dashed line. In which case do the two sets seem "closest"? Probably not in (a). In (b), the dashed set certainly seems close to the solid one; it is part of the solid set. However, many points



(a) (b) (c)

Figure A.1: Closeness of sets.

of the solid set seem distant from the dashed set. Case (c) seems intuitively right. Most points of the solid set are close to the dashed one, and vice-versa. More precisely, each set overlaps the other set rather well. We use these ideas to begin to construct our metric.

Unless otherwise specified, (X, d) will denote a metric space with no other properties.

Notation A.2.1 Let $x \in X, B \subset X$. Define the distance from x to B by

$$d(x, B) = \inf_{b \in B} d(x, b).$$

Hence, if $x \in B$, $d(x, B) = 0$.

Notation A.2.2 Let $A, B \subset X$. Define the distance from A to B by

$$d(A, B) = \sup_{a \in A} d(a, B).$$

This seems reasonable since if $A \subset B$, A should be close to B and by this definition we would have $d(A, B) = 0$. Unfortunately, this function is not a metric. For example, $d([0, \frac{1}{2}], [\frac{1}{3}, 1]) = \frac{1}{3}$ but $d([\frac{1}{3}, 1], [0, \frac{1}{2}]) = \frac{1}{2}$. Symmetry is lacking, which motivates the following construction [4, 23, 43]:

Definition A.2.3 Let $A, B \subset X$. Define the *Hausdorff distance* between A and B by

$$h(A, B) = \max\{d(A, B), d(B, A)\}.$$

This is the function d_Y we seek. It satisfies the intuitive notion of two sets being close, which was found in (c) on page 164. The function h is almost a metric. We use the following table to help rule out certain sets from $P(X)$:

Problem: $h(\emptyset, [0,1]) = ?$	Solution: Only consider non-$\emptyset$ sets.
Problem: $h([0,1), [0,1]) = 0$	Solution: Only consider closed sets.
Problem: $h([0,1], [0,\infty)) = \infty$	Solution: The closed sets must be compact.

Notation A.2.4 Define $\mathcal{H}(X)$ to be the set of all non-empty, compact subsets of X.

Theorem A.2.5 Let (X, d) be a metric space. Then $(\mathcal{H}(X), h)$ is a metric space. Furthermore, if (X, d) is complete then so is $(\mathcal{H}(X), h)$.

We now note a few properties about the Hausdorff metric which enable us to justify the definition of contractivity of an iterated function system as seen in Definition 1.2.19. Again, let (X, d) be a metric space.

Notation A.2.6 Let $Con(X, d, s)$ denote the set of all contractive maps with contractivity at least s.

Lemma A.2.7 Let $w \in Con(X, d, s)$, then $\hat{w} \in Con(\mathcal{H}(X), h, s)$.

Proposition A.2.8 Let (X, d) be a metric space and let

$$\mathbf{w} = \{w_n \in Con(X, d, c_n) : n = 1, 2, \ldots, N\}.$$

Then $\hat{\mathbf{w}} \in Con(\mathcal{H}(X), h, c)$, where $c = \max\{c_n : n = 1, 2, \ldots, N\}$.

Proposition A.2.8 implies a crucial result for IFS.

Theorem A.2.9 (BCMP for IFS) Let $\mathbf{w}$ be an N-map IFS with contractivity c. Then $\hat{\mathbf{w}} \in Con(\mathcal{H}(X), h, c)$. Furthermore, $\hat{\mathbf{w}}$ has a unique fixed point $A_{\hat{\mathbf{w}}} \in \mathcal{H}(X)$ which is also its attractor.

Definition A.2.10 The fixed point of $\hat{\mathbf{w}}$ is called the *attractor* of $\hat{\mathbf{w}}$.

This yields the following version of Proposition 1.2.17 for iterated function systems [4]:

Theorem A.2.11 (The Collage Theorem) Let $\mathbf{w}$ be an N-map IFS with contractivity $0 \leq c < 1$. Suppose $L \in \mathcal{H}(X)$ and $\epsilon > 0$ are such that $h(L, \hat{\mathbf{w}}(L)) \leq \epsilon$. Then $h(L, A_{\hat{\mathbf{w}}}) \leq \frac{\epsilon}{1-c}$.

The distance $h(L, \hat{\mathbf{w}}(L))$ is often called the *collage distance*. The Collage Theorem is important for the Inverse Problem of approximating sets seen in Section 1.2.1. By the Collage Theorem, one could try to construct an IFS $\mathbf{w}$ which takes L close to itself. The attractor of $\mathbf{w}$ would then be close to L.

It is possible that $c \approx 1$ which, in turn, implies that the constant $\epsilon/(1-c)$ can be large. Thus there is no guarantee that the collage distance is small and the approximation may be quite poor. To make $c \approx 0$, one can use maps with small contractivity factors. However, this might increase the number of maps needed to describe the approximation (hence reducing the compression). This fact is relevant when compression is a principal factor.

In order to calculate fractal images using the theoretical machinery that has been developed, one can use the following algorithm, a consequence of Theorem A.2.9 [4]:

Corollary A.2.12 (The Deterministic Algorithm) Let $\mathbf{w}$ be an N-map IFS with $\mathbf{w} = \{w_j : j = 1, 2, \ldots, N\}$. Let $A_0 \in \mathcal{H}(X)$. Compute $A_n = \hat{\mathbf{w}}^{\circ n}(A)$ by $A_{n+1} = \bigcup_{j=1}^{n} \hat{w}_n(A_n)$

for $n = 1, 2, \ldots$. Then the sequence $(A_n) \subset \mathcal{H}(X)$ converges to the attractor of the IFS in $\mathcal{H}(X)$.

A.3 Examples of IFS Attractors

In practice, affine IFS contraction maps are used to simplify calculations. Let $M_n(\mathbb{R})$ denote the set of all $n \times n$ matrices over $\mathbb{R}$ and $\mathbb{R}^n$ be the usual Euclidean n-space.

Definition A.3.1 Let $X \subset \mathbb{R}^n$, $n \in \mathbb{N}^+$. A map $w : X \to \mathbb{R}^n$ is called an *affine transformation* if $\exists A \in M_n(\mathbb{R})$ and $b \in \mathbb{R}^n$ such that

$$w(x) = Ax + b \quad \forall x \in X.$$

In general, given vector spaces X and Y, an *affine transformation* $f : X \to Y$ is a map of the form

$$f(x) = Ax + b$$

where A is a linear transformation from X to Y and $b \in Y$.

Example A.3.2 Let $X = [0, 1]$ and let $w_i(x) = \frac{1}{3}(x + 2i), i = 0, 1$. Then $A_{\hat{\mathbf{w}}} = \mathcal{C}$.

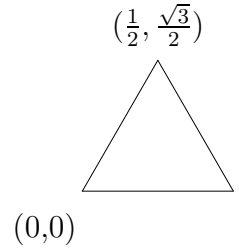
Example A.3.3 Let $X = [0, 1]^2$. Define the following maps:

$$w_1(x, y) = \left(\frac{x}{2}, \frac{y}{2}\right),$$
$$w_2(x, y) = \left(\frac{x}{2} + \frac{1}{2}, \frac{y}{2}\right),$$

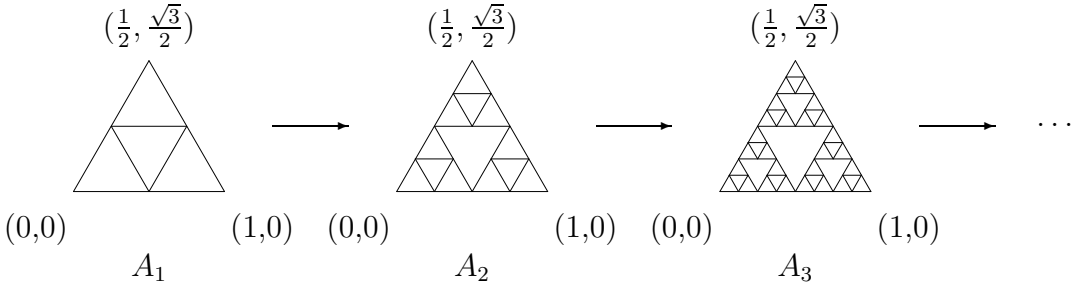
and

$$w_3(x, y) = \left(\frac{x}{2} + \frac{1}{4}, \frac{y}{2} + \frac{\sqrt{3}}{4}\right).$$

To find the attractor of $\hat{\mathbf{w}}$, we use the Deterministic Algorithm. We are allowed to make any choice of A_0. Therefore, let A_0 be the following triangle:



Then, using the algorithm, we obtain the following sequence of sets:



This sequence converges to the Sierpinski gasket [63].

An *affine IFS* $\mathbf{w} = \{w_i\}$ is an IFS where each w_i is affine. Often, affine IFS in $\mathbb{R}^2$ will be written in a table to facilitate their description. Consider an IFS consisting of the maps

$$w_i(x, y) = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e_i \\ f_i \end{pmatrix}, \quad i = 1, 2, \ldots, N.$$

Instead of writing them as above, they are written in a table such as:

$$\begin{array}{cccccc} a_1 & b_1 & c_1 & d_1 & e_1 & f_1 \\ a_2 & b_2 & c_2 & d_2 & e_2 & f_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_N & b_N & c_N & d_N & e_N & f_N \end{array}$$

We now recall the definition of a similitude.

Definition A.3.4 A transformation $w : \mathbb{R}^2 \to \mathbb{R}^2$ is called a *similitude* if it is an affine transformation of the form

$$w(x, y) = r \begin{pmatrix} \cos\theta & \pm\sin\theta \\ \sin\theta & \mp\cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$$

where $(e, f) \in \mathbb{R}^2, r \neq 0, \theta \in [0, 2\pi)$. The constant r is called the *scaling* of w, or its *scaling factor*, and θ is called its *angle of rotation*.

Proposition A.3.5 If $w(x) = Ax + b, A \in M_2,\ b, x \in \mathbb{R}^2$ is a similitude in $\mathbb{R}^2$, then its contractivity factor is $|\det A|$.

Corollary A.3.6 If $w : \mathbb{R}^2 \to \mathbb{R}^2$ is a similitude as above and $|\det A| < 1$, then $w \in Con(X, d)$.

To apply this theory to images, i.e. computer images, one can think of an *image* as being a compact subset of $\mathbb{R}^n$. One can model a computer screen by $X = [0, 1]^2$ and define an image on the screen to be a set A in X, with points being screen pixels. If $x \in A$, the associated pixel is plotted white. If $x \notin A$, leave the pixel black. Hence a white screen represents $A = [0, 1]^2$ (see 3.1).

Suppose an IFS acts on the screen. When the IFS is iterated, the points of A move about the screen. Looking at $\hat{\mathbf{w}}(A)$, we see that $x \in \hat{\mathbf{w}}(A)$ if $\exists i \in 1, 2, \ldots, N$ such that $x = w_i(y)$ for some $y \in A$. Hence, after one iteration of $\hat{\mathbf{w}}$, a pixel is plotted white if there is a white pixel mapped to it by the IFS. We can therefore think of IFS as mapping black and white images to black and white images.

Unfortunately, as they say, the world is not black and white. What is needed is an IFS-type method which allows the pixels to take on grey-level values. We also need maps which move pixels around and then scale their grey-levels. This is the theory of IFSM [28] (see Definition 1.2.21 on page 16).

Bibliography

[1] R. B. Ash. *Measure, Integration, and Functional Analysis.* Academic Press, New York, 1972.

[2] G. Bachman and L. Narici. *Functional Analysis.* Academic Press, 1966.

[3] C. Bandt. Self-similar sets 5. Integer matrices and fractal tilings of $\mathbb{R}^n$. *Proc. Amer. Math. Soc.*, 112:549–562, 1991.

[4] M. F. Barnsley. *Fractals Everywhere.* Academic Press, Boston, second edition, 1993.

[5] M. F. Barnsley and S. Demko. Iterated function systems and the global construction of fractals. *Proc. R. Soc. Lond. A*, 399:243–275, 1985.

[6] M. F. Barnsley and L. Hurd. *Fractal Image Compression.* A K Peters, 1993.

[7] M. F. Barnsley, V. Ervin, D. Hardin and J. Lancaster. Solution of an inverse problem for fractals and other sets. In *Proceedings of the National Academy of Science*, volume 83, pages 1975–1977, 1986.

[8] G. R. Belitskii and Y. I. Lyubich. *Matrix Norms and Their Applications.* Birkhauser, Boston, 1988.

[9] J. J. Benedetto and M. W. Frazier, editors. *Wavelets: Mathematics and Applications*. Studies in Advanced Mathematics. CRC Press, 1994.

[10] G. Beylkin. On the representation of operators in bases of compactly supported wavelets. *SIAM J. Numer. Anal.*, 29(6):1716–1740, 1992.

[11] B. Bollobás. *Linear Analysis: An Introductory Course*. Cambridge University Press, 1992.

[12] J. Burke. *Connections*. Macmillan, London, 1978.

[13] J. Burke. *The Day the Universe Changed*. Little, Brown & Company, 1986.

[14] J. Burke. *The Axemaker's Gift: A Double-Edged History of Human Culture*. G.P. Putnam's Sons, 1995.

[15] C. S. Burrus, R. A. Gopinath and H. Guo. *Introduction to Wavelets and Wavelet Transforms, A Primer*. Prentice Hall, 1998.

[16] C. Cabrelli, U. Molter and E. R. Vrscay. Recurrent iterated function systems: Invariant measures, a collage theorem and moment relations. In H.-O. Peitgen, J. M. Henriques and L. F. Penedo, editors, *Fractals in the Fundamental and Applied Sciences*, pages 71–80, North-Holland, Amsterdam, 1991.

[17] P. M. Centore and E. R. Vrscay. Continuity of attractors and invariant measures for iterated function systems. *Canadian Mathematical Bulletin*, 37(3):315–329, 1994.

[18] J. B. Conway. *A Course in Functional Analysis*. Springer, New York, second edition, 1997.

[19] X. Dai and S. Lu. An MRA yields an orthonormal wavelet basis - an operator-theoretic proof. *Appl. Math. J. Chinese Univ. Ser. B*, 11(1):77–84, 1996.

[20] I. Daubechies. *Ten Lectures on Wavelets*. CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, 1992.

[21] M. Davio, J. P. Deschamps and C. Gossart. *Complex Arithmetic*. Philips MBLE Research Lab. Report R369, May 1978.

[22] N. Dunford and J. T. Schwartz. *Linear Operators Part I: General Theory*. John Wiley and Sons, New York, third edition, 1996.

[23] K. J. Falconer. *The Geometry of Fractal Sets*. Cambridge University Press, Cambridge, U. K., 1985.

[24] F. Falzon and S. Mallat. Analysis of low bit rate image transform coding. *IEEE Trans. Signal Proc.*, 46(4):944–970, April 1998.

[25] Y. Fisher, editor. *Fractal Image Compression*. Springer-Verlag, 1995.

[26] Y. Fisher, editor. *Fractal Image Encoding and Analysis*. NATO ASI Series. Springer-Verlag, Heidelberg, 1998.

[27] B. Forte and E. R. Vrscay. Inverse problem methods for generalized fractal transforms. In Y. Fisher, editor, *Fractal Image Encoding and Analysis*, Proceedings of the NATO Advanced Study Institute, Trodheim, Norway, July 8-17, 1995.

[28] B. Forte and E. R. Vrscay. Solving the inverse problem for function and image ap-

proximation using iterated function systems. *Dynamics of Continuous and Impulsive Systems*, 1:177–231, 1995.

[29] B. Forte and E. R. Vrscay. Theory of generalized fractal transforms. In Y. Fisher, editor, *Fractal Image Encoding and Analysis*, Proceedings of the NATO Advanced Study Institute, Trodheim, Norway, July 8-17, 1995.

[30] W. Gilbert. Radix representations of quadratic fields. *J. Math. Anal. Appl.*, 83:264–274, 1981.

[31] W. Gilbert. Fractal geometry derived from complex bases. *Math. Intelligencer*, 4(2):78–86, 1982.

[32] W. Gilbert. Geometry of radix representations. In C. Davis, B. Grünbaum and F. A. Sherk, editors, *The Geometric Vein, The Coxeter Festschrift*, pages 129–139. Springer-Verlag, 1982.

[33] W. Gilbert. Arithmetic in complex bases. *Mathematics Magazine*, 57(2):77–81, 1984.

[34] W. Gilbert. Complex based number systems. Unpublished, May 1994.

[35] W. Gilbert. The division algorithm in complex bases. *Canadian Mathematical Bulletin*, 39(1):47–54, 1996.

[36] K. Gröchenig and A. Haas. Self-similar lattice tilings. *J. Fourier Anal. Appl.*, 1(2):131–170, 1994.

[37] K. Gröchenig and W. R. Madych. Multiresolution analysis, haar bases, and self-similar tilings of R^n. *IEEE Trans. Inform. Theory*, 39:556–568, 1992.

[38] B. Grünbaum and G. C. Shephard. *Tilings and Patterns*. W.H. Freeman and Company, New York, 1986.

[39] D. Gulick. *Encounters with Chaos*. McGraw Hill, 1992.

[40] C. Handy and G. Mantica. Inverse problems in fractal construction: Moment method solution. *Physica*, D43:17–36, 1990.

[41] C. E. Heil and D. F. Walnut. Continuous and discrete wavelet transforms. *SIAM Review*, 31:628–666, 1989.

[42] A. S. Householder. *The Theory of Matrices in Numerical Analysis*. Blaisdell Publishing Co., 1964.

[43] J. E. Hutchinson. Fractals and self-similarity. *Indiana University Journal of Mathematics*, 30:713–747, 1981.

[44] J. J. Hwang and K. R. Rao. *Techniques and Standards for Image, Video and Audio Coding*. Prentice Hall, 1996.

[45] N. Jacobson. *Basic Algebra*, volume I. W. H. Freeman, New York, second edition, 1985.

[46] S. Jaffard. Multifractal formalism for functions part I: Results valid for all functions. *SIAM J. Math. Anal.*, 28(4):944–970, 1997.

[47] S. Jaffard. Multifractal formalism for functions part II: Self-similar functions. *SIAM J. Math. Anal.*, 28(4):971–998, 1997.

[48] Q. Jiang and L. Peng. Toeplitz type operators on wavelet subspaces. *J. Math. Anal. Appl.*, 207(2):462–474, 1997.

[49] R. H. Kasriel. *Undergraduate Topology*. Saunders, 1986.

[50] I. Kátai and B. Kovács. Canonical number systems in imaginary quadratic fields. *Acta Math. Acad. Sci. Hungaricae*, 37:159–164, 1981.

[51] I. Kátai and J. Szabó. Canonical number systems for complex integers. *Acta. Sci. Math. (Szeged)*, 37:255–260, 1975.

[52] R. Kenyon. Self-replicating tilings. In P. Walters, editor, *Symbolic Dynamics and its Applications*, Contemporary Mathematics, pages 239–263. Amer. Math. Soc., 1992.

[53] R. Kenyon, J. Li, R. Strichartz and Y. Wang. Geometry of self-affine tiles II. *Indiana J. Math.*, 48(1):24–42, 1999.

[54] J. Kovačević and M. Vetterli. Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for $\mathbb{R}^n$. *IEEE Trans. Inform. Th.*, 38(2):533–555, 1992.

[55] J. C. Lagarias and Y. Wang. Radix expansions for the orthant $\mathbb{R}^n$. *Trans. of AMS*, 348:99–117, 1996.

[56] J. C. Lagarias and Y. Wang. Self-affine tiles in $\mathbb{R}^n$. *Advances in Math.*, 121(1):21–49, 1996.

[57] G. Lemarié and Y. Meyer. Ondelettes et bases hilbertiennes. *Revista Matematica Iberoamericana*, 2:1–18, 1986.

[58] G. B. Lewellen. Self-similarity. *Rocky Mountain Journal of Mathematics*, 23(3):1023–1040, Summer 1993.

[59] N. Lu. *Fractal Imaging*. Academic Press, New York, 1997.

[60] W. R. Madych. *Multiresolution Analyses, Tiles, and Scaling Functions*, pages 233–243. Probabilistic and stochastic methods in analysis, with applications. Kluwer Acad. Publ., 1992.

[61] S. Mallat. Multiresolution approximation and wavelet orthonormal bases of $L^2(\mathbb{R})$. *Trans. AMS*, 315:69–88, 1989.

[62] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, London, second edition, 1999.

[63] B. Mandelbrot. *The Fractal Geometry of Nature*. W. H. Freeman and Company, New York, 1977.

[64] P. R. Massopust. *Fractal Functions, Fractal Surfaces, and Wavelets*. Academic Press, San Diego, 1994.

[65] F. Mendivil. The application of a fast non-separable discrete periodic wavelet transform to fractal image compression. *Proc. Fractals in Engineering*, pages 97–106, June 1999.

[66] F. Mendivil and D. G. Piché. Two algorithms for non-separable wavelet transforms and applications to image compression. In M. Dekking, J. Lévy Véhel, E. Lutton and C. Tricot, editors, *Fractals: Theory and Applications in Engineering*, pages 325–345. Springer Verlag, 1999.

[67] F. Mendivil and E. R. Vrscay. Correspondence between fractal-wavelet transforms and iterated function systems with grey level maps. In E. Lutton, C. Tricot and J. Lévy Véhel, editors, *L'Ingénieur et les fractales*, pages 54–64. Springer Verlag, 1997.

[68] Y. Meyer. *Ondelettes, fonctions splines et analyses graduées*. Rapport CEREMADE n.8703, 1987.

[69] Y. Meyer. *Ondelettes et opérateurs I*. Herman, Paris, 1990.

[70] Y. Meyer. *Wavelets and Operators*. Cambridge studies in advanced mathematics. Cambridge University Press, 1992.

[71] A. Mukherjea and K. Pothoven. *Real and Functional Analysis Part A: Real Analysis*. Plenum Press, New York, second edition, 1984.

[72] A. Mukherjea and K. Pothoven. *Real and Functional Analysis Part B: Functional Analysis*. Plenum Press, New York, second edition, 1984.

[73] W. B. Pennebaker and J. L. Mitchell. *JPEG Still Image Data Compression Standard*. Van Nostrand Reinhold, New York, 1993.

[74] W. Penny. A "binary" system for complex numbers. *J. Assoc. Comp. Mach.*, 12(2):247–248, 1965.

[75] D. G. Piché. IFSM, Wavelets and Fractal-Wavelets: Three Methods of Approximation. Master's thesis, University of Waterloo, Department of Pure Mathematics, May 1997.

[76] D. G. Piché. Wavelet representation and compression of images on fractal tilings. *Proc. Fractals in Engineering*, pages 82–96, June 1999.

[77] H. L. Royden. *Real Analysis*. Macmillan, New York, third edition, 1988.

[78] W. Rudin. *Fourier Analysis on Groups*. Interscience Publishers, New York, 1962.

[79] W. Rudin. *Functional Analysis*. McGraw-Hill, 1973.

[80] A. Said and W. A. Pearlman. A new, fast and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans. Circuits and Systems for Video Tech.*, 6(3):243–250, June 1996.

[81] R. Strichartz and Y. Wang. Geometry of self-affine tiles I. *Indiana J. Math.*, 48(1):1–23, 1999.

[82] P. Topiwala, editor. *Wavelet Image and Video Compression*. Kluwer, Boston, 1998.

[83] E. R. Vrscay. Moment and collage methods for the inverse problem of fractal construction with iterated function systems. In H.-O. Peitgen, J. M. Henriques and L. F. Penedo, editors, *Fractals in the Fundamental and Applied Sciences*, pages 443–461, North-Holland, Amsterdam, 1991.

[84] E. R. Vrscay. Iterated function systems: Theory, applications and the inverse problem, Spring 1995. A course in Fractal Image Compression, University of Waterloo.

[85] E. R. Vrscay. Mathematical theory of generalized fractal transforms and associated inverse problems. In *ImageTech Conference on the Mathematics of Imaging and its Applications*, March 1996.

[86] E. R. Vrscay. A new class of fractal-wavelet transforms for image representation and compression. *Can. J. Elect. Comp. Eng.*, 23:69–83, 1998.

[87] G. G. Walter. *Wavelets and Other Orthogonal Systems with Applications*. CRC Press, 1994.

[88] B. A. Wandell. *Foundations of Vision: Behavior, Neuroscience and Computation*. Sinauer Associates, Inc., Massachusetts, 1995.

[89] R. L. Wheeden and A. Zygmund. *Measure and Integral: An Introduction to Real Analysis*. Marcel Dekker, Inc., 1977.

[90] S. Willard. *General Topology*. Addison-Wesley, 1970.

Glossary

(A, K), **54**
(B, D), **90**
$(C_w(u, k^*), d_w)$, **70**
$(\mathcal{H}(X), h)$, **165**
(X, d), **8**
$(a_t a_{t-1} \cdots a_1 a_0)_b$, **36**
$(a_t a_{t-1} \cdots a_1 a_0 \,.\, a_{-1} a_{-2} \cdots)_b$, **36**
(b, D), **35**
$(\mathbf{w}, \mathbf{\Phi})$, **16**
$(\mathbf{w}_{loc}, \mathbf{\Phi})$, **30**
$(d_{t-1} \ldots d_1 d_0)_{(b,D)}$, **83**
$(d_t d_{t-1} \cdots d_1 d_0)_B$, **90**
(q_i), **62**

$\mathcal{A}$, **163**
$A_{\hat{\mathbf{w}}}$, **166**
A_L, **101**
$\bar{\alpha}^l_{\sigma\theta}$, **127**
$\alpha^l_{\sigma\theta}$, **116**
A_t, **100**
$\tilde{A}$, **101**
a^t_z, **84**
$a^t_{z,(b,D)}$, **83**
A_X, **85**
a_z, **84**

$B(X)$, **20**
$b^{\mathbf{c}}$, **70**
β, **39**

$\mathbb{C}$, **8**
$\mathcal{C}$, **161**
$C(D, p)$, **20**
c_f, **9**
$\mathrm{child}(t, T')$, **104**
χ_Q, **50**
$\circ$, **10**
$C_{loc}(D, p)$, **30**
$Con(\mathcal{H}(X), h, s)$, 165
$Con(\mathbb{R}^n)$, **9**
$Con(X, d)$, **9**
$Con(X, d, s)$, **165**
$Con_1(X, d)$, **15**
C_σ, **116**
$C_w(u, k^*)$, **70**

d, **8**
$d(A, B)$, **164**
$d^N_{\mathbf{\Phi}}(\mathbf{\Phi}_1, \mathbf{\Phi}_2)$, **21**
$d(x, B)$, **164**
$d_2(u, v)$, **130**
Δ, **24**, **126**
Δ_j, **70**
$\Delta_{l,j}$, **73**
$\Delta^{\min}_{j,l}$, **74**
$\Delta_{\sigma,\theta,\omega^l_{\sigma\theta}}$, **126**
$\bar{\Delta}_{\sigma,\theta,\omega^l_{\sigma\theta}}$, **127**
$\det A$, 169
$d_m(f, g)$, **12**
$D^l_{\sigma\theta}$, **116**
D^t, **86**
D_v, **32**
$d_w(\mathbf{c}, \mathbf{d})$, **70**

e_j, **45**
$E^l_{\sigma\theta}(g)$, **116**

F, **12**, **63**
$\mathbf{f}$, **163**
$\mathcal{F}(X)$, **16**
$\hat{\mathbf{f}}$, **163**
$\hat{f}$, **162**
F_i, **79**
f_i, **46**
$f_{i,j}$, **46**
$f_{k,p}$, **68**
$f^{\circ n}$, **10**
f_σ, **115**
$\tilde{f}$, **104**

Γ, 47
$G_{i,j}$, **79**
g^l_j, **48**

$h(A, B)$, **165**
$\mathcal{H}(X)$, **165**
h_j, **47**

K, **53**

L, **85**
$\ell^2(\mathbb{Z}^n)$, **54**
λ, **45**
$L^p_A(X, \mu)$, **25**
$Lip(\mathbb{R}^n)$, **9**
$L^p(X, \mu)$, 18
$L^2(\mathbb{R}^n)$, **7**
$L_{X,(b,D)}$, **85**

M, **63**
$m^{(D)}$, **20**
$M(X)$, **20**
M_2, 169
$M_n(\mathbb{R})$, **167**
$M_n(\mathbb{Z})$, **90**

$\mathbb{N}$, **8**
$\mathbb{N}^+$, **8**

$\bigoplus$, **48**

$P(X)$, **162**
ϕ, **47**
$\Phi^r(z)$, **41**
$\mathbf{\Phi}$, **16**
ψ^l, **46**, 48
$\psi^l_{i,j}$, 46
$\mathrm{PSNR}(u, v)$, **131**

Q, **53**
q, **45**
$Q_{a^L_z}$, **111**
Q_σ, **111**
$\tilde{Q}$, **101**
$Q_{\tilde{X}}$, **101**
Q_X, **101**
Q_z, **101**, 111
$Q \simeq R$, **50**

$\mathbb{R}$, **8**
$\mathrm{RemSet}(b, D, w)$, 42
$\rho(L)$, **98**
$\mathbb{R}^n$, **167**

s, **9**
$s_{-l,(d_{L-1}\ldots d_l)}$, **86**
$S_{a,b,c,d}$, **73**
S_f, **104**
$\mathrm{sib}(t, T')$, **104**
σ, **86**
${}_l\sigma$, **86**
$\sigma\rho$, **86**
$Sim(X, d)$, **15**
$Sim_1(X, d)$, **15**
$s_{m,n}$, **58**
$S^l_{\sigma,a,\gamma,b,\omega}$, **127**

Abbreviations

BCMP
: Banach Contraction Mapping Principle, **11**

IFS
: Iterated function system, **14**

IFSC
: IFS on coefficients, **64**

IFSM
: IFS with grey-level maps, **16**

LIFSM
: Local IFSM, **29**

LIFSW
: Local IFS on wavelets, **67**, **115**

MRA
: Multiresolution analysis, **47**

μ-d-n
: μ-dense and non-overlapping, **22**

PSNR
: Peak-signal-to-noise ratio, **131**

QP
: Quadratic programming, **26**

RMSE
: Root-mean-squared error, **130**

Index