OVERCOMPLETE EXPANSIONS FOR DIGITAL SIGNAL PROCESSING

by

Zoran Cvetković

Memorandum No. UCB/ERL M95/114

18 December 1995

OVERCOMPLETE EXPANSIONS FOR DIGITAL SIGNAL PROCESSING

by

Zoran Cvetković

Memorandum No. UCB/ERL M95/114

18 December 1995

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Abstract

OVERCOMPLETE EXPANSIONS FOR DIGITAL SIGNAL PROCESSING

by

Zoran Cvetković

Doctor of Philosophy in Engineering-Electrical Engineering and Computer Sciences

University of California at Berkeley

Professor Martin Vetterli, Chair

The theory of signal expansions into time-frequency localized atoms was one of the major breakthroughs in signal processing during the past decades. In the continuous-time domain, $L^2(\mathbf{R})$, the focus was on representations based on Weyl-Heisenberg and wavelet bases as well as frames which are their overcomplete counterparts. In parallel, these expansions were studied in the discrete-time domain, $\ell^2(\mathbf{Z})$, in the framework of filter banks and subband coding schemes. This research has been confined mainly to orthonormal and biorthonormal bases which are equivalent to critically sampled filter banks. This thesis is concerned with oversampled filter banks. The motivation stems from applications in which critical sampling or orthogonality is not needed and even imposes restrictive design constraints, and also from applications in which redundant representations are sought per se.

The general issues related to filter banks addressed in this thesis are 1) the necessary and sufficient conditions on filter banks for implementing frame or tight frame decompositions in $\ell^2(\mathbf{Z})$; 2) the feasibility of perfect reconstruction using FIR synthesis filters following an FIR analysis; 3) the parameterization and design of interesting

classes of perfect reconstruction oversampled filter banks; 4) the relation of filter banks to continuous-time signal analysis.

Nonsubsampled filter banks and modulated filter banks are two important particular classes which are studied in more detail. Nonsubsampled filter banks provide a tool for the discrete-time implementation of continuous-time transforms. They allow for a very flexible design and this is illustrated here by a procedure for designing maximally flat two-channel filter banks which yield highly regular wavelets with a given number of vanishing moments. The study of modulated filter banks is confined here to the paraunitary case, which is equivalent to tight Weyl-Heisenberg frames in $\ell^2(\mathbf{Z})$. We give a complete parameterization of these filter banks.

Robustness of overcomplete expansions to quantization is also investigated. A heuristic error analysis, based on the white noise error model, indicates that the mean squared quantization error (MSE) is inversely proportional to the frame redundancy, MSE $= O(1/R)$. This phenomenon has been exploited for a long time in oversampled A/D conversion for improving accuracy of the conversion. One of the results of this thesis is a deterministic analysis of error in oversampled A/D conversion which shows that the classical linear reconstruction is suboptimal and that the quantization error can be reduced in the squared norm as $O(1/R^2)$. Besides, even with the optimal reconstruction, the MSE of the oversampled A/D conversion, as a function of the bit-rate B, behaves as $O(1/B^2)$. We demonstrate that the classical coding used in oversampled A/D conversion is inefficient and propose a coding scheme which asymptotically attains $O(2^{-\beta B})$ error-rate characteristic. As a generalization, we give a deterministic analysis of the quantization error in Weyl-Heisenberg frame expansions which demonstrates that under certain reasonable assumptions the quantization error decays as $O(1/R^2)$ in the squared norm.

The applications considered here are based on signal characterization using information on its singularities. These schemes use nonsubsampled filter banks as a preprocessing tool. One of the central issues in these applications is the signal reconstruction from singularities. We study this problem and propose simple and efficient reconstruction algorithms. Signal interpolation is considered as well, in particular

image interpolation. We propose a locally adaptive interpolation algorithm based on extrapolation of the discrete-time wavelet transform across the scales.

Professor Martin Vetterli
Dissertation Committee Chair

One plaintive little strain mingled with the great music of the world, and with a flower for a prize you came down and stopped at my cottage door.

Tagore

To my parents and my aunt.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my gratitude to my advisor, Professor Martin Vetterli, for teaching me all these beautiful things, thus introducing me to the world of research.

I would like to thank my friends and teachers Ton Kalker and Thao Nguyen, from whom I have also learned a lot, and Grace Chang for her collaboration and for being a ray of sunshine in the Wavelet Group.

I am also in debt to other "wavelets" of our group for their assistance in various technical and survival issues related to my research. I particularly thank Grace Chang, Mike Goodwin, Vivek Goyal, Antonio Ortega and Brian Evans.

My gratitude also goes to Jelena Kovačević for her encouragement and support.

Professors Avideh Zakhor and Alberto Grünbaum, I thank for for showing their interest in my research.

Chapter 1

Introduction

1.1 History of the Subject and Motivation

The idea of time-frequency localized representations goes back to the 1940's and the work of Gabor [14], who proposed decompositions of signals in $L^2(\mathbf{R})$ in terms of modulated Gaussians as opposed to Fourier expansions. It was aimed at overcoming the major drawback of the two traditional signal descriptions, one in time and the other in the frequency, which both achieve infinitely fine resolutions in their respective domains but no resolution in the complementary domains. On the other hand, expansions based on modulated Gaussians, which achieve the lower bound on the uncertainty in the joint time-frequency domain, should facilitate descriptions with good resolution in both time and frequency.

Although the idea was itself a paramount shift towards novel signal processing techniques, a few technical problems were still to be resolved. First, is a Gabor family of functions, obtained by translating a Gaussian in time and frequency, complete in $L^2(\mathbf{R})$? What would be admissible prototype functions other than Gaussians, since in spite of its optimal localization, Gaussian need not be the most appropriate waveform for all applications? For instance, in analysis of signals with compact support, prototype functions which are themselves compactly supported would be a rather convenient choice. Then, given such a complete family of functions, how can

expansion coefficients be calculated, and moreover is there a fast algorithm for this purpose?

Expansions based on different kinds of time-frequency localized waveforms have been subsequently used in physics, geophysics and signal processing. However, not before the 1980's have they received a thorough and rigorous treatment and have the aforementioned problems been satisfactorily solved. An important result was the discovery of the relationship between filter banks and wavelets. Bases of wavelets for $L^2(\mathbf{R})$ can be derived from critically sampled filter banks [10], which are equivalent to an analogous class of bases in $\ell^2(\mathbf{Z})$. Iterated critically sampled filter banks can also be used for the efficient computation of coefficients of wavelet expansions in $L^2(\mathbf{R})$, e.g. by using Mallat's algorithm [24]. Discovery of these relations wasn't just a curiosity wedding the continuous time theory with its discrete time analogue, which had been developing independently as the theory of filter banks and subband coding; it also paved the way for applications of novel signal processing principles and brought some insights into subband coding schemes.

The theory of expansions into time-frequency localized atoms in $L^2(\mathbf{R})$ has been developed beyond the orthogonal or biorthogonal case, focusing on redundant representations based on Weyl-Heisenberg and wavelet frames [12]. However, the theory of filter banks has so far been confined mostly to the critically sampled case [47, 50], that is orthogonal and biorthogonal bases. In this thesis we study a particular class of frames in $\ell^2(\mathbf{Z})$, those which are equivalent to perfect reconstruction FIR oversampled filter banks.

One of the main reasons for studying overcomplete expansions is that the requirement for orthogonality or linear independence imposes considerable constraints which can sometimes be in conflict with other design specifications. Ironically enough, perhaps the most striking example is the fact that Gabor analysis with orthonormal bases with good resolution in both time and frequency is not possible. For a number of applications orthonormality is indeed not needed. For instance, it is hard to believe that orthonormal families of wavelets occur anywhere in nature; therefore they would not be a natural choice in modeling of biological systems or natural phenomena.

Other examples from biological systems are signal descriptions in sensory systems of mammals which are based on highly redundant representations. This redundancy compensates for system imperfections and, in spite of very coarse quantization which occurs in the cortical level, facilitates perception of fine signal structures. Furthermore, redundant representations followed by a sophisticated selection of information can even yield a good compression scheme. This has been demonstrated by Mallat and Zhong with their wavelet modulus maxima based signal compression scheme [28]. Another approach, using matching pursuit algorithm, also by Mallat [27], is based on the idea that there are greater chances for finding compact signal representations if the dictionary of elementary vectors at our disposal is richer. The full potential of these algorithms hasn't been assessed yet, however they already give results comparable to standard compression schemes [32].

It was first pointed out by Morlet (as reported in [10]) that for a given accuracy of representation overcomplete expansions allow for a progressively coarser quantization as redundancy is increased. A particular case of this effect, which has long been exploited in engineering practice, is oversampled A/D conversion, where the underlying frame is a *sinc* frame. Statistical analysis, based on the white noise model for quantization error, shows that the error variance is inversely proportional to the oversampling ratio [3], so that oversampling can be used in order to improve the accuracy of the conversion. An intuitive explanation of this effect for frame expansions in general has been given by Daubechies [10]. Daubechies showed, using a heuristic argument based on the white noise model for quantization error, that the expected value of error energy is inversely proportional to the frame redundancy factor, $E(\|e\|^2) = O(1/r)$. However, she conjectured that the error probably decays faster, as indicated by experimental results reported by Morlet.

One of the goals of this thesis is to provide some new insights into this effect and give a stronger quantitative characterization in the case of the oversampled A/D conversion and the quantization of Weyl-Heisenberg frame expansions. Our analysis is based on the following premises. Quantization error, commonly modeled as a white noise, is not really a white noise but rather a structured "noise." This is especially

true if quantization of highly redundant representations is considered. Deterministic analysis may reveal some clues on that structure which are missed by the stochastic treatment and yield a more accurate error bound. Besides, the previous studies of quantization error assume linear reconstruction, which is suboptimal. For a signal f, which is given in terms of its expansion coefficients as $f = \sum_i c_i \varphi_i$, the linear reconstruction from quantized coefficients $\{\hat{c}_i\}$ gives a recovered version $\hat{f} = \sum_i \hat{c}_i \varphi_i$. If $\{\varphi_i\}$ is an overcomplete set in the considered space of signals, then $\hat{f}$ obtained this way and f need not share the same set of quantized expansion coefficients. This indicates that the linear reconstruction does not utilize all the available information and is therefore suboptimal. It can be expected that a consistent reconstruction, that is a reconstruction which always restores a signal which has the same set of quantized coefficients as the original, will be more accurate. Quantization error is here considered with respect to a consistent reconstruction, and it is shown that in some particular but relevant cases the error is inversely proportional to the square of the expansion redundancy factor, $\|e\|^2 = O(1/r^2)$.

1.2 Overview and Contribution of the Thesis in the Context of Previous Work

A global framework for the material presented in the thesis is set up in Chapter 2. We review the main concepts of linear expansions in Hilbert spaces. The emphasis is on overcomplete expansions, that is those based on frames. Three particular classes which are reviewed are wavelet and Weyl-Heisenberg frames, as well as frames of complex exponentials in $L^2[-\sigma, \sigma]$ spaces.

Oversampled filter banks are studied in Chapter 3. The previous work in this area was concerned only with FIR two channel nonsubsampled filter banks [40, 28], and amounted to the investigation of the perfect reconstruction condition. In Chapter 3, general N channel FIR filter banks followed by subsampling by factor K, $K < N$, are considered. A necessary and sufficient condition on a filter bank to be equivalent to a frame in $\ell^2(\mathbf{Z})$ is derived, as well as a necessary and sufficient condition for the

tightness of that frame. Feasibility of perfect reconstruction using FIR filters is also investigated and a necessary and sufficient condition is formulated. Complete parameterizations of classes of filter banks satisfying these conditions are also given. For a given oversampled filter bank which satisfies the frame condition, the corresponding perfect reconstruction synthesis filter bank is not unique. One of them is dual to the analysis filter bank, that is frames in $\ell^2(\mathbf{Z})$ which are equivalent to the two filter banks are dual frames [10]. An advantage of such a synthesis filter bank is that when reconstruction from noisy subband signals is performed, it projects to zero the noise component which is orthogonal to the range of the subband expansion. We also give a necessary and sufficient condition on an FIR filter bank for its dual to also be FIR and parameterize a class of filter banks satisfying this property. Two special cases, focused on in Chapter 3, are nonsubsampled filter banks and modulated filter banks. Nonsubsampled filter banks, implementing shift invariant redundant transforms, have a considerable potential for signal analysis and are related to continuous-time signal filtering. Furthermore, absence of subsampling in the channels results in a very flexible design. We illustrate this point by giving a design procedure for maximally flat two channel filter banks from which highly regular wavelets with a given number of vanishing moments can be derived.

The treatment of modulated filter banks in this thesis is concerned with FIR paraunitary filter banks, that is tight finite length Weyl-Heisenberg frames in $\ell^2(\mathbf{Z})$. Analogously to the situation in continuous-time, short-time Fourier analysis with reasonable bases, or equivalently nontrivial critically sampled filter banks, is not possible. It is demonstrated in this thesis that if some redundancy is allowed the situation changes, so that it is possible to attain tight Weyl-Heisenberg frames with good localization and some other desirable properties such as symmetry. We give a complete parameterization of these tight frames, which are the tool for short-time Fourier analysis in $\ell^2(\mathbf{Z})$. These frames were a subject of a considerable study, however results were confined to finite dimensional spaces or "almost tight" frames [33, 37, 36, 38, 53].

The problem of quantization of overcomplete expansions is considered in Chapter 4. A rigorous treatment of the effect of quantization error reduction in frames, in

the case of tight Weyl-Heisenberg frames with integer oversampling ratios, is due to Munch [31]. Munch proved that the contribution to error energy due to coefficients which correspond to frame vectors localized in a given bounded region of the time-frequency plane, behaves as $O(1/r)$. Note, however, that the estimate of total error which is obtained from Munch's analysis is unbounded. We assert in this thesis that linear reconstruction is suboptimal if we deal with overcomplete expansions, and that a consistent reconstruction should give an error $\|e\|^2 = O(1/r^2)$. So far, this claim has been experimentally verified for frames in $\mathbf{R}^n$ [16] and also been proven in the case of oversampled A/D conversion of periodic bandlimited signals (trigonometric polynomials) [44, 43, 45]. In Chapter 4, it is proven that this also holds in the case of oversampled A/D conversion of bandlimited signals in $L^2(\mathbf{R})$. Our approach is based on the concept of consistent reconstruction, first pointed out in [44], as well as on a deterministic model which uses results of nonharmonic Fourier analysis. If error-rate properties of oversampled A/D conversion are considered, instead of the error itself, then an efficient lossless encoding of quantized samples becomes crucial. With standard PCM (Pulse Code Modulation), quantization error as the function of the bit rate, B, behaves as $\|e\|^2 = O(1/B^2)$ when oversampling increases. An efficient coding scheme, proposed here, attains $\|e\|^2 = O(2^{-\beta B})$. The purpose of this result is to illustrate the importance of efficient lossless encoding of overcomplete expansions, and perhaps give directions for further research in this subject. Note that the deterministic analysis of oversampled A/D conversion is an important fundamental result per se. Besides, it gives the first proof of the $O(1/r^2)$ error reduction property of frames in a case of an infinite dimensional space. This analysis is further generalized to the quantization error of Weyl-Heisenberg frame expansions in $L^2(\mathbf{R})$, proving again that under reasonable assumptions $\|e\|^2 = O(1/r^2)$.

Applications based on wavelet modulus maxima or wavelet zero crossings representation are very illustrative for various concepts related to overcomplete expansions. These are discussed in Chapter 5. Wavelet modulus maxima and wavelet zero-crossings representations were introduced by Mallat, *et al.* [25, 26, 28] as a tool for extraction of information on singularities, which are considered to be among the

most meaningful features for signal characterization. The representations are based on irregular sampling of the multiscale wavelet transform, which is implemented in discrete-time using iterated nonsubsampled filter banks. The samples are taken at local modulus maxima or zero-crossings, which for a particular class of wavelets occur at points of sharp signal variation. Promising performance of these representations has been demonstrated in applications such as signal denoising and compression [26, 28]. They also provide a convenient framework for studying problems such as signal characterization and reconstruction from multiscale edges or zero-crossings, which have been subjects of considerable attention in the signal processing community in the last two decades. One of the central issues about the two representations, in all of these applications, is signal reconstruction from the information they provide. Original reconstruction algorithms, which outperform the existing ones [25, 28, 4] in either numerical complexity or convergence properties are also described in Chapter 5.

Another application discussed here is in image interpolation. Given a small image, the classical problem of image interpolation is to magnify the image many times without loss in the sharpness of the picture. Some existing methods such as bilinear and spline interpolations generate blurred images since they do not utilize any information relevant to preserving the image clarity. To deblur these images, one could use the standard approach of unsharp masking [20]. Other methods include modeling the edges or filtering with nonlinear filters to boost the high frequencies needed to make an image look sharper. The algorithm described in this thesis, which was developed in collaboration with Grace Chang, is a locally adaptive image interpolation scheme. The interpolating functions do not appear explicitly, but are contained in the enhancement component which is added to an initial linearly interpolated image. The enhancement component is a result of analysis of signal singularities, specifically propagation of local extrema of a wavelet transform across scales. The implementation is based on preprocessing using nonsubsampled filter banks. Images interpolated using this algorithm provide better subjective quality than those obtained with the standard schemes.

The results presented in this thesis are also documented in the publications listed

in Appendix A.

The contributions of the thesis are listed below.

- Development of a framework for the study of oversampled filter banks, i.e. filter bank frames in $\ell^2(\mathbf{Z})$.
- Necessary and sufficient condition on a filter bank to be equivalent to a frame in $\ell^2(\mathbf{Z})$.
- Necessary and sufficient condition on a filter bank to be equivalent to a tight frame in $\ell^2(\mathbf{Z})$.
- Necessary and sufficient condition on a finite impulse response (FIR) filter bank frame to have a dual frame consisting of finite length vectors.
- Necessary and sufficient condition on an FIR filter bank frame to have the minimal dual frame consisting of finite length vectors.
- Complete parameterization of FIR filter bank frames.
- Complete parameterization of tight FIR filter bank frames.
- Complete parameterization of FIR filter bank frames which have dual frames consisting of finite length vectors.
- Parameterization of FIR filter bank frames which have minimal duals consisting of finite length vectors.
- Complete parameterization of tight FIR Weyl-Heisenberg frames in $\ell^2(\mathbf{Z})$.
- Designing procedure for wavelets with high regularity and a given number of vanishing moments.
- Deterministic analysis of oversampled A/D conversion of signals in $L^2(\mathbf{R})$.
- Proof of $O(1/r^2)$ quantization error behavior in oversampled A/D conversion, with r being the oversampling ratio.

- Efficient lossless encoding scheme for oversampled A/D conversion.
- Analysis of error-rate characteristics of oversampled A/D conversion and methods for its improvement.
- Explanation of the quantization error reduction property in overcomplete expansions.
- Deterministic analysis of quantization error in Weyl-Heisenberg frame expansions and proof of its $O(1/r^2)$ behavior as a function of the frame redundancy r.
- Study of discrete-time wavelet extrema and wavelet zero-crossings representations (multiscale edge representations) in the framework of convex representations in $\ell^2(\mathbf{Z})$.
- Efficient algorithm for signal reconstruction from the wavelet extrema representation.
- Efficient algorithm for signal reconstruction from the wavelet zero-crossings representation.
- Implementation of the orthogonal projection operator onto the range of the discrete-time wavelet transform.
- Novel locally adaptive signal interpolation scheme.

Notation

The notation used in this thesis is described below.

The identity operator in a Hilbert space will be denoted by Id.

The Fourier transform of a signal $f(t)$, $\mathcal{F}\{f(t)\}$, will be written as $\hat{f}(\omega)$.

We say that a signal $f(t)$ is σ-bandlimited if

$$\hat{f}(\omega) = 0 \text{ for } |\omega| > \sigma, \text{ and } \|f\|^2 = \int_{-\infty}^{\infty} |f(t)|^2 dt < \infty.$$

Similarly, a signal $f(t)$ is said to be T-timelimited if

$$f(t) = 0 \text{ for } |t| > T, \text{ and } \|f\|^2 = \int_{-\infty}^{\infty} |f(t)|^2 dt < \infty.$$

Convolution of functions in $L^2(\mathbf{R})$ will be denoted by $*$, i.e.

$$f * g = f(t) * g(t) = f * g(t) = \int_{-\infty}^{\infty} f(s)g(t-s)ds.$$

For a sequence φ, $\tilde{\varphi}$ will denote the complex conjugate of the time reversed version of φ, $\tilde{\varphi}[n] = \varphi^*[-n]$. Also, for a filter $H(z)$ with impulse response $h[n]$, $\tilde{H}(z)$ will denote the filter whose impulse response is $\tilde{h}[n]$. Similarly, when used with a matrix $\mathbf{H}(z)$ whose entries are rational functions of the complex variable z, $\tilde{\mathbf{H}}(z)$ will denote the matrix obtained from $\mathbf{H}(z)$ by transposing it, changing all coefficients of the rational functions in $\mathbf{H}(z)$ by their complex conjugates, and substituting z by z^{-1}. If $\tilde{\mathbf{H}}(z) = \mathbf{H}(z)$ we say that $\mathbf{H}(z)$ is *parahermitian*. Polynomial matrix $\mathbf{H}(z)$, whose determinant, $\det\mathbf{H}(z)$, is a nonzero constant is called *unimodular matrix*. Note that here we shall use the term *polynomial* for *Laurent polynomials* in general, that is FIR filters $H(z) = \sum_{i=n}^{m} h_i z^i$, which possibly contain both positive and negative powers of z. The complex conjugate transpose of a vector $\mathbf{v}$ will be denoted as $\mathbf{v}^*$, and when used with a scalar, the $*$ superscript will denote its complex conjugate value.

Chapter 2

Frames in a Hilbert Space

2.1 Notion of Frames

A common signal processing framework is to consider signals as vectors in a Hilbert space $\mathcal{H}$, usually $L^2(\mathbf{R})$, $\ell^2(\mathbf{Z})$ or $\mathbf{R}^n$. Signal processing algorithms often involve some sorts of linear expansions. We may justifiably say that linear expansions are ubiquitous tools of signal processing. Linear expansion refers to representation of a signal as a linear combination of vectors of a family $\{\varphi_j\}_{j\in J}$ which is complete in $\mathcal{H}$. Completeness of $\{\varphi_j\}_{j\in J}$ means that any signal x in the space can be represented as a linear combination,

$$x = \sum_{j\in J} c_j^x \varphi_j. \tag{2.1}$$

In addition to completeness, for applications in signal processing it is essential that $\{\varphi_j\}_{j\in J}$ also has some stability properties. First, the coefficients of the expansion should constitute a sequence of finite energy for any signal in the space,

$$\{c_j^x\}_{j\in J} \in \ell^2(J), \quad \forall x \in \mathcal{H}. \tag{2.2}$$

Moreover signals in any given bounded set cannot have expansion coefficients of arbitrarily high energy. In particular, expansion coefficients of signals in the unit ball should be bounded by some finite constant b,

$$\sum_{j\in J} |c_j^x|^2 \le b \quad \text{whenever} \quad \|x\| \le 1. \tag{2.3}$$

The second requirement is that the expansions should provide a unique description of signals in $\mathcal{H}$, and enable discrimination with a certain reasonable resolution. In other words, expansion coefficients corresponding to two different signals x and y should not be close in $\ell^2(J)$ if x and y are not close themselves. A precise formulation of this requirement is that there exists a constant $a > 0$ such that

$$\sum_{j\in J} |c_j^x - c_j^y|^2 \geq a, \quad \textit{whenever} \quad \|x - y\|^2 \geq 1. \tag{2.4}$$

A family $\{\varphi_j\}_{j\in J}$ which satisfies these requirements is said to be a *frame* in $\mathcal{H}$.

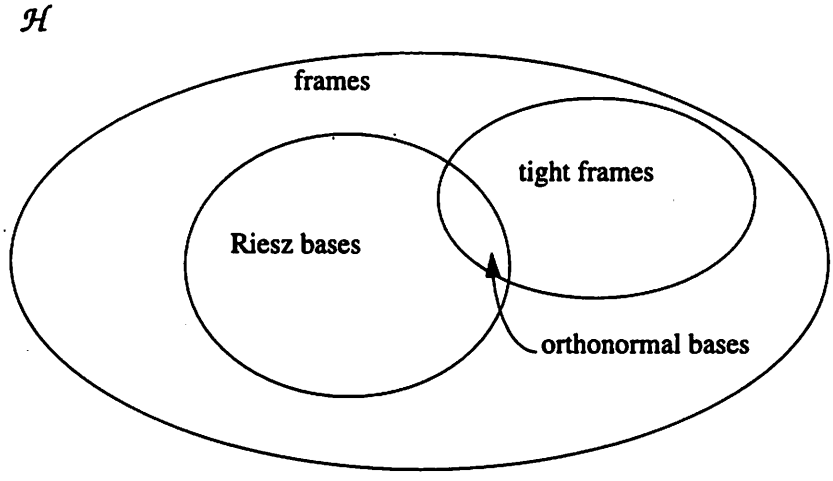

Figure 2.1: *Principal classes of frames in a Hilbert space.*

Three fundamental subclasses of frames are *tight frames*, *Riesz bases* and *orthonormal bases* (see Figure 2.1). A basis in $\mathcal{H}$ is said to be Riesz basis if it is the image of an orthonormal basis under a bounded operator with a bounded inverse. Examples of frames in theses classes for $\mathbf{R}^2$ are shown in Figure 2.2. Note that in finite dimensional spaces, i.e. $\mathbf{R}^n$, completeness of $\{\varphi_j\}_{j\in J}$ implies the stability properties, so most of problems discussed in this chapter are nontrivial only for infinite dimensional spaces. Nonetheless, Figure 2.2 illustrates distinctive properties of the three main subclasses. Both bases[1] and orthonormal bases are linearly independent

[1]The term *basis* will here always mean Riesz basis.

families of vectors. On the other hand, frames and tight frames in general are redundant, overcomplete families in $\mathcal{H}$ and therefore exhibit some linear dependencies. Tight frames and orthonormal bases are subclasses of frames and bases, respectively, characterized by a certain structure, which makes them particularly convenient in signal analysis.

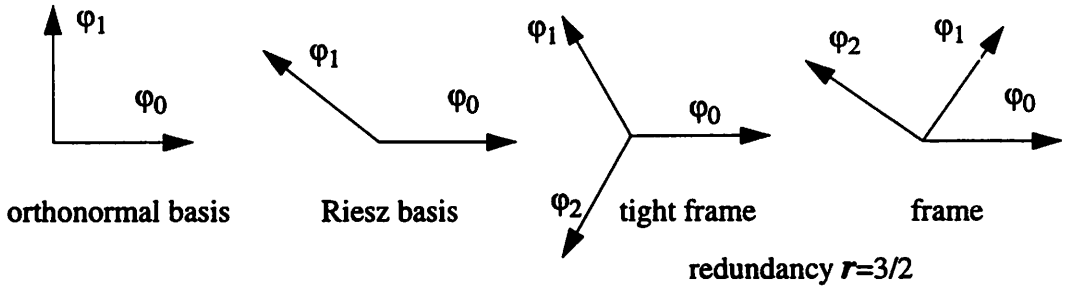

Figure 2.2: *Examples of different frame classes in* $\mathbf{R}^2$.

In this chapter, basic results on frames are reviewed, based on material presented in [10, 12, 13, 56], and some classes of frames which are relevant for further considerations are described.

2.2 Properties of Frames

Expanding a signal with respect to an orthonormal basis $\{\varphi_j\}_{j\in J}$ is equivalent to characterizing it through inner products with the vectors of the basis. This is due to the fact that coefficients of the expansion $x = \sum_{j\in J} c_j^x \varphi_j$ are equal to the inner products, $c_j^x = \langle x, \varphi_j \rangle$. The two representations, expansion and the inner product sequence, are not equivalent for frames in general, but there is still a certain duality. The following definition of a frame in a Hilbert space is based on the inner product description.

Definition 1 *A family of vectors* $\{\varphi_j\}_{j\in J}$ *in a Hilbert spaces* $\mathcal{H}$ *is called a frame if there exist* $A > 0$ *and* $B < \infty$ *such that for all* f *in* $\mathcal{H}$,

$$A\|f\|^2 \leq \sum_{j\in J} |\langle f, \varphi_j \rangle|^2 \leq B\|f\|^2. \tag{2.5}$$

A and B are called frame bounds. □

For a frame $\{\varphi_j\}_{j\in J}$, there exists a dual frame $\{\psi_j\}_{j\in J}$, such that any f in the space can be reconstructed from the inner products $\{\langle f,\varphi_j\rangle\}_{j\in J}$ as

$$f=\sum_{j\in J}\langle f,\varphi_j\rangle\psi_j. \tag{2.6}$$

The frames $\{\varphi_j\}_{j\in J}$ and $\{\psi_j\}_{j\in J}$ have interchangeable roles, so that any f can also be written as a superposition of φ_j vectors, as

$$f=\sum_{j\in J}\langle f,\psi_j\rangle\varphi_j. \tag{2.7}$$

Implications of this expansion formula are actually the stability conditions discussed in the previous section. If A_1 and B_1 are frame bounds of $\{\psi_j\}_{j\in J}$ then (2.4) and (2.3) are satisfied with $a=A_1$ and $b=B_1$. Existence of a dual frame also establishes duality between the inner product description and the series expansion.

In order to establish the existence of a dual frame and give a systematic generating procedure we introduce the frame operator.

Definition 2 *If $\{\varphi_j\}_{j\in J}$ is a frame in $\mathcal{H}$, then the frame operator F is the linear operator from $\mathcal{H}$ to $\ell^2(J)$, defined by*

$$Ff=c,\quad c=\{c_j:c_j=\langle f,\varphi_j\rangle\}_{j\in J}. \tag{2.8}$$

□

The frame condition (2.5) implies that F is a bounded invertible operator from $\mathcal{H}$ to $\ell^2(J)$. The Hilbert adjoint F^* of F is a linear operator from $\ell^2(J)$ to $\mathcal{H}$ satisfying

$$\langle F^*c,f\rangle=\langle c,Ff\rangle \tag{2.9}$$

for all $f\in\mathcal{H}$ and $c\in\ell^2(J)$. If Ff is written in the expanded form, we obtain

$$\langle c,Ff\rangle=\sum_{j\in J}c_j\langle\varphi_j,f\rangle, \tag{2.10}$$

so at least in the weak sense,

$$F^*c = \sum_{j\in J} c_j\varphi_j. \tag{2.11}$$

The linear operator F^*F, given by

$$F^*Ff = \sum_{j\in J}\langle f,\varphi_j\rangle\varphi_j, \tag{2.12}$$

is self-adjoint. It follows from the frame condition (2.5) that

$$A\|f\|^2 \le \langle F^*Ff, f\rangle \le B\|f\|^2 \tag{2.13}$$

for any $f \in \mathcal{H}$, that is, F^*F is positive definite with positive upper and lower bounds. This implies that the inverse $(F^*F)^{-1}$ exists. The image of $\{\varphi_j\}_{j\in J}$ under $(F^*F)^{-1}$, $\{\tilde{\varphi}_j : \tilde{\varphi}_j = (F^*F)^{-1}\varphi_j\}_{j\in J}$, turns out to be a frame dual to $\{\varphi_j\}_{j\in J}$.

Proposition 1 *[10] The $\{\tilde{\varphi}_j\}_{j\in J}$ constitutes a frame with frame bounds $1/B$ and $1/A$. The associated frame operator*

$$\tilde{F} : \mathcal{H} \to \ell^2(J), \quad \tilde{F}f = \{c_j : c_j = \langle f,\varphi_j\rangle\}_{j\in J},$$

satisfies the following:

(i) $\tilde{F}^*F = F^*\tilde{F} = \mathrm{Id}$,

(ii) $\tilde{F}F^* = F\tilde{F}^*$ *is the orthogonal projection operator in* $\ell^2(J)$ *onto* $\mathrm{Ran}(F) = \mathrm{Ran}(\tilde{F})$.

□

This proposition establishes in a constructive manner the existence of a dual frame. If the frame $\{\varphi_j\}_{j\in J}$ is overcomplete, that is φ_j are linearly dependent, then its dual is not unique. To illustrate this consider a linear combination $\sum_{j\in J} a_j\varphi_j = 0$. For any f in $\mathcal{H}$, $\sum_{j\in J} a_j\langle f,\varphi_j\rangle = 0$. If $\{\psi_j\}_{j\in J}$ is a dual frame, and ϕ is any vector in $\mathcal{H}$,

$$\begin{aligned} f &= \sum_{j\in J}\langle f,\varphi_j\rangle\psi_j && (2.14)\\ &= \sum_{j\in J}\langle f,\varphi_j\rangle\psi_j + \sum_{j\in J} a_j\langle f,\varphi_j\rangle\phi && (2.15)\\ &= \sum_{j\in J}\langle f,\varphi_j\rangle(\psi_j + a_j\phi), && (2.16)\end{aligned}$$

so that $\{\psi_j + a_j\phi\}_{j\in J}$ is also dual to $\{\varphi_j\}_{j\in J}$. Another consequence of linear dependencies in $\{\varphi_j\}_{j\in J}$ is that expansions with respect to this frame are not unique. For instance, if $f = \sum_{j\in J} c_j\varphi_j$ and $\sum_{j\in J} a_j\varphi_j = 0$ then also $f = \sum_{j\in J}(c_j + a_j)\varphi_j$.

The dual $\{\tilde{\varphi}_j\}_{j\in J}$ has the distinctive property that for any $c = \{c_j\}_{j\in J}$ in the orthogonal complement of $\mathrm{Ran}(F)$, $\sum_{j\in J} c_j\tilde{\varphi}_j = 0$. This is a consequence of Proposition 1, and is particularly important for reconstruction of signals from coefficients degraded by an additive noise. If a signal f is synthesized from noisy coefficients $\{\langle f, \varphi_j\rangle + n_j\}$ using $\{\tilde{\varphi}_j\}_{j\in J}$ as

$$f' = \sum_{j\in J}(\langle f, \varphi_j\rangle + n_j)\tilde{\varphi}_j, \tag{2.17}$$

the noise component which is in the orthogonal complement of $\mathrm{Ran}(F)$ is automatically reduced to zero. No other dual of $\{\varphi_j\}_{j\in J}$ implicitly performs this projection while effecting the reconstruction.

Another important fact about $\{\tilde{\varphi}_j\}_{j\in J}$ is related to the nonuniqueness of the expansions. Out of all expansion sequences $\{c_j\}$ such that $f = \sum_{j\in J} c_j\varphi_j$, $\{\langle f, \tilde{\varphi}_j\rangle\}$ has the minimum norm. This is established by the following proposition.

Proposition 2 *[10] If $f = \sum_{j\in J} c_j\varphi_j$ for some $c = \{c_j\}_{j\in J}$, then*

$$\sum_{j\in J} |c_j|^2 \geq \sum_{j\in J} |\langle f, \tilde{\varphi}_j\rangle|^2,$$

and equality holds only if $c_j = \langle f, \tilde{\varphi}_j\rangle$ for all $j \in J$. □

Being so special, $\{\tilde{\varphi}_j\}_{j\in J}$ deserves a special name, and is said to be the *minimal dual* of $\{\varphi_j\}_{j\in J}$.[2]

Calculation of the minimal dual frame is not always an easy task, and a closed form solution does not exist, except for some particular cases. However, there does exist a formula which describes a numerical algorithm for the computation of $\{\tilde{\varphi}_j\}_{j\in J}$. It is given by

$$\tilde{\varphi}_j = \frac{2}{A+B}\sum_{k=0}^{\infty} R^k\varphi_j, \tag{2.18}$$

[2]In their pioneering work on frames [13], Duffin and Schaeffer used the term *conjugate frames* for a frame and its minimal dual.

where

$$R = \mathrm{Id} - \frac{2}{A+B}F^*F. \tag{2.19}$$

The speed of convergence of this series depends on the bounds A and B. It can be shown that $\|R\| \leq (B-A)/(B+A)$, so the closer the frame bounds are, the faster the series converges. In the limit, when the frame bounds are equal, all these formulae become much simpler, and the minimal dual frame is given by the first term of the series (2.18), $\tilde{\varphi}_j = A^{-1}\varphi_j$.

A frame with equal frame bounds is called a *tight frame*. For a tight frame and any f in the space

$$\sum_{j\in J} |\langle f, \varphi_j\rangle|^2 = A\|f\|^2. \tag{2.20}$$

This condition implies that

$$\langle F^*Ff, f\rangle = A\|f\|^2, \quad \forall f \in \mathcal{H}. \tag{2.21}$$

Hence, $F^*F = A\mathrm{Id}$[3], which finally gives $\tilde{\varphi}_j = A^{-1}\varphi_j$. Therefore, for all f in $\mathcal{H}$

$$f = \frac{1}{A}\sum_{j\in J}\langle f, \varphi_j\rangle\varphi_j. \tag{2.22}$$

This formula is reminiscent of orthogonal expansions. However, tight frames are in general not orthonormal bases, but overcomplete families in $\mathcal{H}$. If vectors of the tight frame are normalized to unit norm, $\|\varphi_j\| = 1$, the frame bound A gives a measure of redundancy of the frame. Tight frames which are not redundant are orthonormal bases.

Proposition 3 *[10] If $\{\varphi_j\}_{j\in J}$ is a tight frame with frame bound $A = 1$, and $\|\varphi_j\| = 1$ for all $j \in J$, then $\{\varphi_j\}_{j\in J}$ is an orthonormal basis.* □

Analogous relations exist between Riesz bases and frames. They are established by the following propositions. The first one describes the effect of the removal of a vector from a frame.

[3]Note that in the case of complex Hilbert spaces, which we consider, if a linear operator S on $\mathcal{H}$ satisfies $\langle Sx, x\rangle = A\|x\|^2$ for all $x \in \mathcal{H}$, then $Sx = Ax$.

Proposition 4 *[56] The removal of a vector from a frame leaves either a frame or an incomplete set.* □

A frame which ceases to be a frame when any of its vectors is removed is said to be an *exact frame*. It turns out that the class of Riesz bases and the class of exact frames are equivalent.

Proposition 5 *[56] A sequence of vectors in a Hilbert space $\mathcal{H}$ is a Riesz basis if and only if it is an exact frame.* □

As opposed to orthonormal bases, vectors of a Riesz basis are not mutually orthogonal, however a biorthogonal basis always exists.

Proposition 6 *[13] If $\{\varphi_j\}_{j\in J}$ is an exact frame, then it is biorthogonal to its minimal dual frame $\{\tilde{\varphi}_j\}_{j\in J}$, that is*

$$\langle \tilde{\varphi}_i, \varphi_j \rangle = \begin{cases} 1, & i = j \\ 0, & otherwise \end{cases} \tag{2.23}$$

□

At this point we end the overview of general results on frames and transfer our attention to some particular cases which gave rise to the general theory and drove further developments.

2.3 Some Classes of Structured Frames

2.3.1 Frames of Complex Exponentials in $L^2[-\sigma, \sigma]$

The general theory of frames originated from studying the characterizability of bandlimited signals from a sequence of samples [13]. It can be formulated as follows. Given a space of σ-bandlimited signals, it is of interest to characterize sequences of sampling instants which uniquely describe signals in the space and moreover allow for a numerically stable reconstruction. Frames of complex exponentials in $L^2[-\sigma, \sigma]$ are closely related to this problem.

A classical result of Fourier analysis, in communications known as the sampling theorem, asserts that regular sampling of a σ-bandlimited signal $f(t)$ at points $\{n\tau\}_{n\in\mathbf{Z}}$ introduces no loss of information as long as $\tau \leq \pi/\sigma$. Recall that $f(t)$ can be recovered from the samples as

$$f(t) = \frac{1}{\tau}\sum_n f(n\tau) sinc_\sigma(t - n\tau), \tag{2.24}$$

where

$$sinc_\sigma(x) = \frac{\sigma}{\pi}\frac{sin(\sigma x)}{\sigma x}.$$

The value which $f(t)$ attains at the sampling instant $n\tau$ represents the inner product

$$f(n\tau) = \langle f, sinc_\sigma(t - n\tau)\rangle,$$

so that $f(t)$ can be written as

$$f(t) = \frac{1}{\tau}\sum_n \langle f, \varphi_{n,\tau}\rangle \varphi_{n,\tau}(t), \tag{2.25}$$

where

$$\varphi_{n,\tau}(t) = sinc_\sigma(t - n\tau).$$

Therefore, sampling above the Nyquist rate, $\tau < \pi/\sigma$, is equivalent to expanding a signal with respect to

$$\{\varphi_{n,\tau}(t)\}_{n\in\mathbf{Z}},$$

which turns out to be a tight frame for the space of σ-bandlimited signals. Alternatively, this sampling can be viewed as a linear expansion of the Fourier transform of $f(t)$ in terms of complex exponentials $\{e^{j\omega n\tau}\}$, which constitute a tight frame for $L^2[-\sigma,\sigma]$. This interpretation comes from the fact that the space of σ-bandlimited signals is isometrically isomorphic to $L^2[-\sigma,\sigma]$. In general, σ-bandlimited signals can be reconstructed in a numerically stable way from the samples at points $\{\lambda_n\}_{n\in\mathbf{Z}}$ if and only if $\{e^{j\lambda_n\omega}\}$ is a frame in $L^2[-\sigma,\sigma]$. Recall that a sequence of complex exponentials, $\{e^{j\lambda_n\omega}\}$, is said to be a frame in $L^2[-\sigma,\sigma]$ if there exist constants $A > 0$ and $B < \infty$ such that

$$A\int_{-\sigma}^{\sigma} |\hat{f}(\omega)|^2 d\omega \leq \frac{1}{2\pi}\sum_n \left|\int_{-\sigma}^{\sigma} \hat{f}(\omega) e^{j\lambda_n\omega} d\omega\right|^2 \leq B\int_{-\sigma}^{\sigma} |\hat{f}(\omega)|^2 d\omega, \tag{2.26}$$

for any $\hat{f}(\omega) \in L^2[-\sigma, \sigma]$. This can be rephrased as the condition that

$$A\|f(t)\|^2 \leq \sum_n |f(\lambda_n)|^2 \leq B\|f(t)\|^2 \tag{2.27}$$

for every σ-bandlimited function $f(t)$. If this is satisfied we shall also say that $\{\lambda_n\}$ is a frame sequence for the space of σ-bandlimited signals.

Definition 3 *A sequence of real numbers $\{\lambda_n\}$ is said to be a frame sequence for the space of σ-bandlimited signals if $\{e^{j\lambda_n\omega}\}$ is a frame in $L^2[-\sigma, \sigma]$.* □

Conditions under which a sequence $\{e^{j\lambda_n\omega}\}$ is complete and moreover a frame in an $L^2[-\sigma, \sigma]$ space are the core issues of nonharmonic Fourier analysis. An excellent introductory treatment of this subject can be found in [56]. Here, only a few illuminating results are reviewed in order to provide some intuition about frames of complex exponentials.

If samples of a σ-bandlimited signals taken at points $\{\lambda_n\}$ are sufficient for a *numerically stable*, exact reconstruction, then it can be expected that $\{\lambda_n\}$ will retain this property under small perturbations. The next theorem asserts that this is indeed true in the case of the orthonormal basis $\{e^{jn\pi/\sigma}\}$.

Theorem 1 *[22]* (Kadec's $\frac{1}{4}$ Theorem). *If $\{\lambda_n\}$ satisfies*

$$|\lambda_n - n\frac{\pi}{\sigma}| \leq L < \frac{1}{4}\frac{\pi}{\sigma}, \quad n = 0, \pm 1, \pm 2, ...,$$

then $\{e^{j\lambda_n\omega}\}$ is a Riesz basis for $L^2[-\sigma, \sigma]$. □

The bound given in this theorem is tight, i.e. the result does not follow if $\sup_n |\lambda_n - n\pi/\sigma| = \pi/4\sigma$. A more general result on stability of frames of complex exponentials is established by the following theorem.

Theorem 2 *[56] Let $\{e^{j\lambda_n\omega}\}$ be a frame in $L^2[-\sigma, \sigma]$, with bounds $0 < A \leq B < \infty$ and δ a given positive number. If a sequence $\{\mu_n\}$ satisfies $|\lambda_n - \mu_n| \leq \delta$ for all n, then for every σ-bandlimited signal $f(t)$*

$$A(1 - \sqrt{C})^2\|f\|^2 \leq \sum_n |f(\mu_n)|^2 \leq B(1 + \sqrt{C})^2\|f\|^2, \tag{2.28}$$

where

$$C = \frac{B}{A}\left(e^{\sigma\delta} - 1\right)^2. \tag{2.29}$$

□

Corollary 1 *If the set* $\{e^{j\lambda_n\omega}\}$ *is a frame for* $L^2[-\sigma,\sigma]$ *then there is a positive constant* δ *with the property that* $\{e^{j\mu_n\omega}\}$ *is also a frame for* $L^2[-\sigma,\sigma]$ *whenever* $|\lambda_n - \mu_n| \leq \delta$ *for every* n. □

This corollary follows immediately from Theorem 2 if δ in the statement of the theorem is chosen small enough that C is less then 1. For instance, note that there exists a $\delta_{1/4}(\{\lambda_n\},\sigma)$, such that whenever $\delta < \delta_{1/4}(\{\lambda_n\},\sigma)$, $\{e^{j\mu_n\omega}\}$ is a frame with frame bounds $A/4$ and $9B/4$. This fact will be used for derivations in Chapter 4.

Another sufficient condition on $\{\lambda_n\}$ to be a frame sequence is given by Duffin and Schaeffer [13]. Basically, it asserts that $\{\lambda_n\}$ is a frame sequence for the space of σ-bandlimited signals if the λ_n are sufficiently dense on the real axis.

Definition 4 *A sequence* $\{\lambda_n\}$ *of real or complex numbers has uniform density* d, $d > 0$, *if there are constants* $L < \infty$ *and* $\delta > 0$ *such that*

$$\begin{aligned} &|\lambda_n - \tfrac{n}{d}| \leq L, \qquad n \in \mathbf{Z} \\ &|\lambda_n - \lambda_m| \geq \delta > 0 \quad n \neq m. \end{aligned} \tag{2.30}$$

□

We can now formulate the following criterion.

Theorem 3 *[13] If a sequence* $\{\lambda_n\}$ *has a uniform density* $d > \sigma/\pi$, *then* $\{e^{j\lambda_n\omega}\}$ *is a frame in* $L^2[-\sigma,\sigma]$. □

The condition of this theorem requires the uniform density d strictly greater than σ/π. This bound is tight, since there are sequences of uniform density $d = \sigma/\pi$ which are not frame sequences in $L^2[-\sigma,\sigma]$. An example is any sequence obtained from $\{n\pi/\sigma\}$ by dropping one of the points. However, this theorem gives only a sufficient

condition. If some restriction is imposed on L in Definition 4, it is possible to find sequences of uniform density $d = \sigma/\pi$ which are frame sequences in $L^2[-\sigma, \sigma]$. The existence of these sequences is established by Kadec's 1/4 Theorem.

The results on irregular sampling which are reviewed here indicate a tradeoff between regularity and density of sampling. If signals are to be represented with a minimal number of samples, these samples are to be located in strictly defined areas uniformly distributed on the time axis. As soon as some redundancy (oversampling) is introduced, the restrictions on the distribution of the sampling points become much looser.

2.3.2 Weyl-Heisenberg and Wavelet Frames

Frames which are designed for a particular application usually exhibit a specialized structure. Two examples, which arise from engineering practice as well as physics, are Weyl-Heisenberg and wavelet frames in $L^2(\mathbf{R})$.

Weyl-Heisenberg frames are the tool for Short-Time Fourier analysis and have also been used in many areas of theoretical physics under the name of coherent states. They consist of vectors which are obtained by translating and modulating a single prototype window function,

$$\left\{\varphi_{mn}(t) : \varphi_{mn}(t) = e^{jm\omega_0 t}\varphi(t - nt_0)\right\}_{m,n\in Z}. \tag{2.31}$$

Vectors $\varphi_{mn}(t)$ are distributed on a rectangular lattice in the time-frequency plane. The lattice cell size is given by the time and frequency steps of the frame, t_0 and ω_0, respectively (see Figure 2.3). All vectors have the same effective support, which is determined by the shape of the window. The window initially proposed by Gabor [14] was Gaussian, since it achieves the lower bound of uncertainty in the joint time-frequency plane, facilitating signal description with the best possible joint resolution. Other window functions are admissible, too. In fact, the only restriction which the window function of a Weyl-Heisenberg frame with bounds A and B has to obey is

$$A \leq \frac{2\pi}{\omega_0 t_0}\|\varphi\|^2 \leq B. \tag{2.32}$$

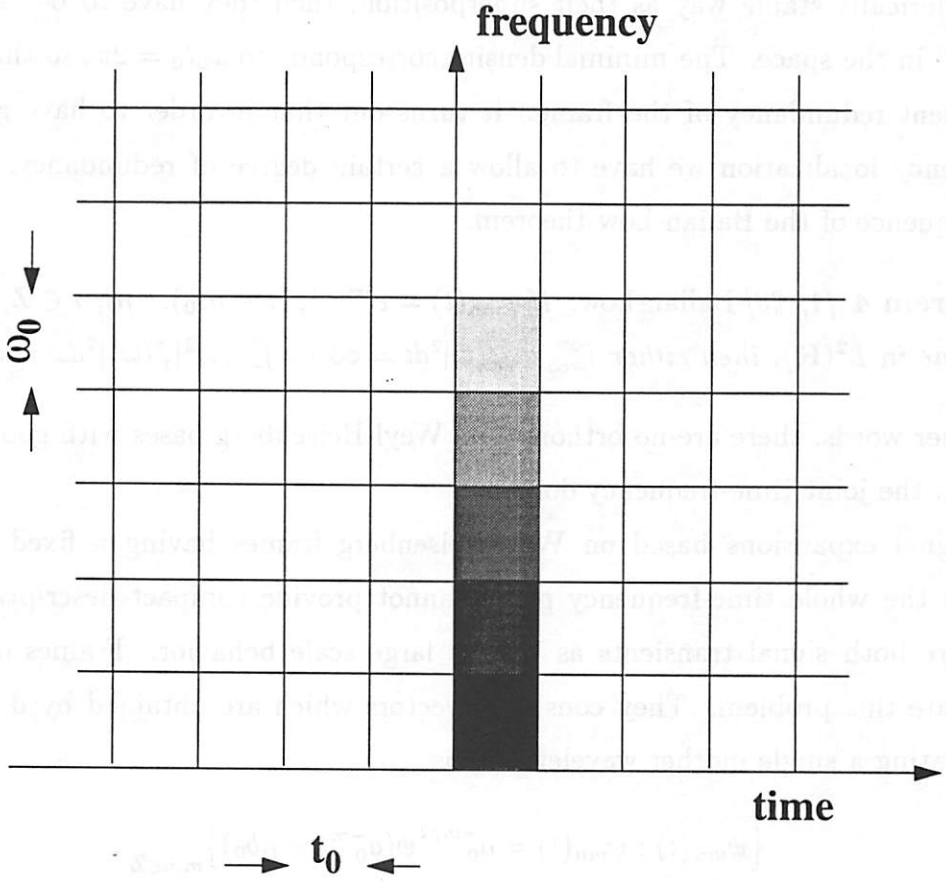

Figure 2.3: *Distribution of vectors of a Weyl-Heisenberg frame in the time-frequency plane. The resolution of the frame, determined by the effective support of frame vectors, is fixed throughout the plane.*

The restriction on the time and frequency steps is

$$\omega_0 t_0 \leq 2\pi. \tag{2.33}$$

The meaning of this condition is that if the Weyl-Heisenberg vectors are to cover the whole time-frequency plane, so that any signal in $L^2(\mathbf{R})$ can be represented in a numerically stable way as their superposition, then they have to be "sufficiently dense" in the space. The minimal density corresponds to $\omega_0 t_0 = 2\pi$, so that $2\pi/\omega_0 t_0$ represent redundancy of the frame. It turns out that in order to have good time-frequency localization we have to allow a certain degree of redundancy. This is a consequence of the Balian-Low theorem.

Theorem 4 *[1, 23]* Balian-Low: *If* $\varphi_{mn}(t) = e^{jm\omega_0 t}\varphi(t - nt_0)$, $m, n \in \mathbf{Z}$, *constitute a frame in* $L^2(\mathbf{R})$, *then either* $\int_{-\infty}^{\infty} t^2|\varphi(t)|^2 dt = \infty$ *or* $\int_{-\infty}^{\infty} \omega^2|\hat{\varphi}(\omega)|^2 d\omega = \infty$. □

In other words, there are no orthonormal Weyl-Heisenberg bases with good localization in the joint time-frequency domain.

Signal expansions based on Weyl-Heisenberg frames having a fixed resolution across the whole time-frequency plane cannot provide compact descriptions which capture both signal transients as well as large scale behavior. Frames of wavelets alleviate this problem. They consist of vectors which are obtained by dilating and translating a single mother wavelet $\psi(t)$ as

$$\left\{\psi_{mn}(t) : \psi_{mn}(t) = a_0^{-m/2}\psi(a_0^{-m}t - nb_0)\right\}_{m,n\in\mathbf{Z}}, \tag{2.34}$$

where $a_0 > 1$ and b_0 are positive constants. The distribution of vectors of a wavelet frame in the time-frequency plane is illustrated in Figure 2.4. The support of the wavelet $\psi_{mn}(t)$ in time is proportional to a_0^m. Going to higher frequencies (smaller values of m) wavelets are better concentrated and distributed with higher densities on the time scale. As a result, wavelet expansions provide signal descriptions with a progressively better time resolution at high frequencies. Hence, they provide an analysis tool which is well suited for extracting information on both fine signal structures (singularities, etc.) and global trends.

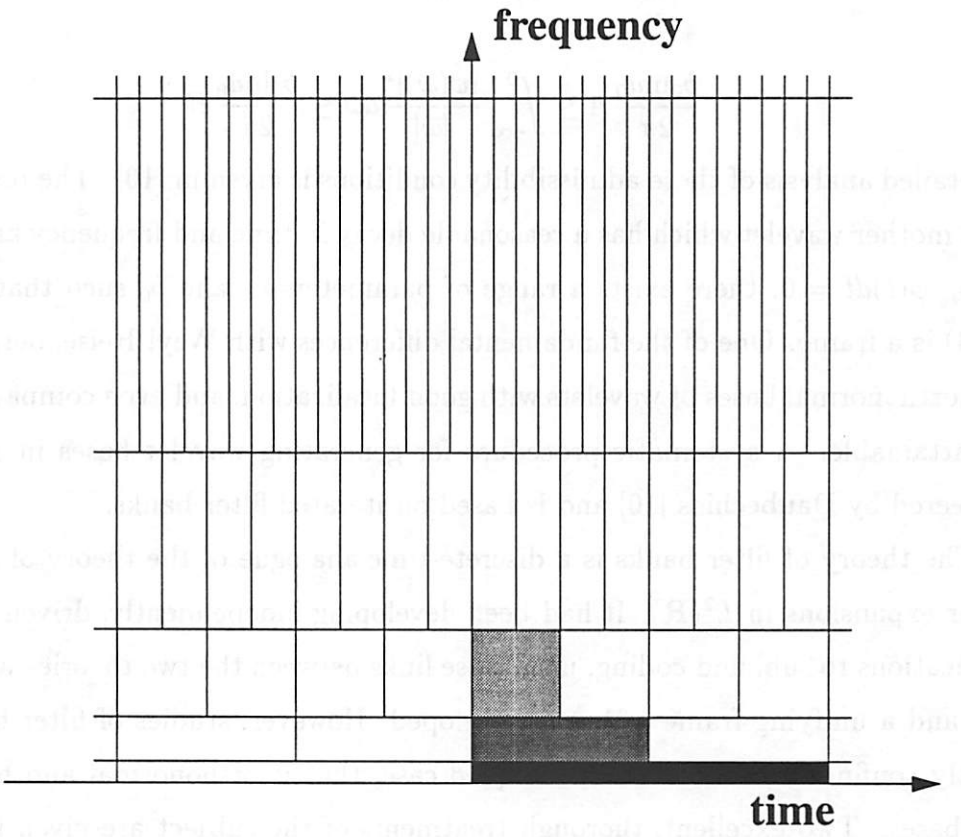

Figure 2.4: *The distribution of vectors of a wavelet frame in the time-frequency plane. Time resolution of the wavelet frame is progressively refined going to higher frequencies at the expense of coarser frequency resolution.*

The admissibility conditions on the mother wavelet $\psi(t)$ for

$$\psi_{mn}(t) = a_0^{-m/2}\psi(a_0^{-m}t - nb_0), \quad m,n \in \mathbf{Z},$$

to constitute a frame in $L^2(\mathbf{R})$ with bounds A and B are given by

$$\frac{b_0 \ln a_0}{2\pi} A \leq \int_0^\infty \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega \leq \frac{b_0 \ln a_0}{2\pi} B, \tag{2.35}$$

and

$$\frac{b_0 \ln a_0}{2\pi} A \leq \int_{-\infty}^0 \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega \leq \frac{b_0 \ln a_0}{2\pi} B. \tag{2.36}$$

A detailed analysis of these admissibility conditions is given in [10]. The result is that for a mother wavelet which has a reasonable decay in time and frequency and satisfies $\int_{-infty}^{\infty} \psi(t)dt = 0$, there exists a range of parameters a_0 and b_0 such that $\{\psi_{mn}\}$ in (2.34) is a frame. One of the fundamental differences with Weyl-Heisenberg frames is that orthonormal bases of wavelets with good localization, and even compact support, are attainable. A systematic procedure for generating wavelet bases in $L^2(\mathbf{R})$ was pioneered by Daubechies [10] and is based on iterated filter banks.

The theory of filter banks is a discrete-time analogue of the theory of structured linear expansions in $L^2(\mathbf{R})$. It had been developing independently, driven mainly by applications to subband coding, until close links between the two theories were unraveled and a unifying framework was developed. However, studies of filter banks were mostly confined to the critically sampled case, that is orthonormal and biorthonormal bases. Two excellent, thorough treatments of the subject are given in [47, 50]. We defer this topic to Chapter 3, which also presents original results on redundant structured expansions in $\ell^2(\mathbf{Z})$.

Chapter 3

Oversampled FIR Filter Banks

3.1 Filter Banks and Linear Expansions in $\ell^2(\mathbf{Z})$

The theory of filter banks [47, 50] provides a convenient framework for both the study and the implementation of an important class of signal decompositions in $\ell^2(\mathbf{Z})$. These are expansions underlying signal analysis through a sliding window using a selected set of elementary waveforms. In general, they have the form

$$x[n] = \sum_{i=0}^{K-1} \sum_{j=-\infty}^{\infty} c_{i,j} \varphi_{i,j}[n], \tag{3.1}$$

where the vectors $\varphi_{i,j}[n]$ denote translated versions of K basic functions, $\varphi_{i,j}[n] = \varphi_i[n - jN]$. The study of families of vectors of this type,

$$\Phi = \{\varphi_{i,j} : \varphi_{i,j}[n] = \varphi_i[n - jN], i = 0, 1, ..., K-1, j \in \mathbf{Z}\} \tag{3.2}$$

amounts to investigation of properties of K channel filter banks followed by subsampling by factor N in each of the channels, as the one shown in Figure 3.1a. In other words, vector families of this type are equivalent to filter banks.

Let the filters $H_0(z)$, $H_1(z)$, ... , $H_{K-1}(z)$ of the analysis filter bank, as the one shown in Figure 3.1a, be selected as complex conjugates of the time reversed versions of the waveforms φ_j, $h_i[n] = \tilde{\varphi}_i[n]$, $\quad i = 0, 1, ..., K-1$. For an input signal f, such a filter bank calculates inner products of f, $< x, \varphi_{i,j} >$, with vectors in the family

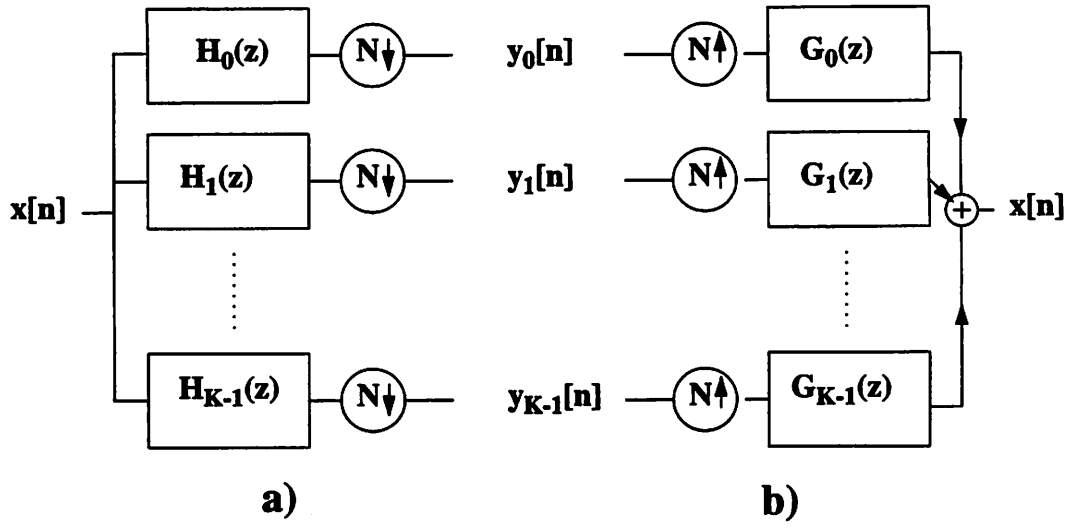

Figure 3.1: *A K channel filter bank with subsampling by factor N in the channels. a) Analysis filter bank. b) Synthesis filter bank.*

Φ. If the vector family Φ is a frame in $\ell^2(\mathbf{Z})$, then the filter bank implements the associated frame operator F, which is in this case a bounded linear operator from $\ell^2(\mathbf{Z})$ to $\ell^2(\{0,1,...,N-1\} \times \mathbf{Z})$,

$$F : \ell^2(\mathbf{Z}) \to \ell^2(\{0,1,...,N-1\} \times \mathbf{Z}). \tag{3.3}$$

We shall also say then that the filter bank implements a *frame decomposition* and refer to Φ as its *associated frame*, or a *filter bank frame*. In filter bank terms, Φ being a frame means that any signal can be reconstructed from the subband components, obtained at the outputs of the filter bank, using stable filters.

On the other hand, a synthesis filter bank as the one in Figure 3.1b, implements the Hilbert adjoint of the frame operator, F^*,

$$x'[n] = \sum_{i=0}^{K-1} \sum_{j=-\infty}^{\infty} y_i[j] \varphi_{i,j}[n] \tag{3.4}$$

provided that impulse responses of the filters $G_1(z)$, ... ,$G_{K-1}(z)$ match the waveforms φ_i, that is $g_i[n] = \varphi_i[n]$. Hence, for the stable filters $G_i(z)$, this filter bank is a bounded linear operator from $\ell^2(\{0,1,...,N-1\} \times \mathbf{Z})$ to $\ell^2(\mathbf{Z})$.

Expansions in $\ell^2(\mathbf{Z})$ with respect to frames Φ, as given in (3.2), are studied in detail in the critically sampled case, $N = K$, which correspond to orthonormal or

biorthonormal bases [50]. Results presented in the remainder of this chapter are concerned with redundant expansions of this type ($N < K$), that is oversampled filter banks, although most of them are valid in the critically sampled case as well.

3.2 Frame Conditions on Filter Banks

A convenient representation of signals and filters in analysis of multirate systems is the so-called *polyphase decomposition*. Consider a K channel filter bank, followed by subsampling by a factor N in the subbands (see Figure 3.1a). The suitable decomposition of an input signal $x[n]$ in this case, is in terms of N polyphase components,

$$X(z) = \sum_{i=0}^{N-1} z^{-i} X_i(z^N) \tag{3.5}$$

where

$$X_i(z) = \sum_{n=\infty}^{\infty} x[nN+i]z^{-n}. \tag{3.6}$$

Each polyphase component $x_i[n]$, given in the z-domain by $X_i(z)$, is obtained subsampling $x[n]$ by the factor N, starting with the i-th point.

Similarly, filters of the filter bank can be represented as

$$H_i(z) = \sum_{j=0}^{N-1} z^j H_{ij}(z^N), \tag{3.7}$$

where

$$H_{ij}(z) = \sum_{n=-\infty}^{+\infty} h_i[nN-j]z^{-n}. \tag{3.8}$$

It can be easily verified [50] that an input signal $x[n]$ produces at the output of this filter bank subband components $y_0[n]$, $y_1[n]$, ... , $y_{K-1}[n]$ which are given in terms of their z-transforms as

$$[Y_0(z)\ Y_1(z)\ \ldots\ Y_{K-1}(z)]^T = \mathbf{H_p}(z)[X_0(z)\ X_1(z)\ \ldots\ X_{N-1}(z)]^T. \tag{3.9}$$

In this expression $\mathbf{H_p}(z)$ denotes the polyphase analysis matrix of the filter bank,

which is defined as

$$\mathbf{H_p}(z) = \begin{bmatrix} H_{00}(z) & H_{01}(z) & \dots & H_{0(N-1)}(z) \\ H_{10}(z) & H_{11}(z) & \dots & H_{1(N-1)}(z) \\ \dots & \dots & \dots & \dots \\ H_{(K-1)0}(z) & H_{(K-1)1}(z) & \dots & H_{(K-1)(N-1)}(z) \end{bmatrix}. \tag{3.10}$$

Another way to look at the filter bank is as a bounded linear operator which maps the polyphase components of the input signal to the set of subband components, in the manner described by the relation (3.9). The input signal can be reconstructed from the subband components if and only if its polyphase components can be recovered after the transform given in (3.9). The frame conditions on an oversampled filter bank will be therefore expressed in terms of properties of its polyphase analysis matrix.

The necessary and sufficient frame and tight frame conditions are given by the following two theorems, which are proven in Appendix 3.7.1 and Appendix 3.7.2, respectively.

Theorem 5 *A filter bank implements a frame decomposition if and only if its polyphase analysis matrix is of full rank on the unit circle.* □

Theorem 6 *A filter bank implements a tight frame decomposition if and only if its polyphase analysis matrix is paraunitary,* $\tilde{\mathbf{H}}_\mathbf{p}(z)\mathbf{H_p}(z) = c\mathbf{I}$. □

An equivalent statement of Theorem 5 is that a filter bank implements a frame decomposition if and only if there exists a matrix $\mathbf{G_p}(z)$ of stable, rational, not necessarily causal, functions such that

$$\mathbf{G_p}(z)\mathbf{H_p}(z) = c\mathbf{I}. \tag{3.11}$$

The matrix $\mathbf{G_p}(z)$ is called the *synthesis polyphase matrix* and its entries $[\mathbf{G_p}(z)]_{i,j} = G_{ij}(z)$ are the polyphase components of filters of the synthesis filter bank (Figure 3.11b),

$$G_i(z) = \sum_{i=0}^{N-1} z^{-j} G_{ij}(z^N), \tag{3.12}$$

which can be used for perfect reconstruction of a signal from the decomposition obtained from the analysis filter bank. In other words, a synthesis filter bank gives a frame which is dual to the frame associated with the analysis filter bank. This dual frame consists of impulse responses of the filters $G_i(z)$ and their translates by integer multiples of N.

If the frame conditions given by Theorem 5 or Theorem 6 are satisfied, in the oversampled case $N < K$, the solution for $\mathbf{G_p}(z)$ of the polyphase equation (3.11), and hence the synthesis filter bank, is not unique. This is in accordance with the fact that for a given analysis frame, the corresponding synthesis frame is not unique. One solution for $\mathbf{G_p}(z)$ is the pseudoinverse of $\mathbf{H_p}(z)$, which is given by

$$\mathbf{H}^{+}(z) = \left(\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H_p}(z)\right)^{-1}\tilde{\mathbf{H}}_{\mathbf{p}}(z). \tag{3.13}$$

It can be easily verified that the frame associated to $\mathbf{G_p}(z) = \mathbf{H}^{+}(z)$ is the minimal dual of the frame associated with the analysis filter bank. Hence, it is important to investigate conditions under which both a filter bank frame and its minimal dual consist of finite length vectors. The following theorem states the necessary and sufficient condition for this to hold.

Theorem 7 *For a frame associated to an FIR filter filter bank, with the polyphase analysis matrix* $\mathbf{H_p}(z)$*, its dual minimal frame consists of finite length vectors if and only if* $\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H_p}(z)$ *is unimodular.* □

Sufficiency of this condition is obvious. It is proven in Appendix 3.7.3 that it is also necessary. This theorem is a generalization of the analogous result for the critically sampled filter banks. According to that result, perfect reconstruction with FIR filters after an analysis by a critically sampled FIR filter bank is possible if and only if the determinant of $\mathbf{H_p}(z)$ is a pure delay [50], which implies that $\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H_p}(z)$ has to be unimodular. Note that in the oversampled case, Theorem 7 does not preclude the existence of an FIR dual frame even if $\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H_p}(z)$ is not unimodular. However, for the reasons discussed in Chapter 2.3.2 we put the emphasize on the reconstruction using minimal dual frames.

Recall that we pointed out in Chapter 2.3.2 that, in general, for a given frame there is no closed form expression for its minimal dual. A remarkable property of the filter bank frames is that the task of finding a dual or the minimal dual amounts to finding a left inverse or a pseudoinverse, respectively, of a matrix of polynomials.

3.3 Parameterizations of Frames in $\ell^2(\mathbf{Z})$

The parameterization of filter bank frames which is given here is based on the Smith form of polynomial matrices [42]. Any polynomial matrix $\mathbf{H_p}(z)$ of dimension $K \times N$ $(K > N)$ can be decomposed as the product

$$\mathbf{H_p}(z) = \mathbf{R}(z)\mathbf{D}(z)\mathbf{C}(z), \tag{3.14}$$

where $\mathbf{R}(z)$ and $\mathbf{C}(z)$ are unimodular matrices of dimensions $K \times K$ and $N \times N$, respectively, while $\mathbf{D}(z)$ is a diagonal $K \times N$ polynomial matrix

$$\mathbf{D}(z) = \begin{bmatrix} d_1(z) & 0 & \dots & 0 \\ 0 & d_2(z) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_N(z) \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}. \tag{3.15}$$

The unimodular matrices can be chosen so that the polynomials $d_i(z)$ are monic and that $d_i(z)$ is a factor of $d_{i+1}(z)$. The matrix $\mathbf{D}(z)$ is called the *Smith form* of $\mathbf{H_p}(z)$. The unimodular matrices $\mathbf{R}(z)$ and $\mathbf{C}(z)$ are products of finitely many elementary matrices

$$\mathbf{R}(z) = \mathbf{R_1}(z)\mathbf{R_2}(z)\cdots\mathbf{R_m}(z)$$

$$\mathbf{C}(z) = \mathbf{C_1}(z)\mathbf{C_2}(z)\cdots\mathbf{C_n}(z).$$

Elementary matrices $\mathbf{R_i}(z)$, $\mathbf{C_j}(z)$ correspond to elementary row (column) operations, and have one of the following forms:

- a permutation matrix, i.e. the identity matrix with two rows permuted;
- a diagonal matrix with elements on the diagonal equal to unity, except for one which is equal to a nonzero scalar;
- a matrix with ones on the main diagonal and a single non-zero entry off the diagonal, which is a polynomial $\alpha(z)$.

An example of the three types of elementary matrices, for the 4×4 case, is given below.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \alpha(z) & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

A complete parameterization of FIR filter bank frames in $\ell^2(\mathbf{Z})$ follows directly from the Smith form, and is stated by the following Proposition.

Proposition 7 *An oversampled FIR filter bank implements a frame decomposition in $\ell^2(\mathbf{Z})$ if and only if the polynomials on the diagonal of the Smith form of its polyphase analysis matrix have no zeros on the unit circle.* □

This proposition follows from the fact that the polyphase matrix and its Smith form have the same rank on the unit circle, since they are related by elementary row (column) operations. Hence, the filter bank implements the frame decomposition if and only if its Smith form is of full rank everywhere on the unit circle, which holds if and only if the polynomials on the diagonal have no zeros on the unit circle.

An important class of filter bank frames are those which have minimal duals consisting of finite length vectors. According to Theorem 7, these are equivalent to polynomial matrices $\mathbf{H_p}(z)$ such that $\tilde{\mathbf{H}}_\mathbf{p}(z)\mathbf{H_p}(z)$ is unimodular. A parameterization of these frames is given by the following proposition.

Proposition 8 *Consider an oversampled FIR filter bank with the polyphase analysis matrix $\mathbf{H_p}(z)$. Then, $\tilde{\mathbf{H}}_\mathbf{p}(z)\mathbf{H_p}(z)$ is unimodular if $\mathbf{H_p}(z)$ has the following form:*

$$\mathbf{H_p}(z) = \mathbf{H_0}\mathbf{R}(z)\mathbf{D}(z)\mathbf{C}(z), \tag{3.16}$$

where the factors $\mathbf{H_0}$, $\mathbf{R}(z)$, $\mathbf{D}(z)$, $\mathbf{C}(z)$ *have the following form:*

- $\mathbf{H_0}$ - *a* $K \times N$ $(K > N)$ *matrix of scalars, such that* $\tilde{\mathbf{H}}_0\mathbf{H}_0 = c\mathbf{I}$,
- $\mathbf{R}(z)$ *and* $\mathbf{C}(z)$ - $N \times N$ *unimodular matrices, products of finitely many elementary matrices,*
- $\mathbf{D}(z)$ - *an* $N \times N$ *diagonal matrix of polynomials, with nonzero monomials on the diagonal.*

On the other hand, any unimodular parahermitian matrix of polynomials, $\mathbf{P}(z)$, *which is positive definite on the unit circle, can be factored as* $\mathbf{P}(z) = \tilde{\mathbf{H}}_\mathbf{p}(z)\mathbf{H}_\mathbf{p}(z)$, *where* $\mathbf{H}_\mathbf{p}(z)$ *is of the form given in (3.16).* □

This result is proven in Appendix 3.7.4.

As we noted in the previous section, for perfect reconstruction using an FIR filter bank after analysis with an oversampled FIR filter bank, $\tilde{\mathbf{H}}_\mathbf{p}(z)\mathbf{H}_\mathbf{p}(z)$ need not be unimodular. The necessary and sufficient condition for an FIR synthesis is given by the following proposition which is proven in Appendix 3.7.5. This proposition also gives a complete parameterization of frames associated to FIR filter banks which have duals consisting of finite length vectors.

Proposition 9 *Perfect reconstruction with FIR filters after analysis by an oversampled FIR filter bank is possible if and only if polynomials on the diagonal of the Smith form of the polyphase analysis matrix are monomials.* □

As it was shown in the previous subsection, tight filter bank frames are equivalent to paraunitary polynomial matrices. A $K \times N$ paraunitary matrix $(K > N)$ can always be embedded into a $K \times K$ paraunitary matrix [35]. The parameterization of the rectangular paraunitary polyphase matrices, that is filter bank tight frames in $\ell^2(\mathbf{Z})$, which we give in the following proposition follows directly from one of the factorizations of square paraunitary matrices studied by Vaidyanathan [47].

Proposition 10 *A $K \times N$ $(K > N)$ polynomial matrix $\mathbf{H_p}(z)$ is paraunitary if and only if it has the decomposition*

$$\mathbf{H_p}(z) = \mathbf{V_M}(z)\mathbf{V_{M-1}}(z)\cdots\mathbf{V_1}(z)\mathbf{H_0}. \tag{3.17}$$

The elementary building blocks, $\mathbf{V_i}(z)$, have the following form,

$$\mathbf{V_i}(z) = \mathbf{I} - \mathbf{v_i}\mathbf{v_i}^* + z^{-1}\mathbf{v_i}\mathbf{v_i}^*, \tag{3.18}$$

where $\mathbf{v_i}$ denotes a unit norm vector, while $\mathbf{H_0}$ is a $K \times N$ matrix of scalars such that $\tilde{\mathbf{H}}_\mathbf{0}\mathbf{H_0} = c\mathbf{I}$. □

The scope of frames in $\ell^2(\mathbf{Z})$ which can be derived from filter banks goes beyond families of vectors of the type given in (3.2). An abundance of frames can be generated by iteratively growing filter bank trees, in the manner used for generating wavelet packets [51]. The significance of iterated filter structures is in the fact that they can generate waveforms with almost any reasonable localization in the time-frequency plane which allows for signal analysis with a great flexibility of time or frequency resolution.

3.4 Nonsubsampled Filter Banks

3.4.1 Frame Conditions on Nonsubsampled Filter Banks

The motivation for the study of nonsubsampled filter banks stems from applications where critical sampling is not needed but rather imposes severe design constraints. Signal analysis and modeling of biological systems are examples of such applications. Nonsubsampled filter banks also find their place in applications which require shift invariant signal representations, which is known to be inconsistent with the subsampling in the filter bank channels [41]. Highly redundant representations followed by a sophisticated sampling strategy can even yield compression schemes. Examples are wavelet modulus maxima and wavelet zero-crossings representations [25, 26] which are proposed for singularity detection, signal denoising and compression,

and are based on preprocessing by nonsubsampled filter banks. The crucial property of the nonsubsampled filter banks in these applications is that they produce dense sequences of samples resulting in representations which are close approximations of continuous-time transforms.

In the case of nonsubsampled filter banks, the frame conditions as well as the condition for the feasibility of an FIR synthesis after an FIR analysis have a peculiar form. These conditions are reviewed in this section. The flexibility of the design of nonsubsampled filter banks is in this thesis illustrated for the case of maximally flat two channel filter banks which implement tight frame decompositions in $\ell^2(\mathbf{Z})$. These design considerations are deferred to Chapter 5.3.3, where applications of non-subsampled filter banks are discussed.

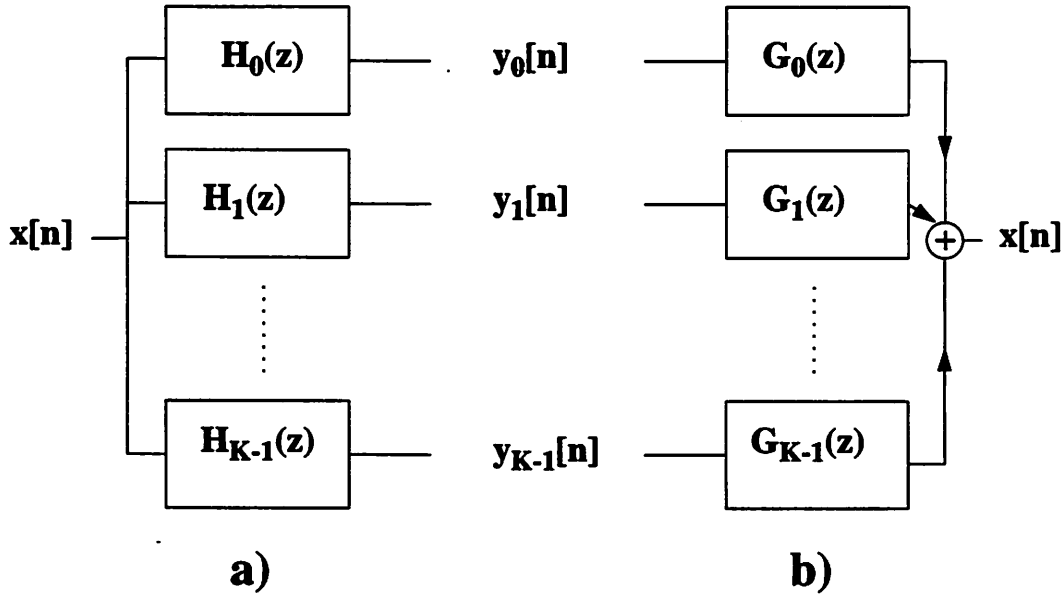

Figure 3.2: *A K channel nonsubsampled filter bank. a) Analysis filter bank. b) Synthesis filter bank.*

The polyphase analysis matrix of a nonsubsampled filter bank (Figure 3.2a) is a column vector whose entries are the analysis filters themselves,

$$\mathbf{H_p}(z) = [H_0(z)\ H_1(z)\ \ldots\ H_{K-1}(z)]^T. \tag{3.19}$$

Perfect and stable reconstruction is possible provided that there exist stable filters

$G_0(z)$, $G_1(z)$, ... , $G_{K-1}(z)$ satisfying

$$H_0(z)G_0(z) + H_1(z)G_1(z) + ... + H_{K-1}(z)G_{K-1}(z) = 1. \tag{3.20}$$

This equation expresses reconstruction by a synthesis filter bank with filters $G_i(z)$ (Figure 3.2b). The necessary and sufficient condition for the existence of such filters is given by the following corollary of Theorem 5.

Corollary 2 *A nonsubsampled filter bank implements a frame decomposition if and only if its analysis filters have no zeros in common on the unit circle.* □

The frame condition does not generally guarantee the possibility of FIR reconstruction.

Corollary 3 *Perfect reconstruction using FIR filters after an FIR analysis by a nonsubsampled filter bank is possible if and only if the analysis filters have no zeros in common.* □

This result is a corollary of Proposition 9. As a special case of Theorem 6 we have the following result about nonsubsampled filter banks and tight frames.

Corollary 4 *A nonsubsampled filter bank implements a tight frame decomposition if and only if its analysis filters are power complementary:*

$$H_0(z)\tilde{H}_0(z) + H_1(z)\tilde{H}_1(z) + ... + H_{K-1}(z)\tilde{H}_{K-1}(z) = 1. \tag{3.21}$$

□

In the case of nonsubsampled filter banks, a frame associated with an FIR filter bank has the minimal dual consisting of finite length vectors only if the frame is tight. This result is an immediate corollary of Theorem 7.

Corollary 5 *For a frame associated with an FIR nonsubsampled filter bank, its minimal dual frame consists of finite length vectors if and only if the analysis filters are power complementary, that is if and only if the frame is tight.* □

3.4.2 Nonsubsampled Filter Banks and Continuous-Time Signal Analysis

Filter banks, as it will be shown here, give samples of continuous time transforms of signals in $L^2(\mathbf{R})$, provided that an appropriate discrete-time version of the signal is available at the input. The usefulness of nonsubsampled filter banks comes from the fact that the highly redundant representations they generate can be close discrete-time approximations of continuous-time signal transforms. In this section we identify the underlying continuous-time filters.

Discretization of a continuous-time signal $f \in L^2(\mathbf{R})$ usually amounts to projecting it onto an *approximation space* V_φ, which is spanned by integer translates of a single function $\varphi(x)$, called *the generating function.* The generating function should satisfy a condition

$$0 < \alpha \leq \sum_{k \in \mathbf{Z}} |\hat{\varphi}(\omega + 2k\pi)|^2 \leq \beta < \infty \tag{3.22}$$

which means that its integer translates constitute a Riesz basis for V_φ. The projection, $f_p(x)$, of an $f(x)$ onto V_φ, can be represented by a sequence of parameters $f_\phi[n]$ which are coefficients of its expansion:

$$f_p(x) = \sum_{n \in \mathbf{Z}} f_\phi[n]\varphi(x - n). \tag{3.23}$$

The sequence $f_\phi[n]$ represents a discrete-time version of the signal, and is obtained by sampling $f(x)$ prefiltered by an appropriate filter $\phi(x)$,

$$f_\phi[n] = \int_{-\infty}^{+\infty} f(x)\phi(n - x)dx. \tag{3.24}$$

Integer translates of the time-reversed version of $\phi(x)$ constitute another bases for V_φ which is dual to $\{\varphi(x - n), n \in \mathbf{Z}\}$ [46] so that $f_p(x)$ can be alternatively represented as

$$f_p(x) = \sum_{n \in \mathbf{Z}} f_\varphi[n]\phi(n - x), \tag{3.25}$$

where

$$f_\varphi[n] = \int_{-\infty}^{+\infty} f(x)\varphi(x - n)dx. \tag{3.26}$$

Shannon's sampling occurs when the *approximation space* is the space of bandlimited signals generated by translates of the *sinc* function. Another particular case is the V_0 space of Mallat's multiresolution analysis, with $\varphi(x)$ being the corresponding *scaling function*. For detailed analysis of various aspects of generalized sampling the reader is referred to work by Unser and Aldroubi [46].

For a continuous-time signal $f(x)$ discretized as in (3.24) consider subsequent processing by the nonsubsampled filter bank which consists of filters $H_0(z)$, $H_1(z)$, ... , $H_{K-1}(z)$. Subband components f_k, $k = 1, 2, \ldots K-1$ generated by the filter bank are given in the Fourier domain by the following formula,

$$\hat{f}_k(e^{j\omega}) = H_k(e^{j\omega}) \sum_{k \in \mathbf{Z}} \hat{f}(\omega + 2k\pi)\hat{\phi}(\omega + 2k\pi). \tag{3.27}$$

This expression indicates that the filter bank performs regular sampling of signals obtained by filtering $f(x)$ by a set of filters $\varphi_0(x)$, $\varphi_1(x)$, ..., $\varphi_{K-1}(x)$, which are given by

$$\hat{\varphi}_k(\omega) = H_k(e^{j\omega})\hat{\phi}(\omega). \tag{3.28}$$

This characterizes continuous-time transforms which underlie nonsubsampled filter bank analysis in the given approximation space V_φ. An interesting particular case is wavelet analysis implemented by iterated two channel filter banks. It is discussed in more detail in the next section.

3.5 Wavelet Analysis in $\ell^2(\mathbf{Z})$

3.5.1 Wavelet Bases and Critically Sampled Filter Banks

Wavelet bases in $\ell^2(\mathbf{Z})$ are derived from critically sampled iterated two-channel filter banks. Such a filter bank is shown in Figure 3.3. The corresponding wavelet family is given by

$$\left\{\psi_{mn}[i] : \psi_{mn}[i] = \psi_m[i - n2^{m+1}], m = 0, 1, ..., J-1, \psi_{Jn}[i] = \psi_J[i - n2^J]\right\}_{n \in \mathbf{Z}}, \tag{3.29}$$

where $\psi_m[n]$ is the complex conjugate of the time-reversed version of the equivalent filter from the input of the filter bank to the m-th output. In particular, if $H_0(z)$ and $H_1(z)$ denote the prototype filters of the filter bank, then the equivalent filters are

$$\begin{aligned} & H_1(z), \\ & H_0(z)H_1(z^2), \\ & \cdots \\ & H_0(z)H_0(z^2)\cdots H_0(z^{2^{J-2}})H_1(z^{2^{J-1}}), \\ & H_0(z)H_0(z^2)\cdots H_0(z^{2^{J-1}}). \end{aligned} \tag{3.30}$$

$\{\psi_{mn}\}$ is a basis if and only if the elementary building block of the iterated filter bank, that is $H_0(z)$ and $H_1(z)$ followed by subsampling by 2, constitute a so-called *perfect reconstruction filter bank*. This condition is equivalent to the requirement that the determinant of the polyphase matrix of the building block has to be a pure delay, that is $\det\mathbf{H_p}(z) = z^{-l}$ [50]. The polyphase analysis matrix is given by 3.8 and 3.10, for $N = K = 2$. Filters $H_0(z)$ and $H_1(z)$ which constitute an orthonormal

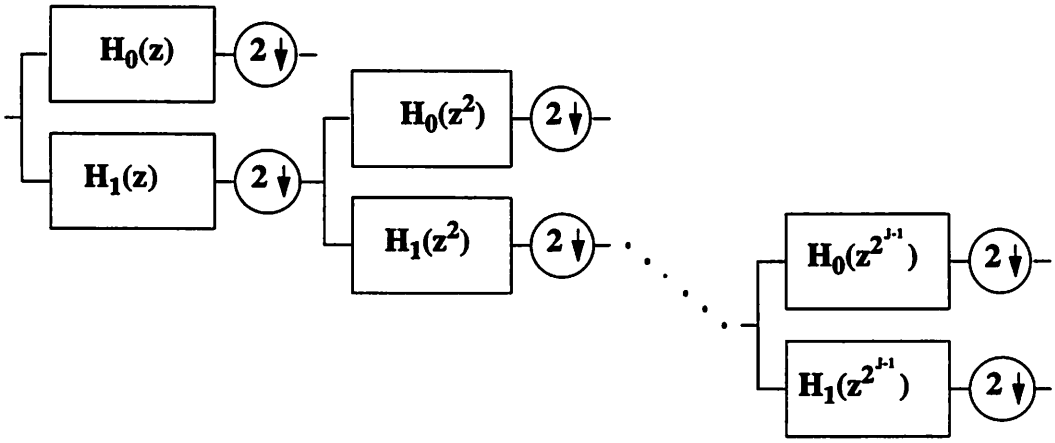

Figure 3.3: *Two-channel iterated critically sampled filter bank.*

filter bank give a family ψ_{mn} which is an orthonormal wavelet basis in $\ell^2(\mathbf{Z})$. The orthonormality is equivalent to paraunitariness of the polyphase analysis matrix, and can be further expressed as the conjunction of the following two conditions [47, 50]

$$H_0(z)H_0(z^{-1}) + H_0(-z)H_0(-z^{-1}) = 2, \tag{3.31}$$

$$H_1(z) = -z^{2k+1}H_0(-z^{-1}), \quad k \in \mathbf{Z}. \tag{3.32}$$

The analogy between wavelet bases in the discrete and the continuous-time domain comes from the same type of tiling of the time-frequency plane (see Figure 2.3), although one can not speak of discrete-time wavelet bases, as given by (3.29), which are obtained by dilating and translating a single filter in time. The most striking link between these discrete-time operators and wavelets in $L^2(\mathbf{R})$ is the fact that there is a systematic procedure for generating wavelet bases in $L^2(\mathbf{R})$ from iterated filter banks. This procedure was originally proposed by Daubechies [11] for generating compactly supported wavelets from FIR orthogonal filter banks, and subsequently verified to be valid for general FIR perfect reconstruction filter banks [49, 8]. The details of the construction are as follows.

Consider a two channel octave band iterated filter bank of depth i, as the one shown in Figure 3.3. The equivalent filters of the two lowest branches are

$$H_0^{(i)}(z) = \prod_{k=0}^{i-1} H_0\left(z^{2^k}\right), \tag{3.33}$$

$$H_1^{(i)}(z) = H_1\left(z^{2^{i-1}}\right) \prod_{k=0}^{i-2} H_0\left(z^{2^k}\right). \tag{3.34}$$

Then associate with the impulse responses $h_0^{(i)}$ and $h_1^{(i)}$ of $H_0^{(i)}(z)$ and $H_1^{(i)}(z)$ the continuous-time functions $\phi^{(i)}(x)$ and $\psi^{(i)}(x)$:

$$\phi^{(i)}(x) = 2^{\frac{i}{2}} h_0^{(i)}(n), \ \frac{n}{2^i} \leq x < \frac{n+1}{2^i}, \tag{3.35}$$

$$\psi^{(i)}(x) = 2^{\frac{i}{2}} h_1^{(i)}(n), \ \frac{n}{2^i} \leq x < \frac{n+1}{2^i}. \tag{3.36}$$

The elementary interval is divided by 2^i so that the continuous-time functions remain compactly supported as $i \to \infty$. The factor $2^{\frac{i}{2}}$ which multiplies $h_0^{(i)}$ and $h_1^{(i)}$ is needed in order to keep the L^2 norms of ϕ^i and ψ^i inside finite bounds. Assume that $\phi^{(i)}(x)$ and $\psi^{(i)}(x)$ converge in the L^2 sense to the limits

$$\phi(x) = \lim_{i \to \infty} \phi^{(i)}(x), \tag{3.37}$$

$$\psi(x) = \lim_{i \to \infty} \psi^{(i)}(x), \tag{3.38}$$

which are piecewise smooth functions in $L^2(\mathbf{R})$. For an in-depth review of the conditions which ensure this convergence the reader is referred to [10]. The function $\psi(x)$

obtained as the limit of the above construction is the mother wavelet of a wavelet basis, while $\phi(x)$ is the so-called *scaling function* associated with the basis.

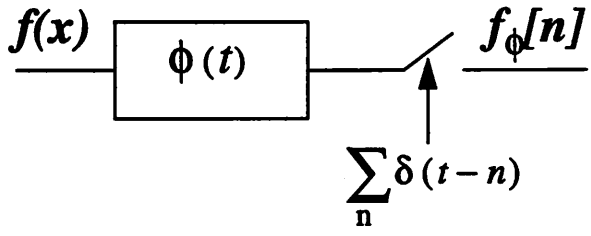

Figure 3.4: *Discretization of a continuous-time signal for discrete-time implementation of the Continuous Wavelet Transform.*

Not only can a wavelet be derived from a two channel filter bank, but a filter bank can be used for the computation of the coefficients of the corresponding wavelet expansion. This result is known as Mallat's algorithm [24]. The algorithm requires that a continuous-time signal $f(x)$ be discretized as in (3.24), which amounts to filtering with the scaling function and sampling at points $t = ..., -1, 0, 1, 2, ...$ (see Figure 3.4). If such a discrete-time version of $f(x)$ is fed into the iterated filter bank, then at the k-th output of the filter bank we obtain the sequence $f_k[n]$ which gives inner products of $f(x)$ with a dilated version $\frac{1}{2^{k/2}}\psi\left(\frac{x}{2^k}\right)$ of the wavelet $\psi(x)$. Precisely,

$$f_k[n] = \int_{-\infty}^{\infty} f(x)\frac{1}{2^{k/2}}\psi\left(n - \frac{x}{2^k}\right)dx. \tag{3.39}$$

This formula was originally proven based on orthogonality of the wavelet $\psi(x)$ with respect to its dilates and translates. However, it follows immediately from the considerations in the previous section (see formula (3.28)) and the wavelet construction.

3.5.2 Nonsubsampled Filter Banks and Wavelet Analysis

The major advantage of wavelet analysis over Fourier and Short-Time Fourier transforms is that it can isolate events such as singular signal points or other fine details. Strictly speaking this is the property of the continuous wavelet transform (CWT). The continuous wavelet transform of a signal $f(x)$ is defined as

$$Wf(s,x) = f * \psi_s(x), \tag{3.40}$$

where $\psi_s(x) = \frac{1}{\sqrt{s}}\psi\left(\frac{x}{s}\right)$. Critically sampled octave-band filter banks, as described in the previous subsection, give samples of the CWT at the grid

$$(s,x) \in \{(2^i, 2^i j) \ : \ i = 1,2,\ldots,J \ j \in \mathbf{Z}\}, \tag{3.41}$$

where J is the depth of the filter bank tree. For applications such as singularity detection it is advantageous to remove the subsampling from the filter bank channels in order to obtain samples of the CWT at a denser grid.

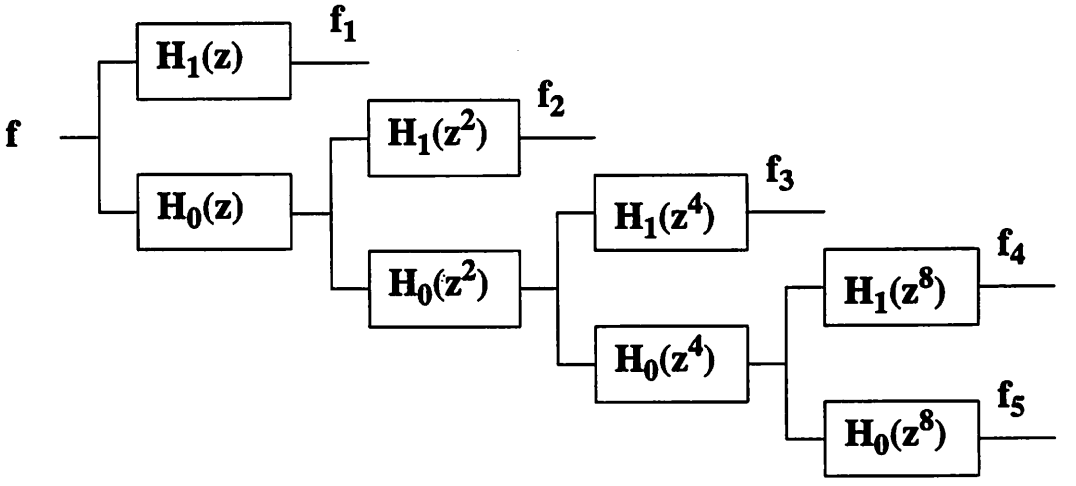

Figure 3.5: *An octave-band nonsubsampled iterated filter bank for discrete-time wavelet transform, of depth* $J = 4$.

A nonsubsampled filter bank, as illustrated in Figure 3.5 produces samples of the CWT at the grid

$$(s,x) \in \{(2^i, j) \ : \ i = 1,2,\ldots,J \ j \in \mathbf{Z}\}. \tag{3.42}$$

Note that in addition to the different sampling grid there is also a slight difference in the construction of the wavelet with respect to the critically sampled case, which amounts to different scaling factors. In the nonsubsampled case, the wavelet $\psi(x)$ and the scaling function $\phi(x)$ are obtained as the limits of

$$\phi^{(i)}(x) = 2^i h_0^{(i)}[n], \ \frac{n}{2^i} \le x < \frac{n+1}{2^i}, \tag{3.43}$$

$$\psi^{(i)}(x) = 2^i h_1^{(i)}[n], \ \frac{n}{2^i} \le x < \frac{n+1}{2^i}, \tag{3.44}$$

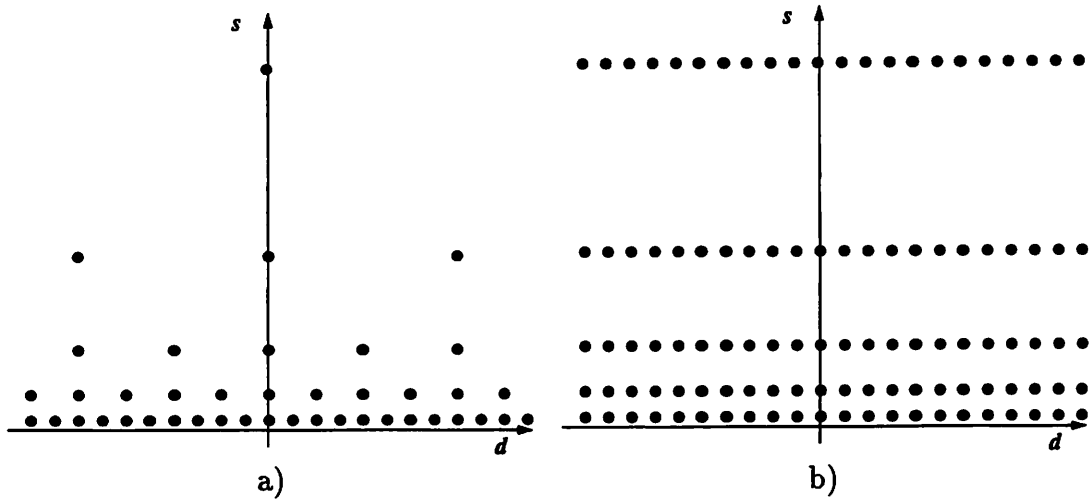

Figure 3.6: *Sampling grids of the CWT implemented by discrete-time processing. a) Grid corresponding to two-channel critically sampled iterated filter banks. b) Grid corresponding to two-channel nonsubsampled iterated filter banks.*

where $h_0^{(i)}$ and $h_1^{(i)}$ are again impulse responses of iterated filters

$$H_0^{(i)}(z) = \prod_{k=0}^{i-1} H_0\left(z^{2^k}\right), \tag{3.45}$$

$$H_1^{(i)}(z) = H_1\left(z^{2^{i-1}}\right) \prod_{k=0}^{i-2} H_0\left(z^{2^k}\right). \tag{3.46}$$

The wavelet transform which underlies the analysis with the nonsubsampled filter bank has the same form except for the scaling factor and is given by

$$Wf(s,x) = f * \psi_s(x), \tag{3.47}$$

where $\psi_s(x) = \frac{1}{s}\psi\left(\frac{x}{s}\right)$.

Besides giving the CWT on a denser grid, nonsubsampled filter banks allow for more flexible design than the critically sampled ones. The design freedom attained by nonsubsampled filter banks is demonstrated in Chapter 5.3.3 through a design of highly regular wavelets with a given number of vanishing moments which are used for singularity detection and discrimination.

3.6 Short-Time Fourier Analysis in $\ell^2(\mathbf{Z})$

3.6.1 Modulated Filter Banks

Along with the development of Gabor's original scheme of the short-time Fourier representations, i.e. Weyl-Heisenberg expansions, it was observed in the 1980's that such time-frequency representations can be stable only in the overcomplete case [2, 21]. Another incarnation of the same phenomenon is expressed by the Balian-Low theorem [1, 23] which implies that there are no orthogonal Weyl-Heisenberg bases which have good localization in both time and frequency. However, as soon as some redundancy is introduced, the picture changes drastically and, as demonstrated by Daubechies [10], good localization of tight Weyl-Heisenberg frames is attainable.

In $\ell^2(\mathbf{Z})$ Weyl-Heisenberg frames have the form

$$\Phi = \{\varphi_{lm} : \varphi_{lm}[n] = e^{j\frac{2\pi}{K}ln}\varphi[n - mN]\}, \tag{3.48}$$

where $\varphi[n]$ denotes the prototype window function. Note that redundant families of this type correspond to $K > N$, while the Weyl-Heisenberg bases are obtained for $K = N$. Such a family of vectors is equivalent to a K channel modulated filter bank, with subsampling by N in the channels (see Figure 3.1). Filters of the equivalent filter bank are modulates of a prototype filter $V(z)$, $H_i(z) = V(e^{j\frac{2\pi i}{K}}z)$, which is given in the time domain as the complex conjugate of the time-reversed version of $\varphi[n]$. A fact about Weyl-Heisenberg bases in $\ell^2(\mathbf{Z})$ analogous to the one described by the Balian-Low theorem has been observed by Vetterli [48].

Theorem 8 *[48] The Weyl-Heisenberg family given in (3.48) derived from a finite-support $\varphi[n]$ is a basis in $\ell^2(\mathbf{Z})$ if and only if $\varphi[n]$ has exactly N nonzero coefficients.*

□

A consequence of this result is that there are no critically sampled modulated filter banks with finite impulse responses which exhibit good frequency selectivity. The purpose of this section is to investigate the existence of tight Weyl-Heisenberg frames in $\ell^2(\mathbf{Z})$, with finite support in time and good localization in frequency.

In the following, a complete parameterization of tight Weyl-Heisenberg frames in $\ell^2(\mathbf{Z})$ with arbitrary rational oversampling ratios is given within the framework of modulated filter banks. Oversampled filter banks are much less constrained than critically sampled ones and consequently allow for the design of frequency selective FIR filters, which can be linear-phase as well. In other words, there exist tight Weyl-Heisenberg frames with symmetries and good localization in the time-frequency domain.

Consider a K channel modulated filter bank, based on the prototype filter $V(z)$, with the subsampling factor $N < K$. Let M be the least common multiple of K and N and let J and L be the two integers satisfying $JK = LN = M$. The M-component polyphase representation of $V(z)$ has the form:

$$V(z) = \sum_{j=0}^{M-1} z^{-j} V_j(z^M). \tag{3.49}$$

Elements of the polyphase analysis matrix can be expressed as

$$H_{ij}(z) = \sum_{l=0}^{L-1} W_K^{-i(j+lN)} z^{-l} V_{j+lN}(z^L), \tag{3.50}$$

where $W_K = exp(j2\pi/K)$. This gives the following factorization:

$$\mathbf{H_p}(z) = \mathbf{F}_K \mathbf{V}(z), \tag{3.51}$$

where $\mathbf{F}_K$ is the $K \times K$ DFT matrix and

$$\mathbf{V}(z) = [I_K ... I_K] \cdot \mathrm{diag}\left(V_0(z^L), V_1(z^L), ..., V_{M-1}(z^L)\right) \begin{bmatrix} I_N \\ z^{-1} I_N \\ \cdots \\ z^{-(L-1)} I_N \end{bmatrix}. \tag{3.52}$$

I_n here stands for the $n \times n$ identity matrix.

By inspection of the above factorization one can see that the elements of $\mathbf{V}(z)$ are given by

$$[\mathbf{V}(z)]_{ij} = z^{-q(i,j)} V_{i+p(i,j)K}(z^L) \quad i = 0, 1, ..., K-1, \quad j = 0, 1, ..., N-1 \tag{3.53}$$

where $p(i,j)$ and $q(i,j)$ are integers satisfying

$$i + p(i,j)K = j + q(i,j)N, \quad p(i,j) \leq J-1, \quad q(i,j) \leq L-1. \tag{3.54}$$

Note that equation (3.54) cannot be satisfied for every pair of integers i and j. In fact, for each j there are exactly J indices, i, which satisfy (3.54). Consequently J nonzero elements are regularly spread in each row of $\mathbf{V}(z)$ and L nonzero elements are regularly spread in each column of $\mathbf{V}(z)$. Three possible cases are illustrated by the following examples.

Example 1 *K is a multiple of N. In this case, $J = 1$, so that there is a single nonzero element in each row of $\mathbf{V}(z)$. For $K = 6$ and $N = 3$ we have:*

$$\mathbf{V}(z) = \begin{bmatrix} V_0(z^2) & 0 & 0 \\ 0 & V_1(z^2) & 0 \\ 0 & 0 & V_2(z^2) \\ z^{-1}V_3(z^2) & 0 & 0 \\ 0 & z^{-1}V_4(z^2) & 0 \\ 0 & 0 & z^{-1}V_5(z^2) \end{bmatrix}.$$

□

Example 2 *K and N are coprime. In this case, $J = N$, so that all elements of $\mathbf{V}(z)$ are nonzero. For $K = 3$ and $N = 2$ we obtain:*

$$\mathbf{V}(z) = \begin{bmatrix} V_0(z^3) & z^{-1}V_3(z^3) \\ z^{-2}V_4(z^3) & V_1(z^3) \\ z^{-1}V_2(z^3) & z^{-2}V_5(z^3) \end{bmatrix}.$$

□

Example 3 *Neither of the above cases, e.g. $K = 6$ and $N = 4$ yields $\mathbf{V}(z)$ equal to*

the following matrix:

$$\begin{bmatrix} V_0(z^3) & 0 & z^{-1}V_6(z^3) & 0 \\ 0 & V_1(z^3) & 0 & z^{-1}V_7(z^3) \\ z^{-2}V_8(z^3) & 0 & V_2(z^3) & 0 \\ 0 & z^{-2}V_9(z^3) & 0 & V_3(z^3) \\ z^{-1}V_4(z^3) & 0 & z^{-2}V_{10}(z^3) & 0 \\ 0 & z^{-1}V_5(z^3) & 0 & z^{-2}V_{11}(z^3) \end{bmatrix}$$

□

3.6.2 Tight Weyl-Heisenberg Frames in $\ell^2(\mathbf{Z})$

As it was stated earlier in this chapter (see Theorem 6 in Section 3.2), a filter bank implements a tight frame decomposition if and only if its polyphase analysis matrix is paraunitary. According to the definitions of $\mathbf{H_p}(z)$ and $\mathbf{V}(z)$ the paraunitariness condition can be expressed also as:

$$\tilde{\mathbf{V}}(z)\mathbf{V}(z) = I_N.$$

Noting that some of the elements of $\tilde{\mathbf{V}}(z)\mathbf{V}(z)$ are identically zero and taking symmetries into account, it can be observed that the paraunitariness condition imposes $N + \frac{N(J-1)}{2}$ distinct constraints. We now investigate the implications of these constraints for the three cases presented in the previous subsection.

Case 1 *K is a multiple of N*
$V(z)$ is constrained only by

$$\sum_{l=0}^{L-1} \tilde{V}_{i+lN}(z)V_{i+lN}(z) = c, \quad i = 0, 1, ..., N-1 \qquad (3.55)$$

This indicates that the polyphase components

$$\{V_i(z), V_{i+N}(z), ..., V_{i+(L-1)N}(z)\},$$

for each $i = 0, ..., N-1$ are power complementary. It follows that they are the polyphase components of a filter $F_i(z)$, which is orthogonal to its translates by multiples

of L. Hence, tight Gabor frames with an integer oversampling ratio K/N are generated by arbitrarily selecting N orthogonal filters $F_i(z)$ and identifying their polyphase components with the polyphase components of $V(z)$. □

Case 2 *K and N are coprime*

In this case, the polyphase components of $V(z)$ are, up to a time delay, equal to entries of a full $K \times N$ paraunitary matrix. In order to show this we need to transform $\mathbf{V}(z)$. Let $\mathbf{D}_r(z)$ and $\mathbf{D}_l(z)$ be diagonal matrices of monomials, selected so that elements in the first row and column of $\mathbf{D}_l(z)\mathbf{V}(z)\mathbf{D}_r(z)$ are equal to polyphase components of $\mathbf{V}(z)$. These matrices are given by $\mathbf{D}_l(z) = \mathrm{diag}(1, z^{q(1,0)}, \ldots, z^{q(K-1,0)})$ and $\mathbf{D}_r(z) = \mathrm{diag}(1, z^{q(0,1)}, \ldots, z^{q(0,N-1)})$. For instance, for matrix $\mathbf{V}(z)$ in Example 2, $\mathbf{D}_l(z) = \mathrm{diag}(1, z^2, z)$, $\mathbf{D}_r(z) = \mathrm{diag}(1, z)$, giving

$$\mathbf{D}_l(z)\mathbf{V}(z)\mathbf{D}_r(z) = \mathbf{U}(z^L) = \begin{bmatrix} V_0(z^3) & V_3(z^3) \\ V_4(z^3) & z^3 V_1(z^3) \\ V_2(z^3) & V_5(z^3) \end{bmatrix}.$$

In general, entries of $\mathbf{D}_l(z)\mathbf{V}(z)\mathbf{D}_r(z)$ are equal to

$$[\mathbf{D}_l(z)\mathbf{V}(z)\mathbf{D}_r(z)]_{i,j} = z^{-q(i,j)+q(i,0)+q(0,j)} V_{i+p(i,j)K}(z^L). \tag{3.56}$$

It follows from (3.54) that

$$(-p(i,j) + p(i,0) + p(0,j))\, K = (-q(i,j) + q(i,0) + q(0,j))\, N,$$

which implies that $-q(i,j) + q(i,0) + q(0,j)$ can be only 0 or L. Therefore we can write

$$\mathbf{D}_l(z)\mathbf{V}(z)\mathbf{D}_r(z) = \mathbf{U}(z^L), \tag{3.57}$$

where $\mathbf{U}(z)$ is a matrix whose nonzero entries are up to a delay equal to polyphase components of $V(z)$. Therefore, an arbitrary paraunitary $K \times N$ matrix $\mathbf{U}(z)$ gives a window function $V(z)$ of a tight Gabor frame. Any FIR tight Weyl-Heisenberg frame in $\ell^2(\mathbf{Z})$ with the oversampling ratio K/N can be obtained in this manner. □

Case 3 *Neither of the above cases*

It can be easily observed that paraunitariness of $\mathbf{V}(z)$ is equivalent to paraunitariness of its submatrices $\mathbf{V}_i(z)$, $i = 0, ..., N/J - 1$, each of dimension $L \times J$. Note that L and J are coprime and all submatrices $\mathbf{V}_i(z)$ have the same structure, in terms of the distribution of delay elements, as a $\mathbf{V}(z)$ matrix of the size $L \times J$. For instance, in Example 4, $\mathbf{V}(z)$ is paraunitary if and only if its submatrices

$$\mathbf{V}_0(z) = \begin{bmatrix} V_0(z^3) & z^{-1}V_6(z^3) \\ z^{-2}V_8(z^3) & V_2(z^3) \\ z^{-1}V_4(z^3) & z^{-2}V_{10}(z^3) \end{bmatrix} \qquad \mathbf{V}_1(z) = \begin{bmatrix} V_1(z^3) & z^{-1}V_7(z^3) \\ z^{-2}V_9(z^3) & V_3(z^3) \\ z^{-1}V_5(z^3) & z^{-2}V_{11}(z^3) \end{bmatrix} \tag{3.58}$$

are paraunitary (compare these submatrices with $\mathbf{V}(z)$ in Example 3). According to considerations in Case 1, it follows that the paraunitariness of the submatrices $\mathbf{V}_i(z)$ is equivalent to the paraunitariness of some matrices $\mathbf{U}_i(z)$, $i = 0, ..., N/J - 1$, whose entries are up to a delay equal to polyphase components of $V(z)$. For the matrix $\mathbf{V}(z)$ in Example 4, corresponding matrices $\mathbf{U}_i(z)$ are given by

$$\mathbf{U}_0(z) = \begin{bmatrix} V_0(z) & V_6(z) \\ V_8(z) & zV_2(z) \\ V_4(z) & V_{10}(z) \end{bmatrix} \qquad \mathbf{U}_1(z) = \begin{bmatrix} V_1(z) & V_7(z) \\ V_9(z) & zV_3(z) \\ V_5(z) & V_{11}(z) \end{bmatrix} \tag{3.59}$$

Therefore, a prototype lowpass filter $\mathbf{V}(z)$ of any tight Gabor frame with the oversampling ratio K/N can be obtained by starting with arbitrary N/J paraunitary matrices of the size $L \times J$ and identifying their entries, up to a delay, with M polyphase components of $V(z)$. □

The results presented in this section give a complete parameterization of FIR tight Gabor frames in $\ell^2(\mathbf{Z})$ for arbitrary rational oversampling ratios. Design of the tight frames then amounts to an optimization procedure under these constraints. Recall that, in the case of perfect reconstruction FIR modulated filter banks with critical sampling, the polyphase components of $V(z)$ can not have more than one nonzero coefficient [48], which is too restrictive to obtain good frequency selectivity. In the

oversampled case no similar restriction is imposed. The acquired design freedom is illustrated by the following example.

Example 4 *Consider the case $N = 2$ and $K = 4$. With the additional requirement that the prototype filter $V(z)$ is symmetric, the design consists of finding a single filter $F(z)$ which is orthogonal to its translates by multiples of 2. In terms of their polyphase components, $F(z)$ and $V(z)$ are given by*

$$F(z) = F_0(z^2) + z^{-1}F_1(z^2)$$
$$V(z) = V_0(z^4) + z^{-1}V_1(z^4) + z^{-2}V_2(z^4) + z^{-3}V_3(z^4)$$

The design constraints are thus satisfied by taking $V_0(z)$ and $V_2(z)$ to be equal to $F_0(z)$ and $F_1(z)$, respectively, and $V_1(z)$ and $V_3(z)$ to be their time reversed versions. One solution, obtained from a 4-tap filter $F(z)$, is shown in Figure 3.7. Of course,

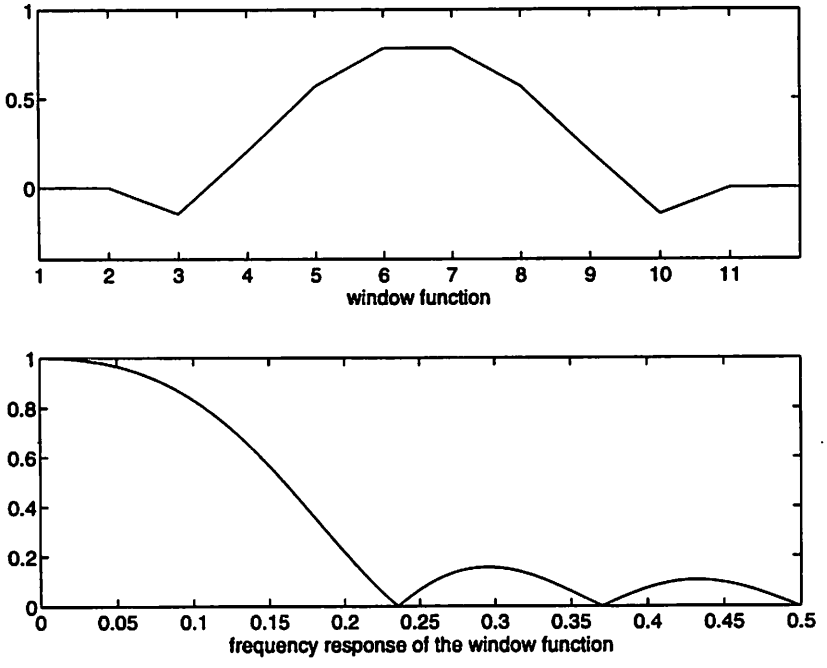

Figure 3.7: *An example of an 8-tap window function for tight Gabor frames with the oversampling factor $K/N = 2$.*

better frequency selectivity could be achieved with longer filters, $F(z)$, since in this case there is no restriction on the filter length. □

3.7 Appendix

3.7.1 Proof of Theorem 5

The norm of a discrete-time signal $x[n]$ in terms of its polyphase components is given by

$$\|x\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{i=0}^{N-1}|X_i(e^{j\omega})|^2 d\omega \tag{3.60}$$

Subband components of $x[n]$ produced at the output of the filter bank will be denoted by $y_i,\ i = 0, 1, \ldots, K-1$. Recall that they are given in the Z-transform domain by

$$[Y_0(z)\ Y_1(z)\ \ldots\ Y_{K-1}(z)]^T = \mathbf{H_p}(z)[X_0(z)\ X_1(z)\ \ldots\ X_{N-1}(z)]^T. \tag{3.61}$$

a) Necessary condition

Suppose $\mathbf{H_p}(e^{j\omega_0})\mathbf{a} = \mathbf{0}$ for some vector $\mathbf{a}$. Consider a sequence of signals $\{x_j\}_{j=1}^{\infty}$ such that the polyphase representation of x_j has the form

$$[X_0^{(j)}(e^{j\omega})\ X_1^{(j)}(e^{j\omega})\ \ldots\ X_{N-1}^{(j)}(e^{j\omega})]^T = p_j(\omega)\mathbf{a}, \quad \omega > 0$$

where

$$p_j(\omega) = \begin{cases} \frac{\sqrt{2\pi j}}{\|a\|} & \omega_0 - \frac{1}{4j} \le \omega \le \omega_0 + \frac{1}{4j} \\ 0 & \text{otherwise} \end{cases}.$$

Then $\|x_j\|^2 = 1$ but $\sum_{i=0}^{K-1}\|y_i^{(j)}\|^2 \to 0$ as $j \to \infty$. Therefore there is no $A > 0$ such that

$$A\|x\|^2 \le \sum_{i=0}^{K-1}\|y_i\|^2$$

for all $x \in \ell^2(\mathbf{Z})$. b)Sufficient condition. Consider

$$\sum_{i=0}^{K-1}\|y_i\|^2 = \sum_{i=0}^{K-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}|\sum_{j=0}^{N-1}H_{i,j}(e^{j\omega})X_j(e^{j\omega})|^2 d\omega$$

Since we consider FIR filter banks, the following holds:

$$\sum_{i=0}^{K-1}\|y_i\|^2 \ \le\ \sum_{i=0}^{K-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{j=0}^{N-1}|H_{i,j}(e^{j\omega})X_j(e^{j\omega})|^2 d\omega \tag{3.62}$$

$$= \sum_{j=0}^{N-1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=0}^{K-1} |H_{i,j}(e^{j\omega})|^2 |X_j(e^{j\omega})|^2 d\omega \tag{3.63}$$

$$\leq \sum_{j=0}^{N-1} \frac{1}{2\pi} \int_{-\pi}^{\pi} K\beta |X_j(e^{j\omega})|^2 d\omega \tag{3.64}$$

where $\beta = \sup_{i,j} \sup_{\omega} |H_{i,j}(e^{j\omega})|^2$. Therefore

$$\sum_{i=0}^{K-1} \|y_i\|^2 \leq B\|x\|^2, \quad 0 < B = K\beta < \infty, \quad \forall x \in \ell^2(\mathbf{Z}).$$

Suppose further that $\mathbf{H_p}(z)$ is of full rank everywhere on the unit circle. Then $\tilde{\mathbf{H}}_\mathbf{p}(z)\mathbf{H_p}(z)$ is also of full rank on the unit circle and hence

$$\mathbf{H_p}^+(z) = \left(\tilde{\mathbf{H}}_\mathbf{p}(z)\mathbf{H_p}(z)\right)^{-1} \tilde{\mathbf{H}}_\mathbf{p}(z)$$

is a matrix of rational functions of z which are bounded (no poles) on the unit circle. We have that

$$[X_0(z)\ X_1(z)\ \ldots\ X_{N-1}(z)]^T = \mathbf{H_p}^+(z)[Y_0(z)\ Y_1(z)\ \ldots\ Y_{K-1}(z)]^T.$$

Since

$$\|x\|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=0}^{N-1} |X_i(e^{j\omega})|^2 d\omega$$

using the same argument as in the above we can show that

$$\|x\|^2 \leq \alpha \sum_{j=0}^{K-1} \|y_j\|^2, \quad 0 < \alpha < \infty.$$

This proves that

$$A\|x\|^2 \leq \sum_{i=0}^{K-1} \|y_i\|^2 \leq B\|x\|^2,\ 0 < A \leq B < \infty, \quad \forall x \in \ell^2(\mathbf{Z}).$$

3.7.2 Proof of Theorem 6

a) Sufficiency of the condition is obvious.

b) Necessary condition.

Let $M_{i,j}(e^{j\omega}) = \left[\tilde{\mathbf{H}}_{\mathbf{p}}(e^{j\omega})\mathbf{H}_{\mathbf{p}}(e^{j\omega})\right]_{i,j}$. First, we want to prove all elements on the diagonal of $\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H}_{\mathbf{p}}(z)$ have to be equal to a constant. To check this consider signals x which have only one nonzero polyphase component. Then

$$\|x\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |X_i(e^{j\omega})|^2 d\omega$$

and

$$\sum_{i=0}^{K-1} \|y_i\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} M_{i,i}(e^{j\omega})|X_i(e^{j\omega})|^2 d\omega.$$

Then, for the tight frame condition to hold we must have $M_{i,i}(e^{j\omega}) = c$, for $i = 0, 1, \ldots, N-1$.

Now suppose that for some i, j, $M_{i,j}(e^{j\omega}) \equiv 0$ does not hold. Consider x such that only $X_i(z)$ and $X_j(z)$ are nonzero. Then

$$\sum_{i=0}^{K-1} \|y_i\|^2 = c\|x\|^2 + \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Re}\left(M_{i,j}(e^{j\omega})X_i^*(e^{j\omega})X_j(e^{j\omega})\right) d\omega.$$

From this it follows that

$$\int_{-\pi}^{\pi} \mathrm{Re}\left(M_{i,j}(e^{j\omega})X_i^*(e^{j\omega})X_j(e^{j\omega})\right) d\omega = 0, \quad \forall X_i^*(e^{j\omega}), X_j(e^{j\omega})$$

and therefore $M_{i,j}(e^{j\omega}) \equiv 0$.

3.7.3 Proof of Theorem 7

Given the filter bank polyphase analysis matrix $\mathbf{H}_{\mathbf{p}}(z)$, let

$$\mathbf{H}_{\mathbf{p}}(z) = \mathbf{R}(z)\mathbf{D}(z)\mathbf{C}(z)$$

be its Smith form decomposition. Here, $\mathbf{R}(z)$ and $\mathbf{C}(z)$ are products of corresponding elementary matrices in (3.14), while

$$\mathbf{D}(z) = \begin{bmatrix} \mathbf{D_1}(z) \\ \mathbf{0} \end{bmatrix}, \tag{3.65}$$

with $\mathbf{D_1}(z)$ being a diagonal matrix of monic polynomials. Suppose that the minimal dual of the frame associated to this filter bank consists of finite length vectors, i.e.

that $\mathbf{H}^{+}(z) = \left(\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H}_{\mathbf{p}}(z)\right)^{-1}\tilde{\mathbf{H}}_{\mathbf{p}}(z)$ is a polynomial matrix. Substituting the Smith decomposition in the equation for the pseudoinverse $\mathbf{H}^{+}(z)$ we obtain:

$$\mathbf{C}(z)^{-1}\left(\tilde{\mathbf{D}}(z)\tilde{\mathbf{R}}(z)\mathbf{R}(z)\mathbf{D}(z)\right)^{-1}\tilde{\mathbf{D}}(z)\tilde{\mathbf{R}}(z) = \mathbf{H}^{+}(z). \tag{3.66}$$

This expression further gives

$$\left(\tilde{\mathbf{D}}_{\mathbf{1}}(z)\mathbf{R}_{11}(z)\mathbf{D}_{\mathbf{1}}(z)\right)^{-1}\left[\begin{array}{cc}\tilde{\mathbf{D}}_{\mathbf{1}}(z) & \mathbf{0}\end{array}\right] = \mathbf{C}(z)\mathbf{H}^{+}(z)\tilde{\mathbf{R}}(z)^{-1} \tag{3.67}$$

where $\mathbf{R}_{11}(z)$ denotes the upper-left corner $N \times N$ submatrix of $\tilde{\mathbf{R}}(z)\mathbf{R}(z)$. It follows that

$$\left[\begin{array}{cc}\mathbf{D}_{\mathbf{1}}(z)^{-1}\mathbf{R}_{11}(z)^{-1} & \mathbf{0}\end{array}\right] = \mathbf{C}(z)\mathbf{H}^{+}(z)\tilde{\mathbf{R}}(z)^{-1}. \tag{3.68}$$

The matrix $\tilde{\mathbf{R}}(z)$ is unimodular and therefore its inverse is a polynomial matrix. Consequently, the right hand side of equation (3.68) is also a polynomial matrix. Therefore $\mathbf{D}_{\mathbf{1}}(z)^{-1}\mathbf{R}_{11}(z)^{-1}$ is a polynomial matrix. This implies that $\det\left(\mathbf{D}_{\mathbf{1}}(z)\mathbf{R}_{11}(z)\right) = c \cdot z^{k}$, for some constant c, and an integer k. However, this is possible only if polynomials on the diagonal of $\mathbf{D}_{\mathbf{1}}(z)$ are monomials.

Taking this fact into account, the expression $\left(\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H}_{\mathbf{p}}(z)\right)^{-1}\tilde{\mathbf{H}}_{\mathbf{p}}(z) = \mathbf{H}^{+}(z)$ can be transformed into

$$\left(\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H}_{\mathbf{p}}(z)\right)^{-1}\left[\begin{array}{cc}\tilde{\mathbf{C}}(z)\tilde{\mathbf{D}}_{\mathbf{1}}(z) & \mathbf{0}\end{array}\right] = \mathbf{H}^{+}(z)\tilde{\mathbf{R}}(z)^{-1}. \tag{3.69}$$

From this we obtain

$$\left(\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H}_{\mathbf{p}}(z)\right)^{-1} = \mathbf{H}^{+}(z)\tilde{\mathbf{R}}(z)^{-1}\left[\begin{array}{c}\tilde{\mathbf{D}}_{\mathbf{1}}(z)^{-1}\tilde{\mathbf{C}}(z)^{-1} \\ \mathbf{0}\end{array}\right]. \tag{3.70}$$

Since both $\tilde{\mathbf{R}}(z)$ and $\tilde{\mathbf{C}}(z)$ are unimodular, and $\tilde{\mathbf{D}}_{\mathbf{1}}(z)$ is the diagonal matrix of monomials, $\tilde{\mathbf{R}}(z)^{-1}$, $\tilde{\mathbf{C}}(z)^{-1}$ and $\tilde{\mathbf{D}}_{\mathbf{1}}(z)^{-1}$ are polynomial matrices. So, $\left(\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H}_{\mathbf{p}}(z)\right)^{-1}$ must be a polynomial matrix as well, which proves that $\det\left(\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H}_{\mathbf{p}}(z)\right) = p(z)$ has to be a monomial. However, this is only possible if $p(z) = constant$, since $p(z) = \tilde{p}(z)$. This proves that $\tilde{\mathbf{H}}_{\mathbf{p}}(z)\mathbf{H}_{\mathbf{p}}(z)$ has to be unimodular.

3.7.4 Proof of Proposition 8

It is obvious that for any $\mathbf{H_p}(z)$ of the form given in (3.16), $\tilde{\mathbf{H}}_\mathbf{p}(z)\mathbf{H_p}(z)$ is unimodular.

Suppose that $\mathbf{P}(z)$ is an $N \times N$ unimodular parahermitian polynomial matrix, positive definite on the unit circle. Any $N \times N$ parahermitian matrix of polynomials $\mathbf{P}(z)$ can be factored as

$$\mathbf{P}(z) = \tilde{\mathbf{H}}(z)\mathbf{H}(z). \tag{3.71}$$

where $\mathbf{H}(z)$ is an $N \times N$ matrix of polynomials [54]. Let $\mathbf{H}(z) = \mathbf{R}(z)\mathbf{D}(z)\mathbf{C}(z)$ be the Smith form decomposition of $\mathbf{H}(z)$. Since $\mathbf{P}(z)$ is unimodular and positive definite on the unit circle, polynomials on the diagonal of $\mathbf{D}(z)$ have to be nonzero monomials. This proves the proposition.

3.7.5 Proof of Proposition 9

Let the Smith decompositions of the polyphase analysis matrix $\mathbf{H_p}(z)$ and a corresponding FIR polyphase synthesis matrix $\mathbf{G_p}(z)$ be

$$\mathbf{H_p}(z) = \mathbf{R_h}(z) \begin{bmatrix} \mathbf{D_h}(z) \\ \mathbf{0} \end{bmatrix} \mathbf{C_h}(z) \quad \mathbf{G_p}(z) = \mathbf{R_g}(z) \begin{bmatrix} \mathbf{D_g}(z) & \mathbf{0} \end{bmatrix} \mathbf{C_g}(z) \tag{3.72}$$

where $\mathbf{D_h}(z)$ and $\mathbf{D_g}(z)$ are diagonal polynomial matrices. From the condition that $\mathbf{G_p}(z)\mathbf{H_p}(z) = \mathbf{I}$ and that $\det\mathbf{R_g}(z) = \det\mathbf{C_h}(z) = 1$, it follows that

$$\det\mathbf{D_g}(z) \cdot \det\mathbf{A}(z) \cdot \det\mathbf{D_h}(z) = 1, \tag{3.73}$$

where $\mathbf{A}(z)$ denotes the $N \times N$ submatrix in the upper-left corner of $\mathbf{C_g}(z)\mathbf{R_h}(z)$. Since $\mathbf{A}(z)$ is a polynomial matrix and both $\mathbf{D_h}(z)$ and $\mathbf{D_g}(z)$ are diagonal polynomial matrices, condition (3.73) can be satisfied only if polynomials on the diagonal of $\mathbf{D_h}(z)$ are monomials.

Conversely, perfect reconstruction can be achieved using the filter bank with the synthesis polyphase matrix $\mathbf{G_p}(z) = \mathbf{C_h}(z)^{-1} \begin{bmatrix} \mathbf{D_h}(z)^{-1} & \mathbf{0} \end{bmatrix} \mathbf{R_h}(z)^{-1}$. If in the Smith decomposition of the analysis polyphase matrix $\mathbf{D_h}(z)$ is a diagonal matrix of monomials, such a synthesis filter bank is obviously FIR.

Chapter 4

Quantization of Overcomplete Expansions

4.1 Noise Reduction in Frames

The fact that overcomplete expansions are less sensitive to degradations than nonredundant expansions comes as no surprise considering that redundancy in engineering systems usually provides robustness. This principle is exploited in modern techniques for A/D conversion, which use oversampling in order to compensate for coarse quantization, resulting in highly accurate conversion overall. Another example, from communications, is channel coding, where signal dependent patterns of control bits are added to data bits in order to combat channel noise. Error correction capability, i.e. the robustness of the code, is proportional to redundancy which is reflected in the number of control bits. Although the principle of redundancy-robustness trade-off seems close to ones intuition, it is not always as simple as it is in the case of the linear codes, to unravel underlying mechanisms and give a quantitative characterization. The first explanation of the noise reduction property of overcomplete expansions was given by Daubechies [10] and we review it here.

Consider a frame $\{\varphi_j\}_{j\in J}$ in a Hilbert space $\mathcal{H}$. Let $\{\tilde{\varphi}_j\}_{j\in J}$ be its minimal dual frame. Expansion coefficients of a signal f in $\mathcal{H}$ with respect to $\{\varphi_j\}_{j\in J}$ are given as

inner products with corresponding vectors of the dual frame,

$$f = \sum_{j \in J} \langle f, \tilde{\varphi}_j \rangle \varphi_j. \tag{4.1}$$

In other words, the expansion coefficients are the image of f under the frame operator $\tilde{F}$ which is associated with the minimal dual frame. Recall that $\tilde{F}$ is an operator from $\mathcal{H}$ to $\ell^2(J)$. If $\{\varphi_j\}$ is a linearly independent set, then the range of $\tilde{F}$ is all of $\ell^2(J)$, that is any point in $\ell^2(J)$ represents a set of the expansion coefficients of some g in $\mathcal{H}$. However, if $\{\varphi_j\}$ is overcomplete, then $\tilde{\varphi}_j$ are linearly dependent. This linear dependency translates to expansion coefficients $\langle f, \tilde{\varphi}_j \rangle$. Consequently, the range of $\tilde{F}$ is no longer all of $\ell^2(J)$, but a proper subspace. A family of expansion coefficients $\langle f, \tilde{\varphi}_j \rangle$ is usually not in this subspace after degradation with some additive noise. Starting with the adulterated coefficients, $\{\langle f, \tilde{\varphi}_j \rangle + n_j\}$, another set closer to the originals can be obtained as the orthogonal projection of $\{\langle f, \tilde{\varphi}_j \rangle + n_j\}$ onto the range of $\tilde{F}$ (see Figure 4.1). Recall that the reconstruction formula

$$f_{rec} = \sum_{j \in J} (\langle f, \tilde{\varphi}_j \rangle + n_j) \varphi_j \tag{4.2}$$

implicitly involves this projection, thus reducing to zero the component of $\{n_j\}$ which is orthogonal to the range of $\tilde{F}$. As the frame redundancy increases, the range of $\tilde{F}$ becomes more and more constrained, "smaller" in some sense, so the noise reduction becomes more effective. These heuristics can be put in a quantitative framework in some particular but nevertheless significant cases. If we consider tight frames in a finite dimensional space, $\mathbf{R}^n$, and a white, zero-mean, additive noise, it can be shown that in this way the expected squared error norm is reduced proportionally to the frame redundancy factor, r [10, 15]. Precisely

$$E(\|f - f_{rec}\|^2) = n\sigma^2/r, \tag{4.3}$$

where σ^2 is the noise variance. Noise reduction in Weyl-Heisenberg and wavelet frames in $L^2(\mathbf{R})$ was also studied by Daubechies [10], under the same assumption of a white, additive, zero-mean noise. Suppose that a signal $f \in L^2(\mathbf{R})$ is well localized in a bounded region of the time frequency plane, so that it can be well approximated

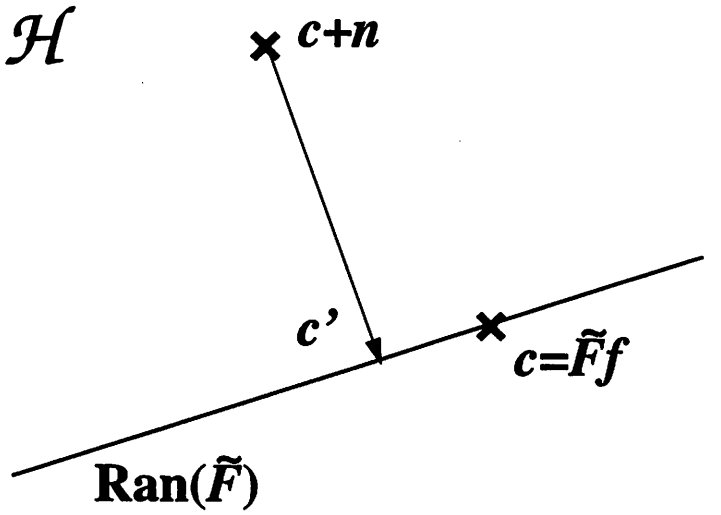

Figure 4.1: *Noise reduction in frames using linear reconstruction. Linear reconstruction of f from noisy coefficients, $\tilde{F}f+\mathbf{n}$ gives a signal f_{rec}. The sequence of expansion coefficients of f_{rec}, $\mathbf{c}' = \tilde{F}f_{rec}$ is the orthogonal projection of degraded coefficients onto the range of $\tilde{F}$, and is therefore closer to the original.*

using a finite number of elements of a wavelet or Weyl-Heisenberg expansion (those localized in the same region). If the expansion coefficients $\{\langle f, \tilde{\varphi}_j\rangle\}_{j\in J}$ are subject to the noise $\{n_j\}_{j\in J}$ then f can be reconstructed as

$$f_{rec} = \sum_{j\in BCJ} (\langle f, \tilde{\varphi}_j\rangle + n_j)\varphi_j. \tag{4.4}$$

The reconstruction error produced in this way can be bounded as

$$E(\|f - f_{rec}\|^2) = \epsilon\|f\|^2 + O(\sigma^2/r), \tag{4.5}$$

where the $\epsilon\|f\|^2$ component is the result of the approximation of f using the finite set of the expansion terms, $\{\varphi_j\}_{j\in BCJ}$. A rigorous proof of this result was Given by Munch [31] for the case of tight Weyl-Heisenberg frames and integer frame redundancy factors. Note that this analysis, although pertaining to $L^2(\mathbf{R})$, has a finite dimensional space flavor, since contribution of the noise on finitely many coefficients is considered.

Based on these results it may be conjectured that the $O(1/r)$ noise reduction property has a more general scope than discussed here. For instance, in $\mathbf{R}^n$ it might hold for frames in general, and that is indicated by results reported in [15]. Or in $L^2(\mathbf{R})$ (4.5) may be valid for frames other than Weyl-Heisenberg frames, and

redundancy factors other than the integer ones. However, no general results have been proven yet.

4.2 Reduction of Quantization Error and Suboptimality of the Linear Reconstruction

The work on the noise reduction effect, reviewed in the previous section, was actually aimed at estimating quantization error which is modeled as an additive white noise. However, results on reconstruction from quantized frame expansions, reported by Morlet (see [10]), indicate that the error decays faster than could be expected from the $O(1/r)$ results. The purpose of this section is to elucidate why a higher reconstruction precision can be expected and how it can be attained.

Quantization of expansions in a Hilbert space $\mathcal{H}$, with respect to a given frame $\{\varphi_j\}_{j\in J}$, is a many to one mapping from $\mathcal{H}$ to $\mathcal{H}$,

$$Q : \mathcal{H} \to \mathcal{H}.$$

It defines a partition of $\mathcal{H}$ into disjoint cells $\{C\}_{i\in I}$. In the case of uniform scalar quantization, each of the cells is defined by a set of convex constraints of the type

$$C_i = \{f : \ (n_{ij} - 1/2)q \le \langle f, \tilde{\varphi}_j \rangle < (n_{ij} + 1/2)q, \ \ j \in J\}, \tag{4.6}$$

where q is the quantization step. For each of the cells, the quantization maps all the signals in the cell to a single signal in its interior, usually its centroid. Roughly speaking, the expected value of quantization error reflects fineness of the partition. One way to refine the partitioning is to tighten the constraints which define cells (4.6) by decreasing the quantization step. Alternatively, for a given quantization step, more constraints can be added, which corresponds to an increase in redundancy of the frame. This gives another explanation of the error reduction property in frames, this time for the quantization error. Effectiveness of the two approaches to partition refinement, that is error reduction, can be assessed based on the results reviewed in the previous section. If the white noise model for the quantization error is accepted, than

the error variance σ^2 is proportional to the square of the maximum error value, $q/2$. According to formulae (4.3) and (4.5), it seems that the quantization step refinement is more effective than the increase in frame redundancy for the error reduction, since the error decreases proportionally to q^2 in the former case, and proportionally to $1/r$ in the latter case. Is this always true, or we are underestimating the potential of frame redundancy for error reduction? Let us look at a simple example for the sake of providing some intuition about this effect.

Quantization of expansions in $\mathbf{R}^2$, with respect to the orthonormal basis

$$\{(0,1),(1,0)\}$$

induces the partition of the plane into square cells

$$\mathcal{C}_i = \{(x,y) : (n_{ix}-1/2)q \leq x < (n_{ix}+1/2)q, (n_{iy}-1/2)q \leq y < (n_{iy}+1/2)q\}. \quad (4.7)$$

Figure 4.2 illustrates partitions obtained after refining the initial partition, with $q = 1$, by reducing q by a factor k or introducing redundant expansions with redundancy factor k, for $k = 2,3,4$. In the first case, each cell of the initial partition is uniformly divided into k^2 square subcells. On the other hand, in the case of increased redundancy the number of subcells varies around k^2. Some of the cells induced by overcomplete expansions are larger while some of them are smaller than the cells of the orthonormal partition for the same k. Based on a simple inspection of the partitions in Figure 4.2 it is hard to tell which of the two approaches gives a finer partition on average. However, if we take a look at the partitions in Figure 4.3, generated by reducing q by the factor k or increasing the redundancy by the factor k^2, we can observe that in the latter case partition refinement is faster, contrary to what would be expected from relations (4.3) and (4.5).

The reason for this discrepancy is that the error estimates in the previous subsection were derived under the assumption of linear reconstruction given by (4.2). However, when dealing with quantization of overcomplete expansions, linear reconstruction is not optimal in the sense that it does not necessarily yield a signal which lies in the same quantization cell with the original. Hence, linear reconstruction does

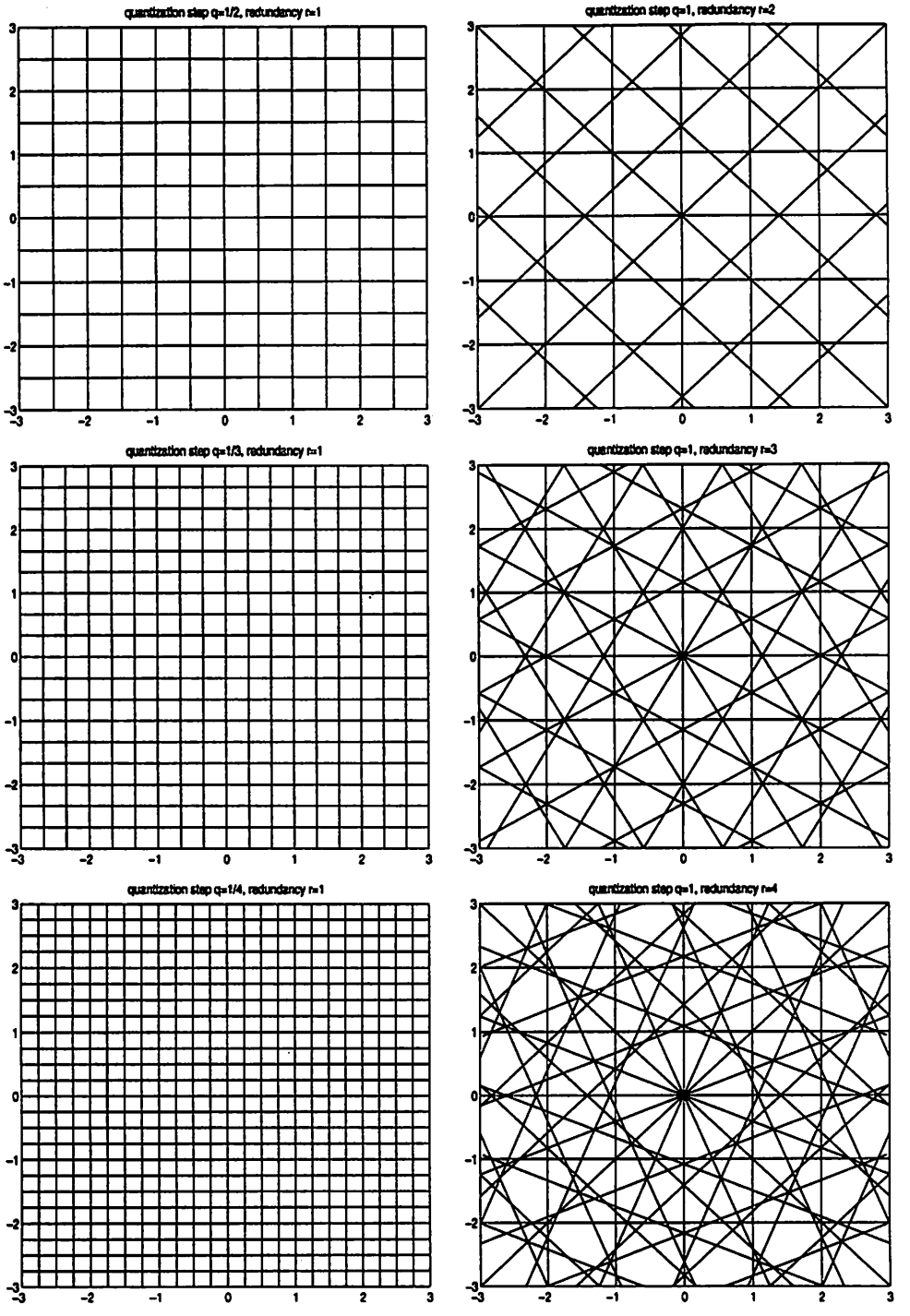

Figure 4.2: *Partitions of* $\mathbf{R}^2$ *induced by quantization of frame expansions. Partitions in the left column are obtained for the orthonormal basis* $\{(1,0),(0,1)\}$ *and quantization steps* $q = 1/2,\ 1/3$ *and* $1/4$. *Partitions in the right column correspond to the quantization step* $q = 1$ *and tight frames* $\{(cos(i\pi/2r), sin(i\pi/2r))\}_{i=0\ldots 2r-1}$, *for* $r = 2,\ 3$ *and* 4.

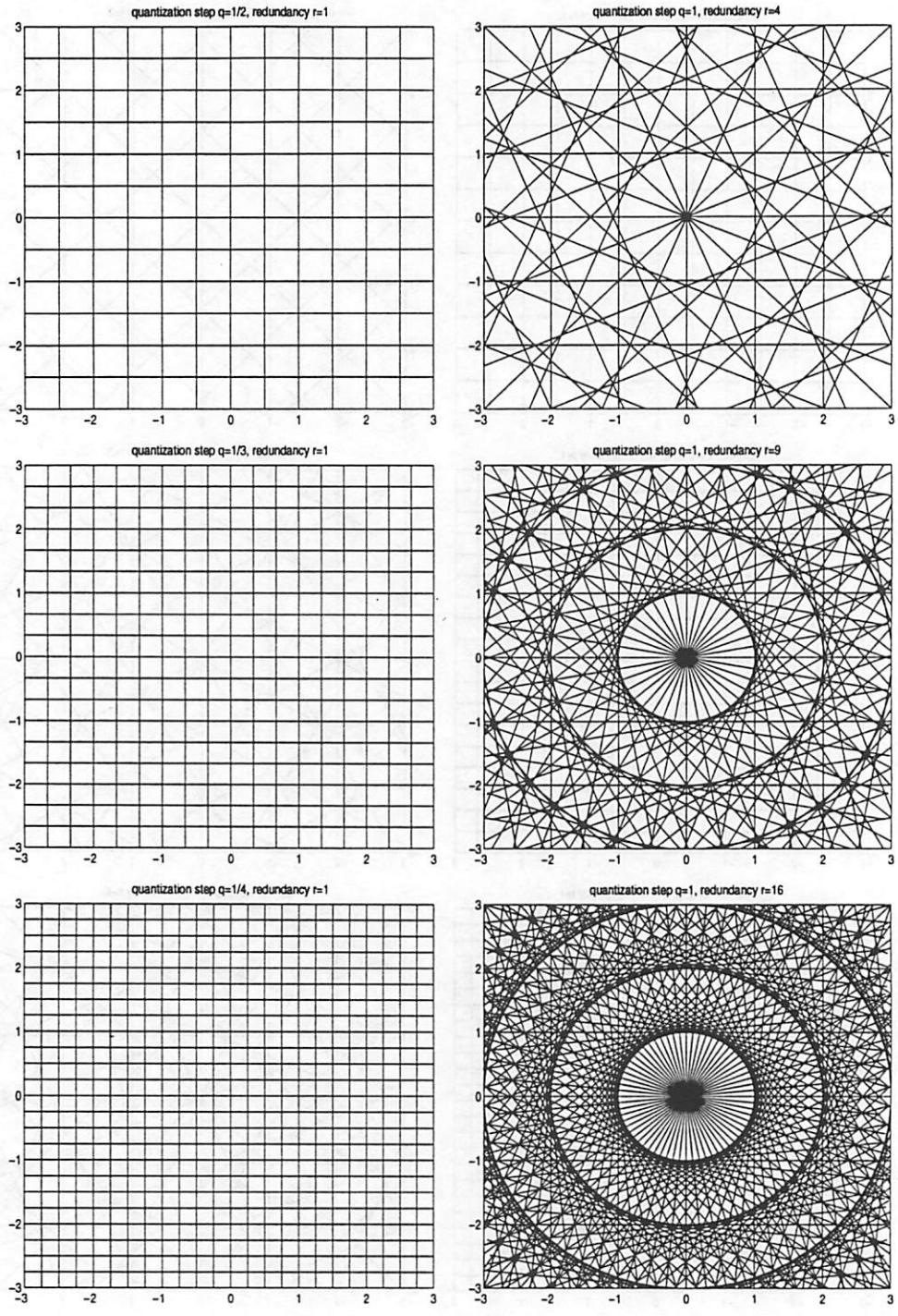

Figure 4.3: *Partitions of* $\mathbf{R}^2$ *induced by quantization of frame expansions. Partitions in the left column are obtained for the orthonormal basis* $\{(1,0),(0,1)\}$ *and quantization steps* $q = 1/2,\ 1/3$ *and* $1/4$. *Partitions in the right column correspond to the quantization step* $q = 1$ *and tight frames* $\{(cos(i\pi/2r), sin(i\pi/2r))\}_{i=0\ldots 2r-1}$, *for* $r = 4,\ 9$ *and* 16.

not fully utilize information which is contained in the quantized set of coefficients, so there is some room for improvement. Figure 4.4 depicts such a scenario. Let $\mathbf{c}$ denote the set of quantized coefficients of a signal f, $\mathbf{c} = Q(\tilde{F}f)$, and let $\mathbf{c}$ be in the partition cell C_i. As the result of linear reconstruction we obtain a signal f_{rec}. Expansion coefficients of f_{rec}, $\tilde{F}f_{rec}$, are obviously in the range of the frame expansions, and are closer to the original ones than $\mathbf{c}$ is, $\|\tilde{F}f_{rec} - \tilde{F}f\| < \|\mathbf{c} - \tilde{F}f\|$. However, $\tilde{F}f$ and $\tilde{F}f_{rec}$ do not lie in the same quantization cell.

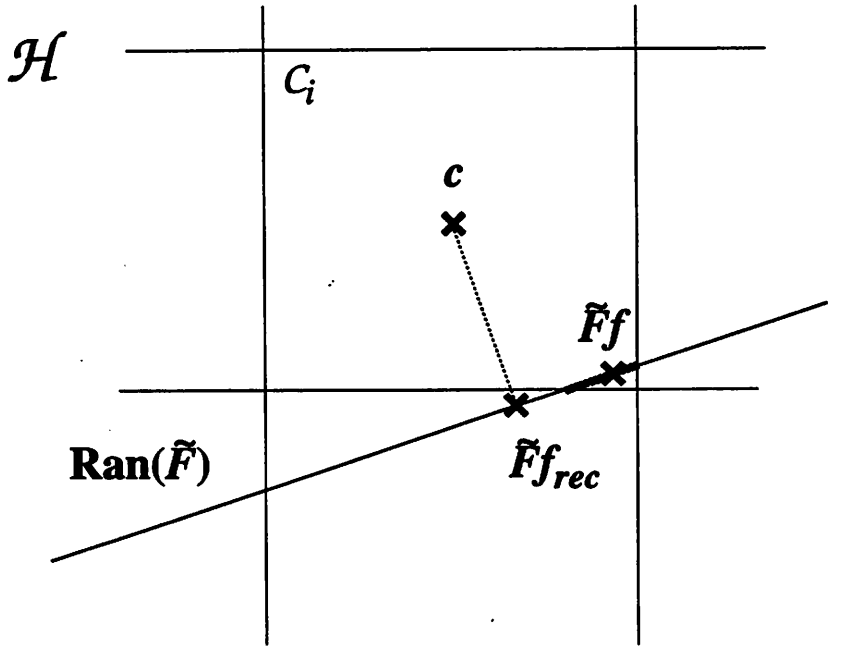

Figure 4.4: *Inconsistency of linear reconstruction from quantized coefficients of an overcomplete expansion. Expansion coefficients of a signal f, $\tilde{F}f$, are located in the quantization cell C_i. Expansion coefficients $\tilde{F}f_{rec}$, which are obtained from the quantized coefficients $\mathbf{c}$ using the linear reconstruction, need not lie in the same quantization cell.*

Reconstruction can be further improved by alternating projections of $\tilde{F}f_{rec}$ onto C_i and $\mathrm{Ran}(\tilde{F})$ until a signal in their intersection is reached. Note that all signals which have expansion coefficients in the intersection $\mathrm{Ran}(\tilde{F}) \bigcap C_i$ produce the same result after the quantization, and therefore cannot be distinguished from f based on the quantized expansion coefficients. These signals constitute the so called *reconstruction*

set of f,

$$\Omega_r(f) = \{g : Q(\tilde{F}g) = Q(\tilde{F}f)\}, \tag{4.8}$$

and are denoted as its *consistent estimates.*

This has been pointed out first by Thao and Vetterli in the context of oversampled A/D conversion [44, 43, 45]. Thao and Vetterli further showed that in the case of the conversion of periodic bandlimited signals (trigonometric polynomials), for any consistent estimate f_c of f $\|f - f_c\|^2 = O(1/r^2)$, where r is the oversampling ratio, provided that f has a sufficient number of quantization threshold crossings. Experimental evidence of the $O(1/r^2)$ behavior was also reported in [18, 19]. The question which naturally arises is whether this result has a wider scope. So far it has been numerically verified for frames in $\mathbf{R}^n$ [15, 16]. In the remainder of this chapter we consider two particular cases in infinite dimensional spaces. These are oversampled A/D conversion of bandlimited signals in $L^2(\mathbf{R})$ and quantization of Weyl-Heisenberg expansions in $L^2(\mathbf{R})$. We prove that under certain reasonable assumptions, consistent reconstruction gives an error e which behaves as $\|e\|^2 = O(1/r^2)$. Considerations in this subsection already have the flavor of deterministic analysis and this is the approach which is taken here in the treatment of the error reduction effect. Statistical analysis and the white noise model, commonly used for the study of quantization error, are not appropriate in the case of overcomplete expansions, since they ignore the deterministic nature of the error and consequently miss some details on its structure which are essential for establishing more accurate bounds.

Before the details of this deterministic analysis are presented a comment on rate-distortion tradeoffs in frame expansions is in order. Besides the suboptimality of linear reconstruction, traditional encoding of quantized coefficients using pulse code modulation (PCM) is also not suitable in conjunction with overcomplete expansions if efficient signal representations are an issue. The reason lies in the fact that in the case of quantized orthonormal expansions reduction of the quantization step k times increases the bit rate by the factor $\log_2 k$, whereas the effect of a k times higher redundancy of a frame expansion is a k times higher bit rate. A sophisticated encoding algorithm is needed to overcome this disadvantage. However, as we pointed

out earlier, frames which appear in engineering practice are not arbitrary frames but rather well structured ones which may facilitate the encoding task. In the following section we discuss this problem in the case of oversampled A/D conversion, and propose an efficient encoding scheme.

4.3 Deterministic Analysis of Oversampled A/D Conversion

4.3.1 Traditional View and Its Limitations

Oversampled analog to digital conversion is the fundamental instance of signal representation based on quantization of overcomplete expansions. The aspects discussed in a general framework in the previous section are considered in this particular case in more detail here.

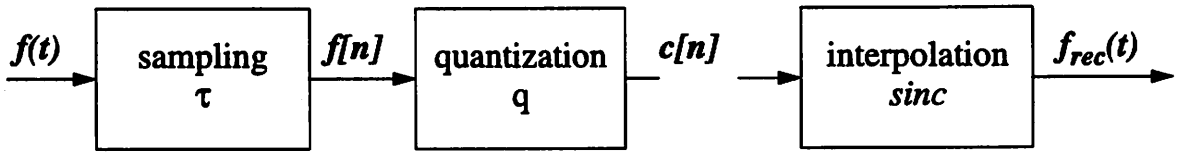

Figure 4.5: *Block diagram of simple A/D conversion followed by classical reconstruction. Input σ-bandlimited signal $f(t)$ is first sampled at a frequency $f_s = 1/\tau$, which is above the Nyquist frequency $f_N = \sigma/\pi$. The sequence of samples $f[n]$ is then discretized in amplitude with a quantization step q. Classical reconstruction gives a signal $f_{rec}(t)$, which is obtained as a low-pass filtered version of some signal having the same digital version as the original $f(t)$, which amounts to a sinc interpolation between quantized samples.*

Oversampled A/D conversion as a process of digital encoding of analog signals involves discretization of a σ-bandlimited analog signal in time, implemented as sampling with a time interval τ, followed by discretization in amplitude, that is quantization of the sequence of samples with a quantization step q. The classical way to reconstruct the discretized analog signal is by low-pass filtering, with $\sigma/2\pi$ cutoff frequency, a pulse train which is modulated by the quantized sequence of samples. A

block diagram of the converter together with a classical reconstruction scheme is illustrated in Figure 4.5. The discretization in time allows for perfect reconstruction of the analog signal as long as it is sampled at or above the Nyquist rate, $\tau < \tau_N$, where $\tau_N = \pi/\sigma$. However the discretization in amplitude introduces an irreversible loss of information. The difference between the original analog signal, f, and its version, f_{rec}, recovered from the digital representation is therefore denoted as the quantization error, $e = f - f_{rec}$. Quantization error is commonly modeled as a uniformly distributed white noise independent of the input. Under certain conditions, such as a large number of quantization levels and a small quantization step q compared to the input amplitude range, this model provides a satisfactory error approximation giving the error variance $\sigma_e^2 = E(e(t)^2) = q^2/12$. Therefore, a natural way to improve accuracy of the conversion would be to reduce the quantization step q. However, the complexity of quantizer circuitry and the precision of analog components impose limits on the quantization step refinement. Modern techniques for high resolution A/D conversion are based on oversampling. Sampling a signal at a rate higher than the Nyquist rate introduces redundancy which can be exploited for error reduction. The white noise model explains this effect in the following way. As a result of oversampling, the quantization error spectrum remains uniform and spreads over a wider frequency range $[-1/\tau, 1/\tau]$, while the error variance remains $q^2/12$. Only the τ/τ_N portion of the error spectrum is in the frequency range of the input signal. Therefore, after the low-pass filtering with σ/π cutoff frequency, the error variance is reduced to $\sigma_e^2 = q^2\tau/12\tau_N$, or

$$E(e(t)^2) = \frac{q^2}{12r}, \tag{4.9}$$

where $r = \tau/\tau_N$ is the oversampling ratio (see Figure 4.6).

The error estimate in (4.9) indicates that although the oversampling can be exploited for improvement of conversion accuracy, it is inferior to quantization refinement. Furthermore, if the error is considered as a function of the bit rate, quantization refinement drastically outperforms oversampling. As an illustration consider the case when the quantization step is halved. This reduces the error by the factor of four, at the expense of an additional bit per sample. In order to achieve the same error

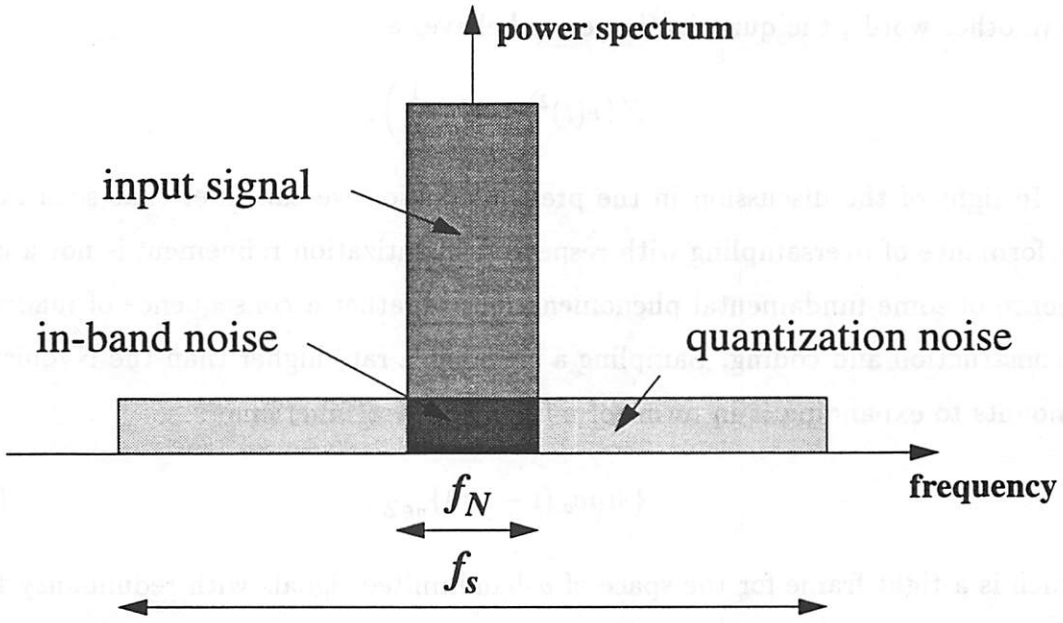

Figure 4.6: *Statistical analysis of the MSE of oversampled A/D conversion with classical reconstruction. Under certain assumptions, the quantization error can be modeled as a white noise. As the sampling interval decreases, the noise spectrum remains flat and spreads out over the whole sampling frequency range, while keeping a constant total power equal to* $\frac{\Delta q^2}{12}$. *For a sampling frequency* f_s, *which is above the Nyquist rate, only the portion* $\frac{\Delta q^2}{12}\frac{f_N}{f_s}$ *is still present in the signal spectrum after the low-pass filtering with* $f_N/2$ *cut-off frequency.*

reduction, the oversampling ratio should be increased four times, and consequently the total number of bits increases by the same factor. This reasoning gives the following results. If the quantization step varies, for a fixed sampling interval τ, the quantization error as a function of the bit rate, B, is given by

$$E\left(e(t)^2\right) = \frac{1}{12}A^2\frac{\tau}{\tau_N}2^{-2\tau B}, \tag{4.10}$$

where A denotes signal amplitude range. Hence, as the quantization step decreases, the error decays as

$$E\left(e(t)^2\right) = O\left(2^{-2\tau B}\right). \tag{4.11}$$

On the other hand if the sampling is successively refined, $\tau \to 0$, for a fixed quantization step,

$$E\left(e(t)^2\right) = \frac{1}{12}\frac{q^2}{\tau_N}\left(\log_2\frac{A}{q}\right)\frac{1}{B}, \tag{4.12}$$

or in other words, the quantization error behaves as

$$E\left(e(t)^2\right) = O\left(\frac{1}{B}\right). \tag{4.13}$$

In light of the discussion in the previous section we can infer that such inferior performance of oversampling with respect to quantization refinement is not a consequence of some fundamental phenomena, but is rather a consequence of inadequate reconstruction and coding. Sampling a signal at a rate higher than the Nyquist rate amounts to expanding it in terms of a family of *sinc* functions

$$\{sinc_\sigma\,(t - n\tau)\}_{n\in\mathbf{Z}}\,, \tag{4.14}$$

which is a tight frame for the space of σ-bandlimited signals with redundancy factor equal to the oversampling ratio. The classical reconstruction based on low-pass filtering is essentially the linear reconstruction (see (4.2)) and is therefore inconsistent and suboptimal. The main point of this chapter is a deterministic analysis of the quantization error which demonstrates that the error of consistent reconstruction can be under reasonable assumptions bounded as $\|e(t)\|^2 = O(1/r^2)$. We also propose an efficient scheme for lossless encoding of quantized samples, which when used together with consistent reconstruction gives the error rate characteristics of the oversampled A/D conversion of the form

$$\|e(t)\|^2 = O\left(2^{-2\beta B}\right). \tag{4.15}$$

The constant β in (4.15) is of the same order of magnitude as the Nyquist sampling interval, so that we have performance comparable to that where the quantization step tends to zero (4.11), and significantly improved compared to that with linear reconstruction and traditional PCM encoding (4.13).

4.3.2 Reconstruction Sets and Asymptotic Behavior

The deterministic analysis of oversampled A/D conversion is based on the study of the structure of the partition which it induces in $L^2[-\sigma,\sigma]$, when considered

as a quantization of an overcomplete expansion. The knowledge on a signal, contained in its digitally encoded representation, defines a set of signals which cannot be distinguished from it after the A/D conversion. Recall from the discussion in the previous section that this set is referred to as the reconstruction set of the considered signal. In this context, the reconstruction set of a signal f can be represented as the intersection of two sets

$$\Omega_r(f) = \mathcal{V}_\sigma \cap \mathcal{C}_q(f), \tag{4.16}$$

where $\mathcal{V}_\sigma$ is the space of σ-bandlimited signals, and $\mathcal{C}_q(f)$ consists of signals in $L^2(\mathbf{R})$ which have amplitudes within the same quantization intervals as f at all sampling instants.

The size of the reconstruction set can be considered as a measure of the uncertainty about the original analog signal. Let us assume that the quantization interval is decreased by an integer factor m. As a result, the reconstruction set becomes smaller, since $\mathcal{C}_{q/m}(f)$ is a proper subset of $\mathcal{C}_q(f)$ (see Figure 4.7). In the limit, when $q \to 0$, there is no uncertainty about the signal amplitude at the sampling points $..., -\tau, 0, \tau, 2\tau, ...$, and the signal can be perfectly reconstructed if sampled above the Nyquist rate. This means that the reconstruction set asymptotically reduces to a single point which is the considered analog signal itself.

This view of the reconstruction set is based on the traditional interpretation of the digital version of an analog signal as a representation which bears information on its amplitude, with the uncertainty determined by the quantization step q, at the sequence of time instants $t = ..., -\tau, 0, \tau, 2\tau, ...$ An alternative interpretation comes into play if the quantization threshold crossings are separated and sampling is sufficiently fine so that at most one quantization threshold crossing can occur in each of the sampling intervals. If this is satisfied, quantized samples of the signal f are uniquely determined by the sampling intervals in which its quantization threshold crossings occur, and vice a versa. So, we can think of the digital representation as carrying information on the instants, with uncertainty τ in time, when the signal assumes values $..., -q, 0, q, 2q, ...$. Hence, the reconstruction set of f can also be

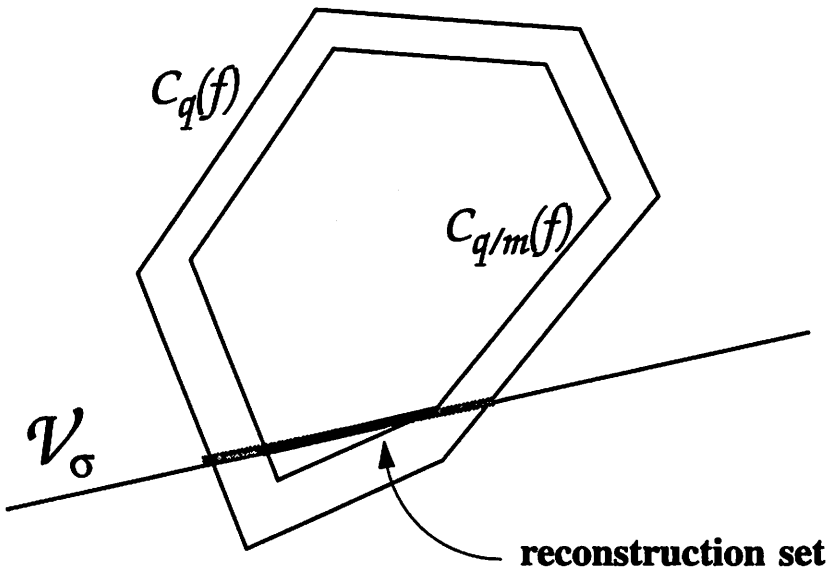

Figure 4.7: *The effect of the quantization refinement on the reconstruction set. The set of signals which share the same digital version with the original, $C_q(f)$, becomes smaller as the result of quantization refinement, $C_{q/m}(f) \subset C_q(f)$. Consequently, the reconstruction set is also reduced giving a higher accuracy of the representation*

viewed as lying in the intersection

$$\Omega_\tau(f) \subseteq \mathcal{V}_\sigma \cap \mathcal{D}_\tau(f), \tag{4.17}$$

where $\mathcal{D}_\tau(f)$ is the set of all signals in $L^2(\mathbf{R})$ which have the same quantization threshold crossing as the original, in all of the sampling intervals where the original goes through a quantization threshold. If in addition we require that signals in $\mathcal{D}_\tau(f)$ cannot have more than one quantization threshold crossing per sampling interval, then equality holds in (4.17).

Consider now the effect of reduction of the sampling interval by an integer factor m. The set of signals sharing the sampling intervals of quantization threshold crossings with the original, for this new sampling interval, is a proper subset of $\mathcal{D}_\tau(f)$, $\mathcal{D}_{\tau/m}(f) \subset \mathcal{D}_tau(f)$. As a result, the size of the reconstruction set is also reduced, implying a higher accuracy of the conversion. This reasoning gives a deterministic explanation for the error reduction as a consequence of oversampling in the A/D conversion. Quantization step refinement, which also improves the accuracy of the conversion, asymptotically gives the reconstruction set reduced to the original signal

only. Does this also happen in this case, when the sampling interval goes to zero? In other words, is it possible to reconstruct perfectly an analog σ-bandlimited signal after simple A/D conversion in the case of infinitely high oversampling?

In the limit, when the sampling interval assumes an infinitely small value, the time instants in which the signal goes through the quantization thresholds are known with infinite precision. The information on the input analog signal in this limiting case is its values at a sequence of irregularly spaced points $\{\lambda_n\}$, which are its quantization threshold crossings. If $\{e^{j\lambda_n\omega}\}$ is complete in $L^2[-\sigma,\sigma]$, then the reconstruction set asymptotically does reduce to a single point, giving the perfectly restored input analog signal. Completeness of $\{e^{j\lambda_n\omega}\}$ in $L^2[-\sigma,\sigma]$ means that any σ-bandlimited signal $f(t)$ is determined by the sequence of samples $\{f(\lambda_n)\}$. However, this does not necessarily mean that $f(t)$ can be reconstructed in a numerically stable way from the samples $\{f(\lambda_n)\}$, unless another constraint on $\{\lambda_n\}$ is introduced which would ensure that any two σ-bandlimited signals which are close at points $\{\lambda_n\}$ are also close in L^2 norms. In other words we want to be sure that the reconstruction error does converge to zero as the oversampling tends to infinity, that is

$$\lim_{\tau\to 0}\|f - f_{rec}\| = 0, \tag{4.18}$$

rather than only asymptotically becoming zero in some odd fashion. A precise formulation of the stability requirement is that $\{e^{j\lambda_n\omega}\}$ constitute a frame in $L^2[-\sigma,\sigma]$. The error bound derived in the following subsection is based on this assumption. In order to illustrate how restrictive this assumption is, at this point we briefly review some of the frame conditions discussed in Section 2.3.1.

Recall that exact frames in a separable Hilbert space, those which cease to be frames when any of the elements is removed, are Riesz bases. Are there any bases of complex exponentials other than Riesz bases? This is still an open problem. Every example of a basis of complex exponentials for $L^2[-\sigma,\sigma]$ so far has been proven to be a Riesz basis ([56], pp. 190-197). So, if we want a reasonable, complete characterization of σ-bandlimited signals from samples at points $\{\lambda_n\}$, then $\{e^{j\lambda_n\omega}\}$ should be a frame in $L^2[-\sigma,\sigma]$.

According to Kadec's 1/4 Theorem, if a sequence of real numbers $\{\lambda_n\}$ satisfies

$$|\lambda_n - n\frac{\pi}{\sigma}| \leq L < \frac{1}{4}\frac{\pi}{\sigma}, \quad n = 0, \pm 1, \pm 2, \ldots, \tag{4.19}$$

then $\{e^{j\lambda_n\omega}\}$ is a Riesz basis for $L^2[-\sigma, \sigma]$. How realistic is this condition in the quantization context? For a quantization step small compared to a signal amplitude, a sufficiently dense sequence of quantization threshold crossings could be expected, thus satisfying (4.19), at least on a time interval before the signal magnitude eventually falls below the lowest nonzero quantization threshold. Suppose that the analog signal $f(t)$ at the input of an A/D converter satisfies: $\hat{f}(\omega)$ is continuous on $[-\sigma, \sigma]$ and $\hat{f}(\sigma) \neq 0$. Such a signal $f(t)$, in terms of its zero-crossings, asymptotically behaves as $sinc_\sigma(t)$. So, if one of the quantization thresholds is set at zero, then for large n, the quantization threshold crossings should be close to points $n\pi/\sigma$.

Another criterion, given by Theorem 3, states that if $\{\lambda_n\}$ has a uniform density $d > \sigma/\pi$, then $\{e^{j\lambda_n\omega}\}$ is a frame in $L^2[-\sigma, \sigma]$. The feasibility of a σ-bandlimited signal with a sequence of quantization threshold crossings having uniform density greater than σ/π is unlikely for a quantization scheme with a fixed quantization step. The amplitude of a bandlimited signal $f(t)$ decays at least as fast as $1/t$, and outside of some finite time interval, $[-T, T]$, it is confined in the range between the two lowest nonzero quantization thresholds. Zero-crossings are the only possible quantization threshold crossings of $f(t)$ on $(-\infty, -T] \cup [T, \infty)$. However, it doesn't seem to be plausible that zero-crossings of a σ-bandlimited signal can have a uniform density greater than the density of zero-crossings of $sinc_\sigma(t)$, which is σ/π. Therefore, in order to meet this criterion the quantization step has to change in time following the decay of the signal. One way to achieve this is a scheme with quantization steps which are fixed on given time segments but eventually decrease with segment order.

4.3.3 $O(1/r^2)$ Error Bound

Consider a σ-bandlimited signal, $f(t)$, at the input of an oversampled A/D converter with sampling interval τ. Suppose that its sequence of quantization threshold crossings, $\{x_n\}$, is a frame sequence for the space of σ-bandlimited signals. If $g(t)$

is a consistent estimate of $f(t)$, there exists a sequence $\{y_n\}$ of quantization threshold crossings of $g(t)$, such that every x_n has a corresponding y_n in the same sampling interval. Hence, for each pair (x_n, y_n), $f(x_n) = g(y_n)$ and $|x_n - y_n| \leq \tau$.

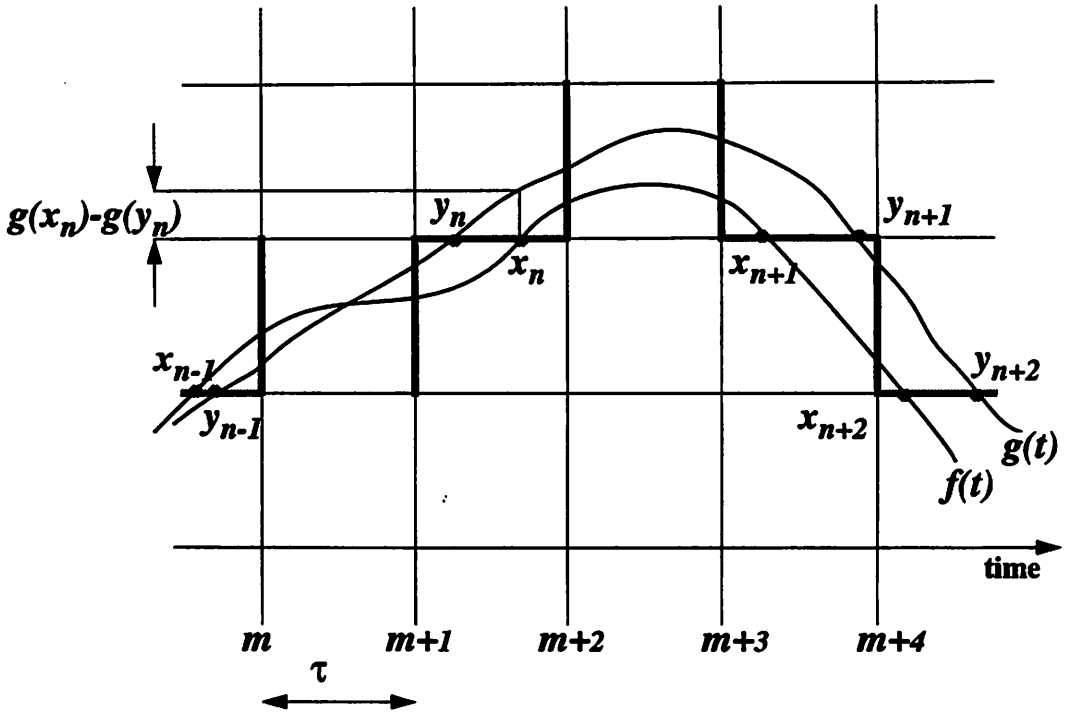

Figure 4.8: *Quantization threshold crossings of an analog signal $f(t)$ and its consistent estimate $g(t)$. The sequence of quantization threshold crossings, $\{x_n\}$, of $f(t)$, uniquely determines its digital version and vice versa, provided that all the crossings occur in distinct sampling intervals. If $f(t)$ goes through a certain quantization threshold at the point x_n, then $g(t)$ has to cross the same threshold at a point y_n which is in the same sampling interval with x_n. At a point x_n, the error amplitude is equal to $|g(x_n) - f(x_n)| = |g(y_n) - g(x_n)|$.*

At the points $\{x_n\}$, the error amplitude is bounded by the variation of $g(t)$ on the interval $[y_n, x_n]$ (without loss of generality we assume that $y_n < x_n$), as shown in Figure 4.8. Since $g(t)$ is bandlimited, which also means that it has finite energy, it can have only a limited variation on $[y_n, x_n]$. This variation is bounded by some value which is proportional to the sampling interval, so

$$|g(x_n) - g(y_n)| \leq c_n \cdot \tau. \tag{4.20}$$

The constant c_n in this relation can be the maximum slope of $g(t)$ on the interval $[y_n, x_n]$, $c_n = \sup_{t\in(y_n,x_n)} g'(t)$. The error signal, $e(t) = g(t) - f(t)$, is itself a σ-bandlimited signal. At the points $\{x_n\}$ its amplitude is bounded by values which are proportional to τ (4.20). If the sequence $\{c_n\}$ is square summable, it can be expected that the energy of the error signal is bounded as $\|e\|^2 \leq const \cdot \tau^2$, or in terms of the oversampling ratio r,

$$\|e\|^2 \leq \frac{const}{r^2}. \tag{4.21}$$

This result is the content of the following theorem.

Theorem 9 *Let $f(t)$ be a real σ-bandlimited signal at the input of an A/D converter with a sampling interval $\tau < \pi/\sigma$. If the sequence of quantization threshold crossings of $f(t)$, $\{x_n\}$, forms a frame sequence for the space of σ-bandlimited signals, then there exists a positive constant δ such that if $\tau < \delta$, for every consistent estimate of $f(t)$, $g(t) \in C^1$,*

$$\|f(t) - g(t)\|^2 \leq k\|f(t)\|^2\tau^2, \tag{4.22}$$

where k is a constant which does not depend on τ. □

Proof Let A and B be bounds of the frame $\{e^{jx_n\omega}\}$ in $L^2[-\sigma, \sigma]$, so that for any σ-bandlimited $s(t)$

$$A\|s(t)\|^2 \leq \sum_n |s(x_n)|^2 \leq B\|s(t)\|^2. \tag{4.23}$$

At the points x_n, the error amplitude is bounded by the variation of the reconstructed signal on the corresponding sampling intervals, as described by the following relations:

$$\begin{aligned} |e(x_n)| &= |f(x_n) - g(x_n)| \\ &= |f(x_n) - g(y_n) + g(y_n) - g(x_n)| \\ &= |g(y_n) - g(x_n)| \\ &\leq \tau \cdot g'(\epsilon_n), \quad min(x_n, y_n) \leq \epsilon_n \leq max(x_n, y_n). \end{aligned} \tag{4.24}$$

Here, ϵ_n denotes a point on the interval $[y_n, x_n]$. The error norm then satisfies

$$\begin{aligned} \|e(t)\|^2 &\leq \frac{1}{A}\sum_n |e(x_n)|^2 \\ &\leq \frac{\tau^2}{A}\sum_n |g'(\epsilon_n)|^2. \end{aligned} \tag{4.25}$$

If τ is smaller than $\delta = \delta_{1/4}(\{x_n\}, \sigma)$ (see the corollary of Theorem 2), then $|\epsilon_n - x_n| < \delta$ and consequently

$$\sum_n |g'(\epsilon_n)|^2 \leq \frac{9B}{4}\|g'(t)\|^2. \tag{4.26}$$

This gives

$$\begin{aligned} \|e(t)\|^2 &\leq \tau^2 \frac{9B}{4A}\|g'(t)\|^2 \\ &\leq \tau^2\sigma^2 \frac{9B}{4A}\|g(t)\|^2 \end{aligned} \tag{4.27}$$

so that the energy of the error can be bounded as

$$\|e(t)\|^2 \leq \frac{\tau^2\sigma^2 9B}{4A}\|g(t)\|^2. \tag{4.28}$$

It remains to find a bound for the norm of $g(t)$.

According to Theorem 2, since $|x_n - y_n| < \tau < \delta_{1/4}(\{x_n\}, \sigma)$, the following holds

$$\begin{aligned} \|g(t)\|^2 &\leq \frac{4}{A}\sum_n g(y_n)^2 \\ &= \frac{4}{A}\sum_n f(x_n)^2 \\ &\leq \frac{4B}{A}\|f(t)\|^2. \end{aligned} \tag{4.29}$$

As a consequence of the last inequality and the error bound in (4.28) we obtain

$$\|e(t)\|^2 \leq 9\sigma^2 \frac{B^2}{A^2} \|f(t)\|^2 \tau^2. \tag{4.30}$$

□

In Appendix 4.5 we give another, more intuitive error bound for the case when quantization threshold crossings form a sequence of uniform density greater than σ/π.

An extension of the result of Theorem 9 to complex bandlimited signals is straightforward, provided that the real and imaginary parts are quantized separately.

Corollary 6 *Let the real and imaginary parts of a complex σ-bandlimited signal $f(t)$, at the input of an A/D converter with the sampling interval $\tau < \pi/\sigma$, be quantized separately. If sequences of quantization threshold crossings of both* $\mathrm{Re} f(t)$ *and* $\mathrm{Im} f(t)$ *form frame sequences for the space of σ-bandlimited signals, then there exists a positive constant δ such that if $\tau < \delta$ then for every consistent estimate $g(t) \in C^1$*

$$\|f(t) - g(t)\|^2 \leq k \|f(t)\|^2 \tau^2, \tag{4.31}$$

for some constant k which does not depend on τ. □

Proof Let $\{x_n^r\}$ and $\{x_n^i\}$ be the sequences of quantization threshold crossings of the real and imaginary parts of $f(t)$, respectively, and $0 < A_r \leq B_r < \infty$, $0 < A_i \leq B_i < \infty$ be the corresponding frame bounds. Then the relation (4.31) holds for

$$\tau < \min\left(\delta_{1/4}\left(\{x_n^r\}, \sigma\right), \delta_{1/4}\left(\{x_n^i\}, \sigma\right)\right)$$

and

$$k = 9\sigma^2\mu, \quad \mu = \max\left(\frac{B_i^2}{A_i^2}, \frac{B_r^2}{A_r^2}\right). \tag{4.32}$$

□

4.3.4 Error-Rate Characteristics with Optimal Reconstruction and Efficient Coding

An efficient scheme for lossless encoding of the digital representation follows from the observation that, for sufficiently fine sampling, the quantized values of signal

samples can be determined from the corresponding sequence of quantization threshold crossings.

Consider a bandlimited signal $f(t)$ and suppose that its quantization threshold crossings are separated, that is there is an $\epsilon > 0$ such that no two threshold crossings are closer than ϵ. Note that as a bandlimited signal of a finite energy $f(t)$ has a bounded slope so that there is always an $\epsilon_1 > 0$ such that $f(t)$ can not go through more than one quantization threshold on any interval shorter than ϵ_1. The condition for separated quantization threshold crossings requires in addition that intervals between consecutive crossings through a same threshold are limited from below away from zero. For a sampling frequency higher than $1/\epsilon$, all quantization threshold crossings of $f(t)$ occur in distinct sampling intervals. Under this condition quantized samples of $f(t)$ are completely determined by its sequence of quantization threshold crossings (see Figure 4.8). Another effect of high oversampling is that quantized values of consecutive samples differ with a small probability. In this case, an economical digital representation would encode incidences of data changes, that is, sampling intervals where quantization threshold crossings occur rather than quantized samples themselves.

Quantization threshold crossings can be grouped on consecutive time intervals of a given length, for instance T. For each of the crossings at most $1 + \log_2(T/\tau)$ bits are then needed to record its position inside the interval T. The height of the threshold crossing can be given with respect to the previous one, so that for this information only one additional bit is needed, to denote direction of the crossing (upwards or downwards). Hence, for recording of the information on quantization threshold crossings on an interval where Q of them occur, $Q\,(2 + \log_2(T/\tau))$ bits are needed (see Figure 4.9). The bit rate is then bounded as

$$B \leq \frac{Q_m}{T}\left(2 + \log_2\left(\frac{T}{\tau}\right)\right), \tag{4.33}$$

where Q_m denotes the maximal number of the crossings on an interval of length T. Note that if the samples themselves are recorded the bit rate increases linearly with the oversampling ratio, so quantization threshold crossings encoding is substantially

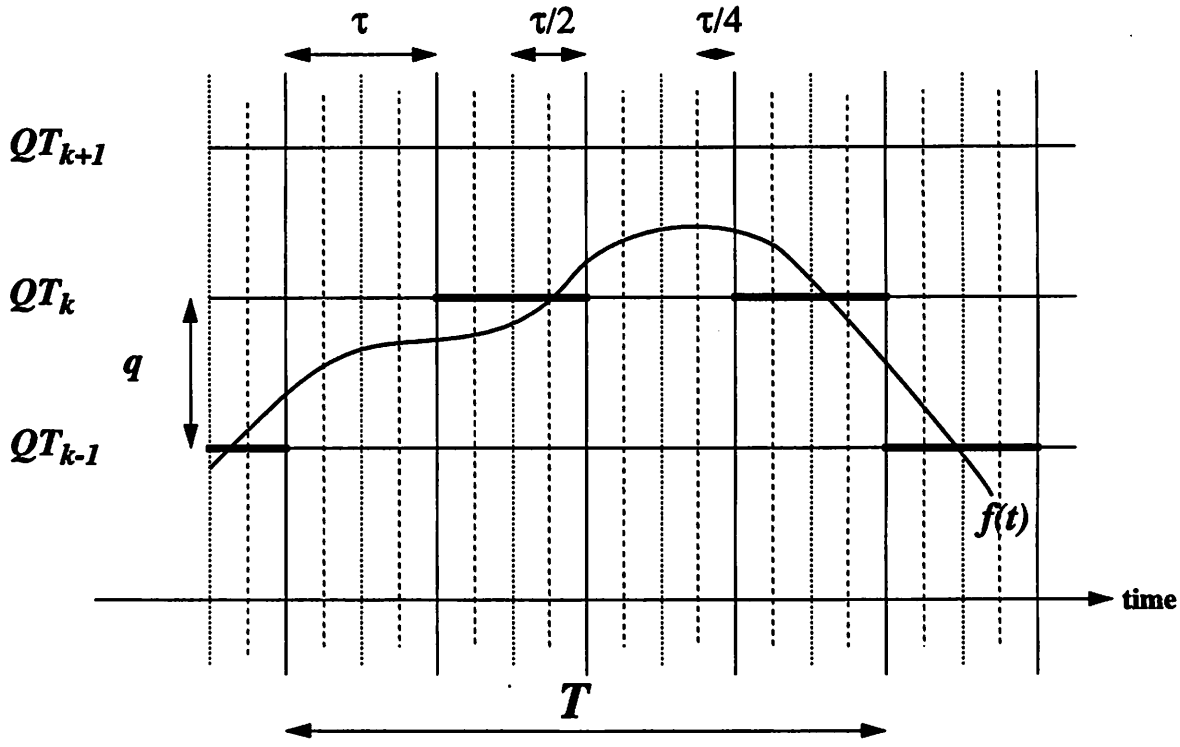

Figure 4.9: *Quantization threshold crossings encoding. Quantization threshold crossings are grouped on intervals of a length T. Refining the sampling interval by a factor 2^k requires additional k bits per quantization threshold crossing to encode its position inside the interval T.*

more efficient. If this efficient coding is used together with consistent reconstruction, then the error-rate characteristics of the oversampled A/D conversion becomes

$$\|e(t)\|^2 = KT^2 2^{-2\frac{T}{Q_m}B}, \tag{4.34}$$

where K is a constant which depends on the input signal, as follows from (4.22) and (4.33).

In order to estimate factor the T/Q_m in this expression, we consider again the two types of quantization threshold crossings, denoting them as d-crossings and s-crossings. A quantization threshold crossing is said to be a d-crossing if it is preceded by a crossing of a different quantization threshold, or an s-crossing if it is preceded by a crossing of the same threshold (see Figure 4.10). The total number of quantization threshold crossings of a σ-bandlimited signal $f(t)$ on an interval T is sum of these

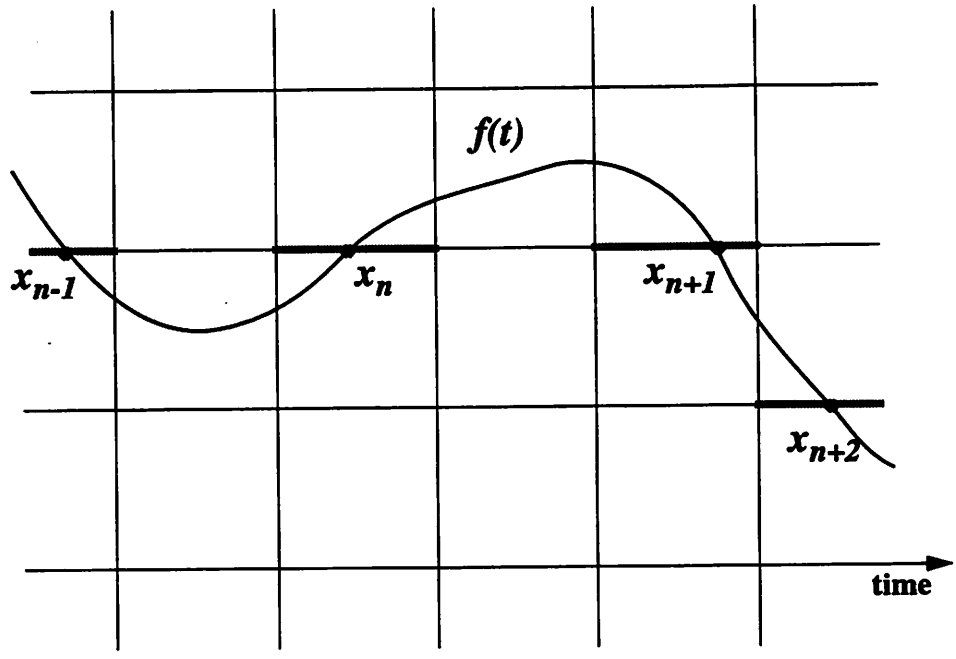

Figure 4.10: *Quantization threshold crossing types. A quantization threshold crossing can be immediately preceded by a crossing of the same quantization threshold, as illustrated by the crossings at points x_n and x_{n+1}. Such threshold crossings are denoted as s-crossings, and each of them is preceded by a point where the consider signal assumes an extremum. The other type of quantization threshold crossings, d-crossings, are those which occur after a crossing of a different quantization threshold. The threshold crossing at x_{n+2} is of this type.*

two types of crossings. The count of d-crossings, Q_d, depends on the slope of $f(t)$ as well as the quantization step size, q. The slope of $f(t)$ can be bounded as

$$|f'(t)| \leq \|f\|\sigma^{\frac{3}{2}}, \tag{4.35}$$

which gives

$$\frac{Q_d}{T} \leq \frac{\sigma^{\frac{3}{2}}}{q}\|f\|. \tag{4.36}$$

For the count of s-crossings, Q_s, we can investigate average behavior. Each of the s-crossings of $f(t)$ is preceded by a point where $f(t)$ assumes a local extremum. If s-crossings constitute a sequence of a uniform density d, then there is a subset of zeros of $f'(t)$ which also constitute a sequence of uniform density d. According to Theorem 3, $d \leq \sigma/\pi$ except in a degenerate case when $f'(t)$ is identically equal to

zero. If we assume that the sequence of quantization threshold crossings of $f(t)$ is a realization of an ergodic process, then as the interval T grows

$$\frac{Q_s}{T} \to c \leq \frac{\sigma}{\pi}. \tag{4.37}$$

Hence the error rate characteristics are given by $\|e(t)\|^2 = O(2^{-\frac{2}{\alpha_1+\alpha_2}B})$ where $\alpha_1 \leq \|f\|\sigma^{\frac{3}{2}}/q$ and α_2 is "close" to σ/π. Recall that for an A/D converter with a fixed sampling frequency $f_s \geq \sigma/\pi$ and quantization step refinement the mean squared error is given by $E(e(t)^2) = O(2^{-\frac{2}{\alpha}B})$, where $\alpha = f_s$.

It is interesting to find the error-rate characteristics of oversampled A/D conversion for the four combinations of reconstruction and encoding, i.e. linear versus consistent reconstruction, and PCM versus quantization threshold crossings encoding. These are given in Table 1.

Table 4.1: *Error-rate characteristics of oversampled A/D conversion, as the sampling interval tends to zero, for the four different combinations of reconstruction and encoding. The quantization error, $e(t)$, is expressed as a function of the bit rate, B.*

	Linear Reconstruction	Consistent Reconstruction
PCM	$E(e(t)^2) = O\left(\frac{1}{B}\right)$	$\|e(t)\|^2 = O\left(\frac{1}{B^2}\right)$
Efficient Encoding	$E(e(t)^2) = O\left(2^{-\frac{T}{Q_m}B}\right)$	$\|e(t)\|^2 = O\left(2^{-2\frac{T}{Q_m}B}\right)$

These results demonstrate the importance of an appropriate lossless coding of the digital representation. According to the given characteristics, for good error-rate performance of oversampled A/D conversion, efficient coding is more important than optimal reconstruction.

The most successful technique for high resolution analog to digital conversion is currently Sigma-Delta modulation, which achieves high error reduction gains based on a sophisticated exploitation of oversampling. The error reduction property of an n-th

order Sigma-Delta modulator is characterized by the relation $E(e(t)^2) = O(\tau^{2n+1})$, in the case of classical reconstruction. However, it has been shown in [45] that in the case of periodic bandlimited signals this result can be improved with a consistent reconstruction which gives $E(e(t))^2 = O(\tau^{2n+2})$. The error-rate characteristic of the Sigma-Delta modulator, with a consistent reconstruction, is then given by $E(e(t))^2 = O(B^{-(2n+2)})$. It can be observed that it is asymptotically outperformed by the simple A/D conversion, even with the linear reconstruction, provided that efficient coding is used.

4.4 $O(1/r^2)$ Error Behavior in Quantization of Weyl-Heisenberg Frame Expansions

Error analysis in the case of quantization of Weyl-Heisenberg frame expansions is an immediate generalization of the results on oversampled A/D conversion. The two cases which are considered first are: 1) frames derived from bandlimited window functions without restrictions on input signals, except that they are in $L^2(\mathbf{R})$; 2) timelimited input signals, with no restrictions on the window function except for the requirement that it is in $L^2(\mathbf{R})$.

CASE 1. σ-bandlimited window function

Let

$$\left\{\varphi_{m,n}(t) : \varphi_{m,n}(t) = \varphi(t - nt_0)e^{jm\omega_0 t}\right\} \tag{4.38}$$

be a Weyl-Heisenberg frame in $L^2(\mathbf{R})$, with the bounds

$$A\|f\|^2 \leq \sum_{m,n\in\mathbf{Z}} |\langle\varphi_{m,n}, f\rangle|^2 \leq B\|f\|^2. \tag{4.39}$$

Frame coefficients $\{c_{m,n} : c_{m,n} = \langle\varphi_{m,n}, f\rangle\}$ of a signal f can be expressed in the Fourier domain as

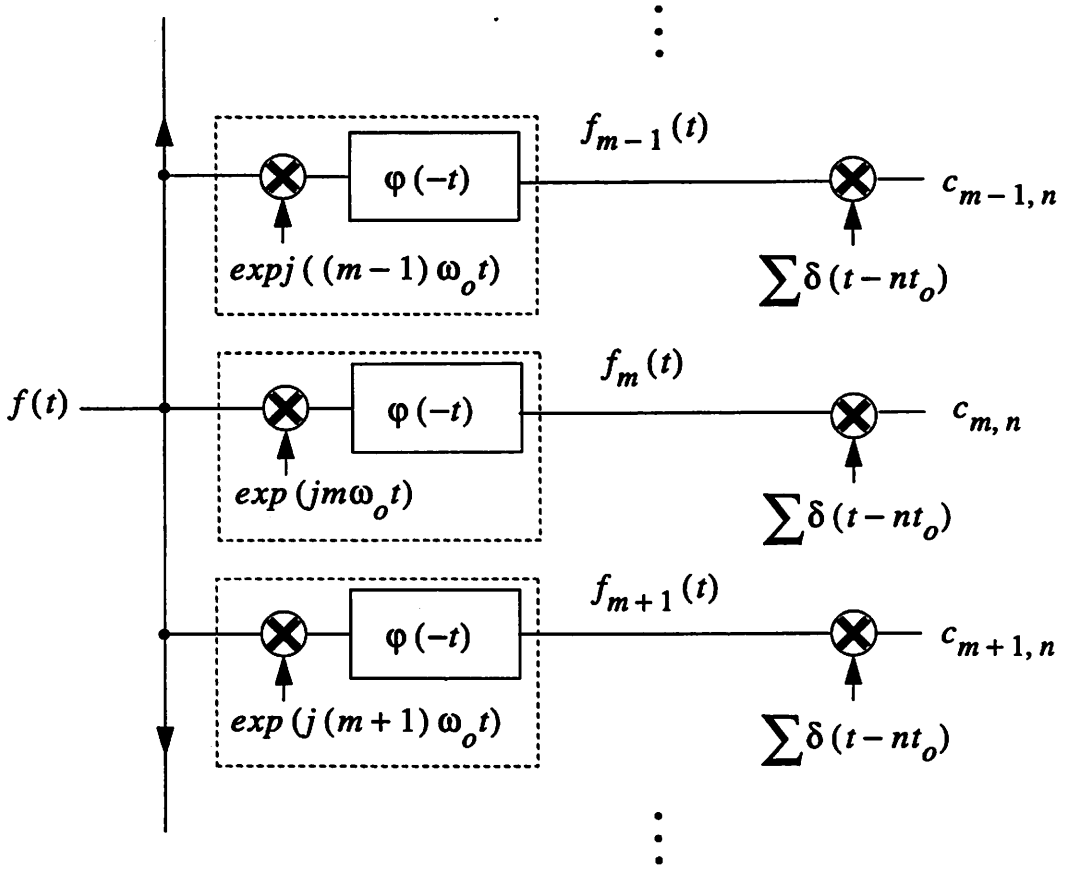

Figure 4.11: *Evaluation of Weyl-Heisenberg frame coefficients. For a fixed m, coefficients $c_{m,n}$ are obtained as samples, with t_0 sampling interval, of the signal $f_m(t)$, which is the result of modulating the input signal with $e^{jm\omega_0 t}$, followed by filtering with $\varphi(-t)$.*

$$c_{m,n} = \int_{-\infty}^{\infty} \hat{f}(\omega - m\omega_0)\hat{\varphi}^*(\omega)e^{j\omega nt_0}d\omega. \tag{4.40}$$

The system which for an input signal gives these coefficients can be viewed as a multichannel system, containing a separate channel for each frequency shift $m\omega_0$, such that the m-th channel performs modulation of an input signal with $e^{jm\omega_0 t}$, then linear filtering with $\varphi(-t)$ and finally sampling at points $\{nt_0\}$. Such a system is shown in Figure 4.11. For a fixed m, coefficients $c_{m,n}$ are samples of the signal

$$f_m(t) = \left(f(t)e^{jm\omega_0 t}\right) * \varphi(-t), \tag{4.41}$$

which will be called the *m-th subband component* of $f(t)$. In the sequel, the m-th subband component of a signal $s(t)$ will be denoted by $s_m(t)$ and the sequence obtained by sampling $s_m(t)$ with the interval t_0 will be denoted by $S_m[n]$, $S_m[n] = s_m(nt_0)$. Using this notation, frame coefficients of $f(t)$ are given by $c_{m,n} = F_m[n]$.

Such an interpretation of Weyl-Heisenberg frame coefficients of the signal $f(t)$ means that their quantization amounts to simple A/D conversion of the subband components of $f(t)$. Note that these coefficients are in general complex, and it is assumed here that real and imaginary parts are quantized separately. If the frame window, $\varphi(t)$, is a σ-bandlimited function, each of the subband components is also a σ-bandlimited signal. In this context, a signal $g(t)$ is said to be a consistent estimate of $f(t)$ if they have the same quantized values of the frame coefficients and each subband component of $g(t)$ is continuously differentiable, $g_m(t) \in C^1$ (note that the subband signals, being bandlimited, are continuously differentiable a.e.). According to the results from the previous section, this indicates that if the frame redundancy is increased by decreasing the time step t_0 for a fixed ω_0, the quantization error of consistent reconstruction should decay as $O(t_0^2)$. This result is established by the following corollary of Theorem 9, and can be expressed in terms of the oversampling ratio, $r = \frac{2\pi}{\omega_0 t_0}$, as $\|e\|^2 = O(1/r^2)$.

Corollary 7 *Let $\{\varphi_{m,n}(t)\}$ be a Weyl-Heisenberg frame in $L^2(\mathbf{R})$, with time step t_0 and frequency step ω_0, derived from a σ-bandlimited window function $\varphi(t)$. Consider quantization of the frame coefficients of a signal $f(t) \in L^2(\mathbf{R})$ and suppose that for a certain ω_0 the following hold:*

i) quantization threshold crossings of both real and imaginary parts of all the subband components $f_m(t) = (f(t)e^{jm\omega_0 t}) * \varphi(-t)$ form frame sequences for the space of σ-bandlimited signals, with frame bounds $0 < \alpha_m^r \leq \beta_m^r < \infty$ and $0 < \alpha_m^i \leq \beta_m^i < \infty$;

ii)

$$\sup_{m \in \mathbf{Z}} \max \left(\frac{\beta_m^r}{\alpha_m^r}, \frac{\beta_m^i}{\alpha_m^i} \right) = M < \infty.$$

Then there exists a constant δ, such that if $t_0 < \delta$, for any consistent estimate $g(t)$ of $f(t)$, the reconstruction error satisfies

$$\|f(t) - g(t)\|^2 \leq k\|f(t)\|^2 t_0^2, \tag{4.42}$$

where k is a constant which does not depend on t_0. □

Proof Let $f(t)$ be reconstructed from its quantized coefficients as $g(t)$. Suppose that $g(t)$ is a consistent estimate of $f(t)$, that is, frame coefficients of $g(t)$ are quantized to the same values as those of $f(t)$. Under the bandlimitedness condition on $\varphi(t)$, all subband signals $f_m(t)$ are also σ-bandlimited, and each $g_m(t)$ is a consistent estimate of the corresponding $f_m(t)$.

Under assumption *i)*, and as a consequence of Corollary 7, for each m there is a δ_m such that the m-th subband error component, $f_m(t) - g_m(t)$, satisfies

$$\|f_m(t) - g_m(t)\|^2 \leq 9\sigma^2 \mu_m \|f_m(t)\|^2 t_0^2, \quad \mu_m = \max\left(\left(\frac{\beta_m^r}{\alpha_m^r}\right)^2, \left(\frac{\beta_m^i}{\alpha_m^i}\right)^2\right). \tag{4.43}$$

For a sampling interval $t_0 \leq \pi/\sigma$ and any $s \in L^2(\mathbf{R})$, norms of subband signals s_m and their sampled versions satisfy

$$\|s_m\|^2 = t_0 \|S_m\|^2. \tag{4.44}$$

The frame condition (4.39) then implies

$$\begin{aligned} \|s\|^2 &\leq \frac{1}{A}\sum_m \|S_m\|^2 \\ &= \frac{1}{At_0}\sum_m \|s_m\|^2, \end{aligned} \tag{4.45}$$

and

$$\begin{aligned} \|s\|^2 &\geq \frac{1}{B}\sum_m \|S_m\|^2 \\ &= \frac{1}{Bt_0}\sum_m \|s_m\|^2. \end{aligned} \tag{4.46}$$

From assumption *ii)* and Theorem 2, it follows that $\delta = \inf_{m\in\mathbf{Z}} \delta_m$ is strictly greater than zero, $\delta > 0$. For $t_0 \leq \delta$, we have the following as a consequence of relations (4.43), (4.45) and (4.46):

$$\begin{aligned} \|f(t) - g(t)\|^2 &\leq \frac{1}{At_0} \sum_m \|f_m(t) - g_m(t)\|^2 \\ &\leq \frac{1}{At_0} \sum_m 9\sigma^2 \mu_m \|f_m(t)\|^2 t_0^2 \\ &\leq 9\sigma^2 M^2 \frac{B}{A} \|f\|^2 t_0^2. \end{aligned} \tag{4.47}$$

This finally gives

$$\|f(t) - g(t)\|^2 \leq k\|f(t)\|^2 t_0^2, \tag{4.48}$$

which for a constant ω_0 can be expressed as $\|f(t) - g(t)\|^2 = O(1/r^2)$. □

CASE 2. T-timelimited signals

Another way to interpret expansion coefficients of $f(t)$ with respect to the frame (4.38) is to consider them as samples of signals

$$\hat{\varphi}_n(\omega) = \left(\hat{\varphi}(\omega) e^{-jnt_0\omega}\right) * \hat{f}(\omega), \tag{4.49}$$

with ω_0 sampling interval. This is illustrated in Figure 4.12. Suppose that $f(t)$ is a T-timelimited signal. Then $\mathcal{F}\{\hat{\varphi}_n(\omega)\}$ is also T-timelimited for each n, which makes all subband signals $\hat{\varphi}_n(\omega)$ bandlimited. From an argument completely analogous to the one in the previous case, it can be concluded that if for some fixed t_0 sequences of quantization threshold crossings of subband signals $\hat{\varphi}(\omega)$ satisfy certain frame properties, the quantization error of consistent reconstruction can be bounded as $\|e(t)\|^2 \leq k\|f(t)\|^2 \omega_0^2$. Since we consider the case when $t_0 = const$ and $\omega_0 \to 0$, this can also be expressed as $\|e\|^2 = O(1/r^2)$. The precise formulation of this result is as follows.

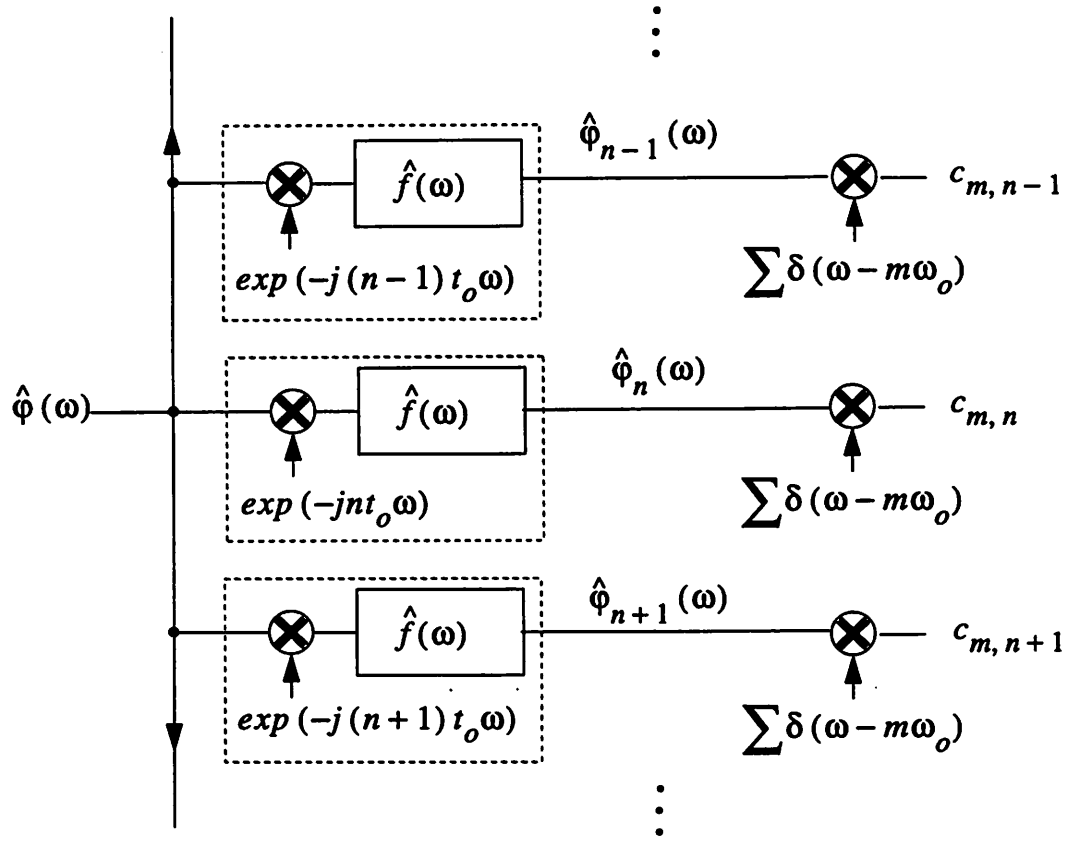

Figure 4.12: *Evaluation of Weyl-Heisenberg frame coefficients. For a fixed n, coefficients $c_{m,n}$ are obtained as samples, with ω_0 sampling interval, of the signal $\hat{\varphi}_n(\omega)$, which is itself the result of modulating $\hat{\varphi}(\omega)$ by $e^{-jnt_0\omega}$, followed by filtering with $\hat{f}(\omega)$.*

Corollary 8 *Let $\{\varphi_{m,n}(t)\}$ be a Weyl-Heisenberg frame in $L^2(\mathbf{R})$, with time step t_0 and frequency step ω_0, derived from a window function $\varphi(t)$. Consider quantization of the frame coefficients of a T-timelimited signal $f(t) \in L^2(\mathbf{R})$. Suppose that for a certain t_0 the following hold:*

i) quantization threshold crossings of both real and imaginary parts of all the subband components $\hat{\varphi}_n(\omega) = (\hat{\varphi}(\omega)e^{-jnt_0\omega}) * \hat{f}(\omega)$ form frame sequences for the space of T-timelimited signals, with frame bounds $0 < \alpha_n^r \leq \beta_n^r < \infty$ and $0 < \alpha_n^i \leq \beta_n^i < \infty$;

ii)

$$\sup_{n\in\mathbf{Z}} \max\left(\frac{\beta_n^r}{\alpha_n^r}, \frac{\beta_n^i}{\alpha_n^i}\right) = M < \infty.$$

Then there exists a constant δ, such that if $\omega_0 < \delta$, for any consistent estimate $g(t)$ of $f(t)$, the reconstruction error satisfies

$$\|f(t) - g(t)\|^2 \leq k\|f(t)\|^2\omega_0^2, \tag{4.50}$$

where k is a constant which does not depend on ω_0. □

If the Weyl-Heisenberg frame is redefined as

$$\left\{\varphi_{m,n}(t) : \varphi_{m,n}(t) = \varphi(t - nt_0)e^{jm\omega_0(t-nt_0)}\right\},$$

then analogous results hold for the cases when input signals are bandlimited or the window function has a compact support in time.

The assumptions on bounded support of frame window functions or input signals in either time or frequency, introduced in the above considerations, are natural assumptions of time-frequency localized signal analysis. A question which naturally arises is whether in the case when both the window function is bandlimited and the considered signals have finite support, the error decays as $\|e\|^2 = O(\omega_0^2 t_0^2)$. Another interesting case is when none of these assumptions is introduced. Is it then also possible to exploit frame redundancy for quantization error reduction so that the error norm tends to zero as the redundancy is increased, or even more $\|e\|^2 = O(1/r^2)$? These are still open problems.

4.5 Appendix

4.5.1 Alternative Proof of Theorem 9

For this proof we need following result by Duffin and Schaeffer [13].

Lemma 1 *[13] Let $\{e^{j\lambda_n\omega}\}$ be a frame in $L^2[-\sigma,\sigma]$. If M is any constant and $\{\mu_n\}$ is a sequence satisfying $|\mu_n - \lambda_n| \leq M$, then there is a number $C = C(M,\sigma,\{\lambda_n\})$ such that*

$$\frac{\sum_n |f(\mu_n)|^2}{\sum_n |f(\lambda_n)|^2} \leq C \tag{4.51}$$

for every σ-bandlimited signal $f(t)$. □

Recall that in order to have a sequence of quantization threshold crossings of a σ-bandlimited signal $f(t)$, having a uniform density $d > \sigma/\pi$, the quantization step q must change in time following the decay of the signal. Since $f(t)$ is square integrable, $q = q(t)$ has to decay at least as fast as $1/t$, although it can be fixed on a given set of time segments.

Let A and B denote again the lower and the upper bound, respectively, of the frame determined by the quantization threshold crossings, $\{x_n\}$, of $f(t)$, and let $g(t)$ be a consistent estimate of $f(t)$. Following the discussion in the proof of Theorem 9 (see (4.25)), the error norm is bounded as

$$\|e(t)\|^2 < \frac{\tau^2}{A}\sum_n |g'(\epsilon_n)|^2. \tag{4.52}$$

Since τ is smaller then the Nyquist sampling interval, $|x_n - \epsilon_n| < \pi/\sigma$ for all n. According to Lemma 1, there exists a constant $C = C(\pi/\sigma,\sigma,\{x_n\})$ such that for all sampling intervals $\tau < \pi/\sigma$

$$\sum_n |g'(\epsilon_n)|^2 \leq C\sum_n |g'(x_n)|^2. \tag{4.53}$$

This gives

$$\|e(t)\|^2 \leq \frac{\sigma^2 BC\tau^2}{A}\|g(t)\|^2. \tag{4.54}$$

Being a consistent estimate of $f(t)$, $g(t)$ can not differ from $f(t)$ by more than $q(n\tau)$, at time instants $\{n\tau\}$. The energy of $g(t)$ can be bounded, by considering its samples at these points, as

$$\|g(t)\|^2 = \tau\sum_n |g(n\tau)|^2 \tag{4.55}$$

$$\leq \tau \sum_n \left(|f(n\tau)|^2 + 2|q(n\tau)f(n\tau)| + |q(n\tau)|^2 \right) \tag{4.56}$$

$$= \|f(t)\|^2 + E_s(\tau). \tag{4.57}$$

Note that

$$E_s(\tau) = \tau \sum_n \left(2|q(n\tau)g(n\tau)| + |q(n\tau)|^2 \right)$$

converges to

$$\lim_{\tau \to 0} E_s(\tau) = \int_{-\infty}^{+\infty} \left(2|q(t)g(t)| + |q(t)|^2 \right) dt, \tag{4.58}$$

which has to be finite since $q(t) = O\left(f(t)\right)$ and $f(t)$ is square integrable. Therefore, $E_s(\tau)$ has to be bounded by some E which does not depend on τ. This gives the following error bound

$$\|e(t)\|^2 \leq \frac{\sigma^2 BC}{A} \left(\|f(t)\|^2 + E \right) \tau^2. \tag{4.59}$$

Chapter 5

Applications

5.1 Foreword

Singularities based signal processing schemes, described in this chapter, illustrate several issues related to overcomplete expansions. They provide examples of applications where redundancy is used to improve accuracy of signal analysis, demonstrate advantages of design flexibility of overcomplete expansions. They also indicate a possibility of efficient signal compression resulting from a sophisticated selection of information provided by redundant expansions. A powerful framework for a number of relevant theoretical problems as well as applications related to signal characterization using information on its singularities was introduced by Mallat et al. [25, 26, 28] as the concept of wavelet modulus maxima and wavelet zero-crossings representations. For particular classes of wavelets, singular events are reflected in the wavelet transform, by inducing evolution of local modulus maxima or zero-crossings across the scales. The two representations, based on irregular sampling of a dyadic wavelet transform at either local extrema or zero-crossings, provide signal descriptions which give explicit information on its singularities and transient phenomena. Due to a fairly broad applicability of the concept, we take a general approach when discussing the two representations while pointing out the relations to concrete problems and applications. The emphasis is on our work, which is focused on algorithms for signal

reconstruction from the two representations and on the development of the concept in the framework of convex representations in $\ell^2(\mathbf{Z})$. Our work was influenced by results of Berman [4], who first posed the problem in purely discrete-time, using $\mathbf{R}^n$ as the model space.

Another application of a similar flavor, described in more detail, is an image interpolation algorithm developed in collaboration with Chang [6, 7]. The algorithm is locally adaptive and is aimed at good rendition of edges. Information needed for applying the appropriate local interpolation is obtained through an analysis of image singularities, based on a redundant wavelet expansion which is implemented using an iterated nonsubsampled filter bank.

5.2 Singularity Detection and Processing Using Wavelets

In this section we briefly review an excellent treatment of this topic given in [25, 26, 28].

5.2.1 Wavelet Transform and Multiscale Edge Detection

Even before the wavelet formalism was developed, multiscale techniques have been used in computer vision for edge detection. The approach is to smooth a signal $f(x) \in L^2(\mathbf{R})$ at different scales using the dilated versions

$$\theta_s(x) = \frac{1}{s}\theta\left(\frac{x}{s}\right)$$

of a low-pass filter $\theta(x)$ and detect the resulting points of sharp variations. These sharp variations of $f * \theta_s(x)$ are considered to be edges of $f(x)$ at the scale s. The purpose of the smoothing at a scale s is to remove signal fluctuations which are small relative to the analysis scale so that sharp variations of the larger structures are better visible. An inflection point of the smoothed signal $f*\theta_s(x)$ produces a local extremum in its first derivative or a zero-crossings in its second derivative (see Figure 5.1).

Hence, an edge in $f(x)$ can be detected as either a local extremum of $d(f * \theta(x))/dx$ or a zero-crossing of $d^2(f * \theta(x))/dx^2$. In a particular case when the smoothing function, $\theta(x)$ is a Gaussian, the extrema detection gives a Canny edge detector [5], while the zero-crossing scheme corresponds to the Marr-Hilderth edge detection [29]. Although the two detection schemes are similar the detection of local extrema has an advantage. A local extremum of $d(f * \theta(x))/dx$ can be either a local maximum or a minimum of the modulus of $d(f * \theta(x))/dx$. Modulus maxima of $d(f * \theta(x))/dx$ correspond to sharp variations of $f(x)$ while modulus minima correspond to points where $f(x)$ is smooth. Hence, the local extrema detection makes information on these different types of signal behavior explicit, whereas with the zero-crossings approach it is not so simple to distinguish between the two types of inflection points.

In order to cast these edge detection schemes into a wavelet transform framework consider wavelets

$$\psi^a(x) = \frac{d\theta(x)}{dx} \quad \text{and} \quad \psi^b(x) = \frac{d^2\theta(x)}{dx^2}. \tag{5.1}$$

For a wavelet $\psi(x)$, let $\psi_s(x)$ denote

$$\psi_s(x) = \frac{1}{s}\psi(\frac{x}{s}).$$

The wavelet transform of $f(x)$ with respect to wavelets $\psi^a(x)$ and $\psi^b(x)$ is defined by

$$W^a f(s,x) = f * \psi^a_s(x) \quad \text{and} \quad W^b f(s,x) = f * \psi^b_s(x), \tag{5.2}$$

respectively. Note that this is the definition of the continuous wavelet transform introduced in (3.47). The wavelet transforms $W^a f(s,x)$ and $W^b f(s,x)$ are respectively proportional to the first and second derivative of $f * \theta(x)$,

$$W^a f(s,x) = f * (s\frac{d\theta_s}{dx})(x) = s\frac{d}{dx}(f * \theta_s)(x), \tag{5.3}$$

$$W^b f(s,x) = f * (s^2\frac{d^2\theta_s}{dx^2})(x) = s^2\frac{d^2}{dx^2}(f * \theta_s)(x). \tag{5.4}$$

Therefore, a sharp signal variation produces modulus maxima in the wavelet transform for a wavelet which is the first derivative of a smoothing function, or zero-crossings if the wavelet is the second derivative of a smoothing function. The positions of these modulus maxima or zero-crossings constitute smooth curves in the

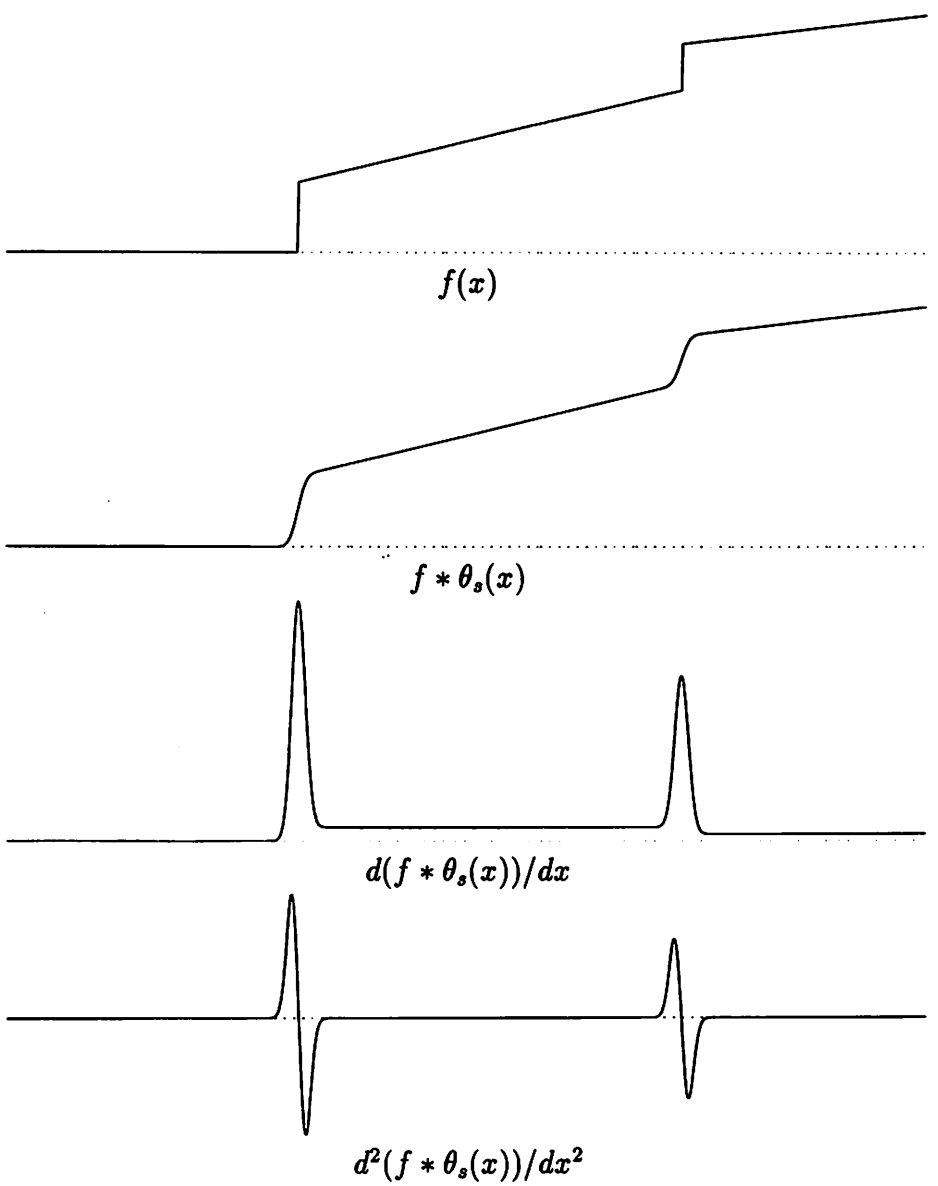

Figure 5.1: *Edge detection. Local extrema of the first derivative of a smoothed signal,* $d(f * \theta_s(x))/dx$*, indicate the presence of sharp variations in the signal. Alternatively, the sharp variations can be detected from zero-crossings of the second derivative of the smoothed signal,* $d^2(f * \theta_s(x))/dx^2$.

scale-space plane (s, x), called *finger prints* [52]. These facts can be demonstrated in a rigorous manner, but they are beyond the scope of this thesis. The purpose of this discussion is only to illustrate the applicability of the concept of wavelet transform modulus maxima to the problem of multiscale edge detection and to provide some intuition results which are presented in the following. The usefulness of wavelet transform in the detection of sharp signal variations is not limited only to being a reformulation of the classical results. Wavelet transform provides a powerful tool not only for detection but also for discrimination of different types of sharp variations and singular behaviors. This aspect of wavelet transform is the content of the next subsection.

5.2.2 Detection and Characterization of Singularities

The notion of a singular point is defined based on the Lipschitz regularity of a function. A function $f(x)$ is said to be *Lipschitz* α at x_0, where $n \leq \alpha \leq n+1$ for some positive integer n, if there exist two constants A and h_0 and a polynomial of order n, $P_n(h)$, such that for $h < h_0$

$$|f(x_0 + h) - P_n(h)| \leq A|h|^{\alpha}. \tag{5.5}$$

The function $f(x)$ is *uniformly Lipschitz* α over the interval (a, b) if there exists a constant A and for any $x_0 \in (a, b)$ there exists a polynomial $P_n(x)$ of order n, such that equation (5.5) is satisfied if $x_0 + h \in (a, b)$. The *Lipschitz regularity* of $f(x)$ at x_0 is the superior bound of all values α such that $f(x)$ is Lipschitz α at x_0. A function is said to be singular at x_0 if it is not Lipschitz 1 at x_0. Lipschitz regularity gives a characterization of function differentiability. It can be shown that a function $f(x)$ which is Lipschitz α, for $\alpha > n$ at a point, is n times differentiable at that point. A function which is continuously differentiable at x_0 is Lipschitz 1 at x_0. Even if the derivative of $f(x)$ is discontinuous but bounded at x_0, $f(x)$ is Lipschitz 1 at x_0 and therefore not considered to be singular at x_0.

Negative Lipschitz exponents for tempered distributions are defined by considering their primitives. Let α be a noninteger real number. A tempered distribution

$f(x)$ of a finite order is is said to be uniformly Lipschitz α on (a,b) if its primitive is uniformly Lipschitz $\alpha+1$ on (a,b). Let us illustrated these definitions in the case of a Dirac which is a frequently encountered tempered distribution. The second order primitive of a Dirac is a piecewise linear function. It is uniformly Lipschitz 1 in a neighborhood of $x=0$, and therefore uniformly Lipschitz α for $\alpha<1$. Consequently, in a neighborhood of 0 a Dirac is uniformly Lipschitz α for $\alpha<-1$. This is not sufficient to conclude that a Dirac is uniformly Lipschitz -1 in a neighborhood of 0, but based on the definition of local regularity we can say that the uniform Lipschitz regularity of a Dirac in a neighborhood of 0 is equal to -1. The Lipschitz characterization of singularities as is discussed so far has a global nature since it deals with uniform regularity of functions over intervals rather than points. In the applications which are considered in this chapter we are concerned with isolated singularities. The definition of Lipschitz exponents can be extended to isolated singularities, and it is said that a distribution $f(x)$ has an isolated singularity Lipschitz α at x_0, if $f(x)$ is uniformly Lipschitz α over an interval (a,b) containing x_0 and is uniformly Lipschitz 1 over any subinterval of (a,b) which does not include x_0. For instance, a Dirac centered at $x_0=0$ has an isolated singularity of Lipschitz regularity -1 at $x_0=0$, while a step function centered at 0,

$$h(x)=\begin{cases}0, & x<0\\ 1, & x\geq 0\end{cases} \tag{5.6}$$

has that of Lipschitz regularity 0.

The Lipschitz regularity of a function can be estimated by considering the asymptotic decay of its Fourier transform [34]. However, Fourier transform indicates only the global regularity of a function and cannot capture the local variation of regularity. On the other hand, the wavelet transform $Wf(s,x)=f*\psi_s(x)$, with respect to a compactly supported wavelet $\psi(x)$, depends only upon values of f in a neighborhood of x whose size is determined by the support of $\psi_s(x)$. The wavelet transform, considered over a range of small scales, is therefore well suited for detection of signal local characteristics. As an illustration, consider again a Dirac $\delta_{x_0}=\delta(x-x_0)$ and a step function $h_{x_0}(x)=h(x-x_0)$. The wavelet transform of the Dirac is the wavelet itself

translated by x_0

$$W\delta_{x_0}(s,x) = \psi_s(x - x_0) = \frac{1}{s}\psi(\frac{x - x_0}{s}). \tag{5.7}$$

From this equation we can observe two things. First, the modulus maxima of $W\delta(s,x)$ form a set of lines in the half-plane $(s,x)_{s>0,x\in\mathbf{R}}$ (scale-space of the wavelet transform) which originate at the point $x = x_0$ on the boundary of the half-plane ($s = 0$). Second, the asymptotic behavior of the values of the modulus maxima as the scale approaches zero can be characterized as $O(s^{-1})$. In the case of the step function, the wavelet transform has the form

$$Wh_{x_0}(s,x) = \theta(\frac{x - x_0}{s}), \tag{5.8}$$

where $\theta(x)$ is the primitive of the wavelet. Similarly to the wavelet transform of δ_{x_0}, Wh_{x_0} has modulus maxima which define lines in the scale-space originating from (or converging to) $(s,x) = (0,x_0)$, the location of the singularity. As opposed to the Dirac, the modulus maxima of Wh_{x_0} retain uniform values across all the scales, so that their asymptotic behavior when s tends to zero can be described as $O(s^0)$. We see that in these two simple cases, the *finger prints* of the wavelet transform modulus maxima point to locations of the isolated singularities and, furthermore, the asymptotic behavior of the modulus maxima values reveals the type of singularity. This observation has a more general scope, and is established below.

The potential of wavelet transform for characterizing local Lipschitz regularities in a given range depends on the number of vanishing moments of the wavelet. We say that a wavelet $\psi(x)$ has n vanishing moments if for all positive integers $k < n$

$$\int_{-\infty}^{\infty} x^k\psi(x)dx = 0. \tag{5.9}$$

The following theorem asserts that a function is uniformly Lipschitz α on a given interval if its wavelet transform with respect to a wavelet with n vanishing moments, $n > \alpha$, has no modulus maxima at small scales in that interval.

Theorem 10 *[26] Let n be a strictly positive integer and $\psi(x)$ a wavelet which has a compact support and n vanishing moments and is n times continuously differentiable. Let $f(x) \in L^1[a,b]$. If there exists a scale $s_0 > 0$ such that for all scales $s < s_0$ and*

$x \in (a,b)$, $|Wf(s,x)|$ has no local maxima, then for any $\epsilon > 0$ and $\alpha < n$, $f(x)$ is uniformly Lipschitz α in $(a+\epsilon, b-\epsilon)$. If $\psi(x)$ is the n-th derivative of a smoothing function, then $f(x)$ is uniformly Lipschitz n on any such interval $(a+\epsilon, b-\epsilon)$. □

The fact that singularities of a function can be located as the origins of the modulus maxima lines of its wavelet transform, as illustrated above, can be given a precise formulation. First we need to define the *closure* of the set of local modulus maxima as the set of points on the real line which are adherent points of the set of modulus maxima when the real line is considered as the boundary of the scale-space half-plane $(s,x)_{s>0,x\in\mathbf{R}}$. In other words, x_0 is in this closure if for any $\epsilon > 0$, there exists wavelet transform modulus maxima at a point (s_1, x_1) such that $|x_0 - x_1| < \epsilon$ and $s_1 < \epsilon$. Now we can formulate the following corollary of Theorem 10.

Corollary 9 *[26] The closure of the set of points where $f(x)$ is not Lipschitz n is included in the closure of the wavelet transform maxima of $f(x)$.* □

Isolated singularities can be discriminated based on the evolution of wavelet transform modulus maxima in the manner described by the following theorem. It is assumed that the wavelet has a compact support, is n-times continuously differentiable and has n vanishing moments.

Theorem 11 *[26] Let $f(x)$ be a tempered distribution whose wavelet transform is well defined over (a,b) and let $x_0 \in (a,b)$. Suppose that there exists a scale $s_0 > 0$ and a constant C such that for $x \in (a,b)$ and $s < s_0$ all the modulus maxima of $Wf(s,x)$ belong to a cone defined by $|x - x_0| \leq Cs$. Then, at all points $x_1 \in (a,b)$, $x_1 \neq x_0$, $f(x)$ is uniformly Lipschitz n in a neighborhood of x_1.*

Let $\alpha < n$ be a non-integer. The function $f(x)$ is Lipschitz α at x_0 if and only if there exists a constant A such that at each modulus maxima at (s,x) in the cone $|x - x_0| \leq Cs$

$$|Wf(s,x)| \leq As^{\alpha}. \tag{5.10}$$

□

Sharp variations of real signals are often not singularities such as a Dirac or a step function, but are rather their smoothed versions. In such cases it is usually important to estimate how smooth are these transitions. As suggested by Mallat et al. in [28], a smooth transition can be modeled as a singular function $u(x)$ convolved with a Gaussian of variance σ^2,

$$f(x) = u * g_\sigma(x), \tag{5.11}$$

where

$$g_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}. \tag{5.12}$$

If $u(x)$ has a singularity at x_0 of Lipschitz exponent α_0, then $f(x)$ has a sharp variation in a neighborhood of x_0 and that variation is characterized by α_0 as well as σ which indicates its smoothness. For a wavelet which is the first derivative of a smoothing function, the evolution of wavelet modulus maxima as the scale tends to zero can be well approximated by

$$|Wf(s,x)| \approx K s s_0^{\alpha_0 - 1}, \tag{5.13}$$

where $s_0 = \sqrt{s^2 + \sigma^2}$ [28]. Therefore, the decay of wavelet modulus maxima provides information on the underlying singularity as well as the Gaussian smoothing.

With this we end the overview of results on singularity characterization from wavelet transform modulus maxima. For more detail the reader is referred to [26].

5.3 Wavelet Modulus Maxima and Wavelet Zero-Crossings Representations

5.3.1 The Concept and Its Applicability

Evaluating the wavelet transform across a continuum of scales is computationally expensive and certainly unnecessary for many singularity based applications. Due to the convenience from its filter bank implementation the wavelet transform is usually computed at dyadic scales $s = 2^j$, $j \in \mathbf{N}$. The wavelet modulus maxima representation [26, 28] of a signal $f(x)$ is obtained by sampling the dyadic wavelet transform

$\{Wf(2^j,x)\}_{j\in\mathbf{N}}$ at the local maxima of $|Wf(2^j,x)|$, which captures information on its singularities or transient behaviors. This irregular sampling also overcomes the problem of the lack of shift invariance, which is one of the major drawbacks of wavelet series expansions which is lack of shift invariance. The wavelet zero-crossings representation [25] is obtained by recording zero-crossing positions of the dyadic wavelet transform and the integral values of $\{Wf(2^j,x)\}_{j\in\mathbf{N}}$ between consecutive pairs of zero-crossings. We saw earlier in this chapter that for a wavelet which is the second derivative of a smoothing function the zero-crossings of the wavelet transform correspond to signal inflection points. However, the zero-crossings themselves do not provide a stable characterization of the signal[1] [25] and the purpose of supplementing the information of multiscale zero-crossings with the integral values is to make the representation stable.

The concept of the wavelet modulus maxima and the wavelet zero-crossings representations is not only convenient for formulating some classical signal processing and computer vision problems, such as signal characterization from zero-crossings or multiscale edges and multiscale edge detection, but is also a powerful framework for studying these problems. An immediate benefit of the wavelet formalism is a systematic filter design procedure for multiscale schemes using iterated two channel filter banks, in the manner described in Section 3.5.1. Furthermore, the filter bank approach was also useful for a fast digital implementation of the underlying continuous time wavelet transform. An important issue related to multiscale edges or zero-crossings is to investigate whether they provide a unique characterization of signals and whether this characterization is stable. The wavelet formalism facilitated the design of novel efficient algorithms for signal reconstruction from multiscale edges or zero-crossings [25, 28, 4, 9] which was partly a result of attempts to experimentally assess the uniqueness and stability of the representations. Although experiments in reconstruction from wavelet modulus maxima and wavelet zero-crossings representations yielded very good results, it was demonstrated by Meyer [30] and Berman

[1]By *stable* we mean that a small perturbation of the representation cannot correspond to an arbitrarily large perturbation of the signal itself.

[4] that these representations in general do not provide a unique characterization of signals. Nonetheless, an answer of a relatively broad scope to the uniqueness problem has been reached. Furthermore, Berman has also proved certain stability properties of the representations [4].[2]

The wavelet modulus maxima representation was incorporated into novel schemes for signal denoising and compression. Both schemes were developed by Mallat, *et al.* [26, 28]. The signal denoising algorithm is based on analyzing the propagation across the scales of wavelet transform modulus maxima of a signal degraded by an additive white noise. This enables the discrimination of the modulus maxima corresponding to important signal structures from the maxima produced by the noise. The denoising then amounts to selecting signal related maxima and reconstructing from the partial wavelet modulus maxima representation. The details of the algorithm, as well as the experimental results demonstrating its effectiveness are given in [26]. The compression scheme is based on a similar idea. Data reduction is achieved by representing signals using partial wavelet modulus maxima representations. Compression factors of this scheme reported by Mallat et al. [28] were around 30 for 256×256 images without significant image degradation. The wavelet modulus maxima image compression scheme is actually a perceptual coding algorithm since it basically preserves information on edges and other sharp variations which are perceptually the most important image characteristics.

For both of these applications it is essential to have an effective and reliable reconstruction algorithm. In the following subsections we study the wavelet modulus maxima and the wavelet zero-crossings representations in the context of convex representations in $\ell^2(\mathbf{Z})$, which is a model space for digital implementation. One appealing property of convex representations is that the reconstruction problem can be at least theoretically easily solved using alternating projections onto convex sets. It turns out that in the case these two presentations this method yields simple and efficient reconstruction algorithms. Wavelet modulus maxima representation as discussed above is

[2]Note that Berman considers wavelet extrema and wavelet zero-crossings representations in $\mathbf{R}^n$, and his results pertain to this space.

not a convex representation, and therefore we consider wavelet extrema representation as its convex modification.

5.3.2 Discrete-Time Wavelet Extrema and Wavelet Zero-Crossings Representations

The wavelet transform for the wavelet modulus maxima or zero-crossings based schemes is implemented using an octave band analysis nonsubsampled filter bank, as discussed in Section 3.5.2 and shown also in Figure 5.2a. Such an analysis FIR filter bank is a bounded linear operator $\mathbf{W} : \ell^2(\mathbf{Z}) \to \ell^2(\mathbf{I})$, $\mathbf{I} = \{1, 2, ..., J+1\} \times \mathbf{Z}$, and is called the *discrete dyadic wavelet transform*. In this chapter we shall refer to the discrete dyadic wavelet transform as the wavelet transform. The operator $\mathbf{W}$ is an ensemble of $J+1$ linear operators $W_j : \ell^2(\mathbf{Z}) \to \ell^2(\mathbf{Z})$, $j = 1, 2, ..., J+1$. The operators W_j are convolution operators with filters

$$
\begin{aligned}
&V_1(z) = H_1(z), \\
&V_2(z) = H_0(z)H_1(z^2), \\
&\cdots \\
&V_J(z) = H_0(z)\cdots H_0(z^{2^{J-2}})H_1(z^{2^{J-1}}), \\
&V_{J+1}(z) = H_0(z)\cdots H_0(z^{2^{J-2}})H_0(z^{2^{J-1}}),
\end{aligned}
$$

respectively. In the following, signals in $\ell^2(\mathbf{Z})$ will be denoted by lower case letters, f, g,... and their wavelet transforms by the corresponding upper case letters, $F = \mathbf{W}f$, $G = \mathbf{W}g$,.... Any vector F in $\ell^2(\mathbf{I})$ will represent a $(J+1)$-tuple of vectors in $\ell^2(\mathbf{Z})$, $F = (F^1, F^2, \ldots, F^{J+1})$, so the jth component of $\mathbf{W}f$ will be denoted by $W_j f = F^j$.

In Section 3.5.2, we demonstrated that if a discrete-time signal $f[n]$ at the input of this filter bank is obtained by sampling a continuous time signal $f_c(x)$ prefiltered with an appropriate scaling function, then the sequences at the output of the filter bank are samples of the continuous-time wavelet transform of $f_c(x)$. In practice, the discrete-time version of a continuous-time signal is usually obtained by sampling the signal prefiltered by a low-pass filter which needs not be the required scaling

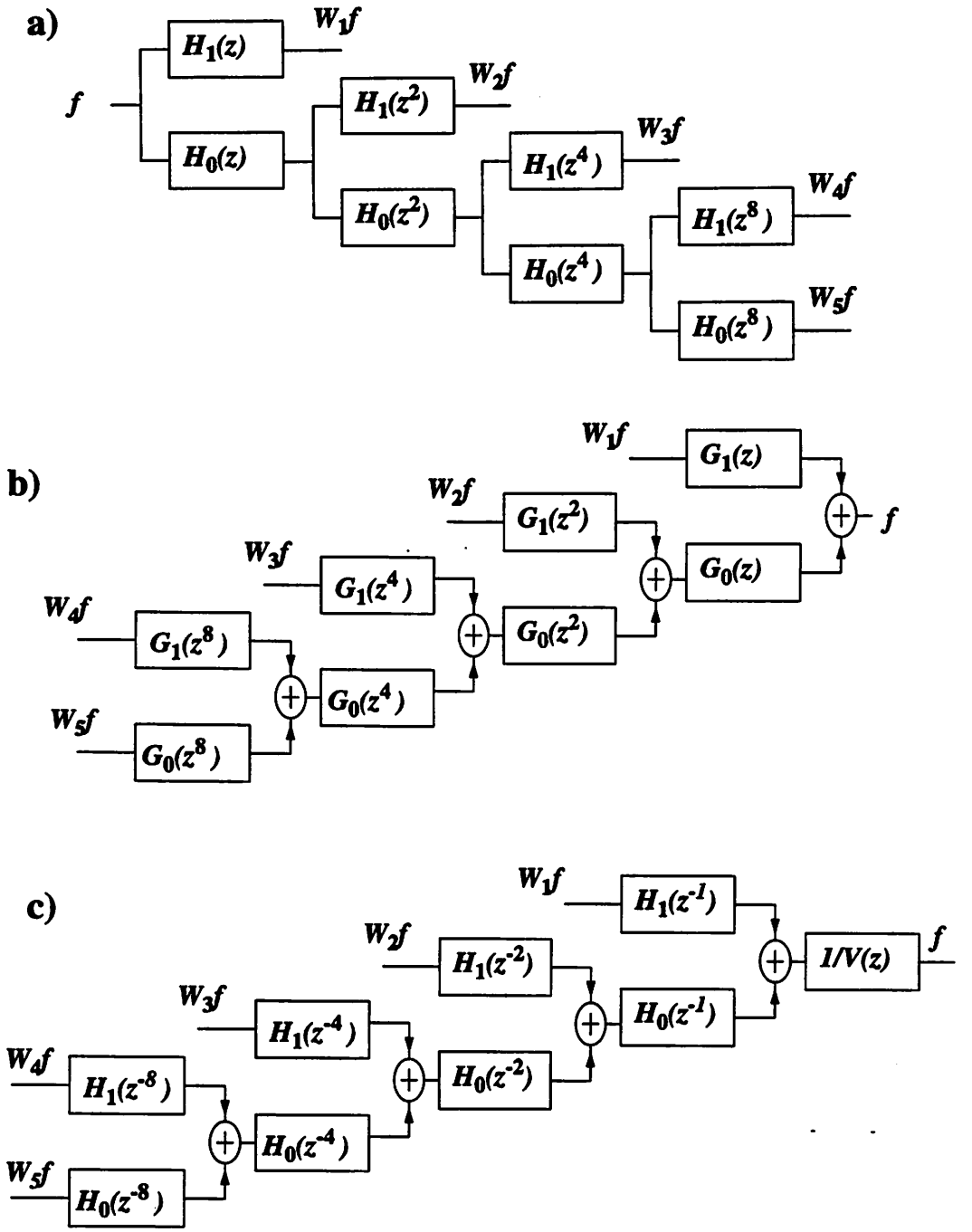

Figure 5.2: *The discrete dyadic wavelet transform. a) Implementation of the* $\mathbf{W}$ *operator; b) Implementation of an inverse of* $\mathbf{W}$*; c) Implementation of the* $\tilde{\mathbf{W}}^*$ *operator.*

function. However, if the scaling function is a reasonable low-pass filter, which is usually the case, the two methods of discretizations give approximately equal results. This illustrates a case where overcomplete expansions, implemented as nonsubsampled filter banks, are essential in providing a close approximation to continuous-time filtering. Redundancy is exploited here for improving the accuracy of signal analysis with respect to possible implementation based on critically sampled filter banks.

The definitions of the wavelet extrema and zero-crossings representations adopted here are essentially those introduced by Berman et al. [4]. Let M_a and M_i denote the operators which give the locations of local maxima and minima of a signal, respectively,

$$M_a f = \{k : f(k+1) \leq f(k),\ f(k-1) \leq f(k)\}, \tag{5.14}$$

$$M_i f = \{k : f(k+1) \geq f(k),\ f(k-1) \geq f(k)\}, \tag{5.15}$$

and M be the operator extracting signal values at its local extrema points

$$Mf = \{f(k),\ k \in M_a f \cup M_i f\}. \tag{5.16}$$

According to this notation, the wavelet extrema representation of a signal f is defined as

$$E_e f = \{M_a W_j f, M_i W_j f, M W_j f,\ j = 1, \ldots, J+1\}, \tag{5.17}$$

meaning that $E_e f$ consists of the indices of local extrema of $W_j f$ and the values of $W_j f$ at these points, for all scales $j = 1, 2, ..., J+1$.

The wavelet extrema representation contains information on both wavelet transform modulus maxima and minima. This does not necessarily lead to a significant increase in the number of points to be considered with respect to the wavelet modulus maxima representation as would appear at first. It turns out that most of the wavelet transform local extrema are actually modulus maxima (there are examples of signals for which the wavelet extrema and modulus maxima representations are the same). In experiments performed on lines from images and on randomly generated signals we observed that taking the modulus maxima instead of all local extrema reduces

the total number of points by only about 10%. Another example of a convex modification of the original modulus maxima scheme is a representation which includes information on positions of local modulus minima without coding their values. Also, either the whole low-pass signal $W_{J+1}f$ or only its local extrema could be kept for this representation. For the sake of conciseness we will adhere to the definition (5.17) since most of the results apply to the other cases as well.

The definition of wavelet zero-crossings representation requires the introduction of two more operators. Let Z denote the operator which provides information on the zero-crossing locations of a sequence,

$$Zf = \{k : f(k) \cdot f(k-1) \leq 0\}. \tag{5.18}$$

and let S be the operator which gives the integral values (the sum of points) between all pairs of consecutive zero-crossings of some sequence. If the total number of zero-crossings of f is denoted by $|Zf|$, and its kth zero-crossing by z_k, the S operator is defined as

$$Sf = \left\{ Sf(k) : Sf(k) = \sum_{j=z_{k-1}}^{z_k-1} f(j),\ k = 1, ..., |Zf| + 1 \right\}. \tag{5.19}$$

It is assumed here that the points $-\infty$ and $+\infty$ are also zero-crossings, denoted by z_0 and $z_{|Zf|+1}$, respectively. In order to ensure that $Sf(1)$ and $Sf(|Zf| + 1)$ are finite, we shall require that the signal f be integrable, that is $f \in \ell^1(\mathbf{Z})$. Therefore, the following definition of wavelet zero-crossings representation is valid for signals in $\ell^1(\mathbf{Z})$ (this is usually the case in practice, where signals with sufficient decay are encountered):

$$E_z f = \{ZW_j f, SW_j f,\ j = 1, \ldots, J+1\}. \tag{5.20}$$

In words, the zero-crossings representation, $E_z f$, consists of the indices of the zero-crossings of $W_j f$ and the integral values of $W_j f$ between consecutive zero-crossings, across all scales $j = 1, 2, ..., J + 1$.

It is worthwhile to point out the mutual relationship between wavelet extrema and wavelet zero-crossings representations. Consider the extrema representation $R_e f$

of some signal $f \in \ell^2(\mathbf{Z})$,

$$R_e f = \{M_a f, M_i f, Mf\} \tag{5.21}$$

and the zero-crossings representation $R_z \Delta f$ of its difference Δf,

$$\Delta f(n) = f(n+1) - f(n), \tag{5.22}$$

defined as

$$R_z \Delta f = \{Z\Delta f, S\Delta f\}. \tag{5.23}$$

According to the definitions of local extrema and zero-crossings, (5.14), (5.15), (5.18), the local extrema of f coincide with the zero-crossings of Δf:

$$M_a f \cup M_i f = Z\Delta f. \tag{5.24}$$

In addition to the equivalence between $M_a f \cup M_i f$ and $Z\Delta f$, Mf and $S\Delta f$ also provide equivalent information on the signal f. With z_k denoting the index (location) of the kth zero-crossing of Δf, $k = 1, 2, ..., |Z\Delta f|$, the following relations can be easily proven:

$$Mf = \left\{f(z_1), f(z_2), \ldots, f(z_{|Z\Delta f|})\right\}, \tag{5.25}$$

$$S\Delta f = \left\{\left(f(z_1) - f(-\infty)\right), \left(f(z_2) - f(z_1)\right), ..., \left(f(+\infty) - f(z_{|Z\Delta f|})\right)\right\}. \tag{5.26}$$

Since in most practical cases $f(-\infty) = 0$ and $f(+\infty) = 0$, information contained in Mf and $S\Delta f$ are equivalent, i.e. one uniquely determines the other. We can now state the relation between the two representations which is an immediate consequence of equalities (5.24), (5.25), (5.26), and the commutativity of the Δ and W_j operators.

Proposition 11 *For signals in $\ell^2(\mathbf{Z})$, the wavelet extrema representation and the wavelet zero-crossings representation of the signal's first difference (5.22) provide an equivalent characterizations of the signal. Consider an arbitrary signal $f \in \ell^2(\mathbf{Z})$ and its difference Δf. Any signal in the reconstruction set of Δf, from its wavelet zero-crossings representation, is the first difference of some signal in the reconstruction set of f, from its wavelet extrema representation. Conversely, the first difference of any signal in the reconstruction set of f, from its wavelet extrema representation, is in the reconstruction set of Δf, from its wavelet zero-crossings representation.* □

5.3.3 Filter Design

Besides giving samples of the continuous wavelet transform at a considerably denser grid than critically sampled filter banks, another important advantage of non-subsampled filter banks in this application is that they allow for design specifications which cannot be attained with critically sampled filter banks. This is demonstrated in this section. First, we review perfect reconstruction conditions on the filter banks for the discrete dyadic wavelet transform.

Although the purpose of the wavelet transform here is to obtain information on signal singularities, it is important that the $\mathbf{W}$ operator is invertible, or that the analysis filter bank allows for a stable perfect reconstruction of the input signal, since that means that during the analysis we do not lose any information of the signal. For perfect reconstruction from the discrete dyadic wavelet transform, $\mathbf{W}f$ based on prototype filters $H_0(z)$ and $H_1(z)$, it is necessary and sufficient that there exist two filters $G_0(z)$ and $G_1(z)$ satisfying

$$H_0(z)G_0(z) + H_1(z)G_1(z) = 1. \tag{5.27}$$

The inverse of the wavelet transform operator, $\mathbf{W}^{-1}$, can be implemented by the filter bank shown in Figure 5.2b, which will be referred to as the non-subsampled synthesis octave band filter bank. A stable reconstruction is possible if and only if the filters $H_0(z)$ and $H_1(z)$ do not have common zeros on the unit circle. For the reconstruction scheme using only FIR filters following FIR analysis, the perfect reconstruction condition (5.27) is equivalent to the constraint that $H_0(z)$ and $H_1(z)$ have no common zeros. These results also follow immediately from the more general considerations in Chapter 3. Recall that the synthesis filters $G_0(z)$ and $G_1(z)$, and therefore the inverse wavelet transform operator $\mathbf{W}^{-1}$, are not unique. An obvious solution for the reconstruction operator $\mathbf{W}^{-1}$ is represented by Figure 5.2c, where $V(z)$ is

$$V(z) = V_1(z)V_1(z^{-1}) + V_2(z)V_2(z^{-1}) + ... + V_{J+1}(z)V_{J+1}(z^{-1}). \tag{5.28}$$

It amounts to filtering the octave band components W_1f, W_2f, ... , $W_{J+1}f$ by $U_1(z) = \tilde{V}_1(z)/V(z)$, $U_2(z) = \tilde{V}_2(z)/V(z)$, ... ,$U_{J+1}(z) = \tilde{V}_{J+1}(z)/V(z)$ respectively,

and adding the resulting sequences. This inverse is the Hilbert adjoint of the dual of $\mathbf{W}$, and will be denoted by $\tilde{\mathbf{W}}^*$. In applications of wavelet modulus maxima representation, signals are usually reconstructed from a partial information of their wavelet transform. In such cases it is important to use $\tilde{\mathbf{W}}^*$ as an inverse of $\mathbf{W}$ since it reduces the error, whereas some other inverses can produce the opposite effect. Also, the algorithms for signal reconstruction from wavelet extrema or wavelet zero-crossings representations use iteratively the orthogonal projection operator onto the range of the wavelet transform, $P_V = \mathbf{W}\tilde{\mathbf{W}}^*$. For these reasons, it is desirable to deal with filter banks for which the corresponding $\tilde{\mathbf{W}}^*$ operator has an FIR implementation, which turns out to be possible if and only if $V(z) = 1$. From the considerations in Chapter 3 it follows that $\tilde{\mathbf{W}}^*$ can be implemented using FIR filters if and only if $H_0(z)$ and $H_1(z)$ are power complementary,

$$H_0(z)H_0(z^{-1}) + H_1(z)H_1(z^{-1}) = 1, \tag{5.29}$$

(see also [9]). In many signal processing tasks linear phase filters are desirable. However, the power complementary condition (5.29) excludes the possibility of non-trivial linear phase FIR designs [47]. If linear phase filters are to be used then it is important to use the synthesis filters $G_0(z)$ and $G_1(z)$ which closely approximate $H_0(z^{-1})$ and $H_1(z^{-1})$.

Another relevant feature of the wavelet transform extrema and zero-crossings schemes is the flatness of filters at zero and at half the sampling frequency. It is related to the number of vanishing moments and to the regularity of the wavelet used in the signal analysis implemented by the filter bank. We say that a filter $H(e^{j\omega})$ has flatness n at frequency ω_0 if $dH(e^{j\omega})/d\omega$ has a zero of multiplicity n at ω_0. If the multiplicity of the zero of $H_0(z)$ at $z = -1$ is N_0, then $H_0(z)$ has flatness $N_0 - 1$ at $\omega = \pi$. The regularity of the wavelet depends on the flatness of the low-pass filter $H_0(z)$ at $z = -1$. Roughly speaking, highly regular wavelets are obtained from filters with maximum number of zeros at $z = -1$. On the other hand, the number of vanishing moments of the wavelet is given by the multiplicity of the zero at $z = 1$ in the high-pass filter $H_1(z)$. In applications which are described in this chapter, we are interested in detecting discontinuities and sharp variations with Lipschitz exponents

smaller than 1. For that purpose, it is sufficient to use a wavelet with a single vanishing moment, that is a wavelet which is the first derivative of a smoothing function. If we want to relate sharp variations in the signal to zero-crossings of its wavelet transform then the wavelet should have two vanishing moments, i.e. it should be the second derivative of a smoothing function. Also, for detection and characterization of singularities with Lipschitz regularities smaller than 1, it is sufficient that the wavelet used is once continuously differentiable. However, in order to obtain clearer description and to facilitate detection of important singularities, it is advantageous that the wavelet does not have more vanishing moments than necessary, and that it is highly regular at the same time. These two requirements are contradictory in the case of wavelets derived from orthogonal (critically sampled) filter banks. However, this is not the case with wavelets derived from nonsubsampled filter banks. This is demonstrated below.

Consider a pair of filters, $H_0(z)$ and $H_1(z)$, satisfying the following conditions:

$$H_0(z)H_0(z^{-1}) + H_1(z)H_1(z^{-1}) = 1, \quad H_0(-1) = 0, \quad H_1(1) = 0. \tag{5.30}$$

If the multiplicity of the zero of $H_0(z)$ at $z = -1$ is N_0 and the multiplicity of the zero of $H_1(z)$ at $z = 1$ is N_1, then the filters have flatness $N_0 - 1$ at $\omega = \pi$ and flatness $N_1 - 1$ at $\omega = 0$. The issue is here to investigate how large N_0 can be given N_1 and the length of the filters, L. The following proposition holds.

Proposition 12 *For a pair of filters $H_0(z)$ and $H_1(z)$ of length L, satisfying (5.30), let N_0 be the multiplicity of the zero of $H_0(z)$ at $\omega = \pi$, and N_1 be the multiplicity of the zero of $H_1(z)$ at $\omega = 0$. It is possible to design $H_0(z)$ and $H_1(z)$ for any pair of N_0 and N_1 such that* $1 \leq N_0 < L$, $1 \leq N_1 < L$, *and* $N_0 + N_1 \leq L$. □

A constructive proof of the above proposition is given in Appendix 5.6.1. The low-pass filter of the maximally flat pair, the case when $N_0 + N_1 = L$, is obtained as a spectral factor of

$$H_0(z)H_0(z^{-1}) = P\left(\left(\frac{1-z}{2}\right)\left(\frac{1-z^{-1}}{2}\right)\right), \tag{5.31}$$

where $P(y)$ is given by (see Appendix 5.6.1)

$$P(y) = (1-y)^{N_0} \left(\sum_{l=0}^{L-1-N_0} \binom{N_0+l-1}{l} y^l \right). \tag{5.32}$$

The high-pass filter $H_1(z)$ is obtained from the factorization

$$H_1(z)H_1(z^{-1}) = 1 - P\left(\left(\frac{1-z}{2}\right)\left(\frac{1-z^{-1}}{2}\right)\right). \tag{5.33}$$

The regularity of wavelets derived from these filters can be estimated using Daubechies' criterion [10, 11]. According to this criterion, a wavelet derived from a filter bank with the low-pass filter

$$H_0(e^{j\omega}) = \left(\frac{1+e^{j\omega}}{2}\right)^{N_0} R(\omega), \tag{5.34}$$

is r times continuously differentiable if

$$B = \sup_{\omega\in[0,2\pi]}|R(\omega)| < 2^{N_0-r-1}. \tag{5.35}$$

For the maximally flat power complementary filter design, $N_0 + N_1 = L$,

$$B = \sup_{y\in[0,1]}\sqrt{Q(y)}, \tag{5.36}$$

where $Q(y)$ is the polynomial defined in (5.50), while its coefficients are given by (5.52) (see the proof of Proposition 12 in Appendix 5.6.1). The case of $N_1 = 1$, corresponds to wavelets with a single vanishing moment, $Q(y) \equiv 1$ and $N_0 = L-1$, implying that the derived wavelet is at least $L-3$ times continuously differentiable. This also follows from the theory of B-splines, since the function $\phi(x)$ in (3.43) in this case is a B-spline. Wavelets with two vanishing moments are obtained from filters with $N_1 = 2$. In this case, the low-pass filter has all but one of its zeros at $\omega = \pi$, and therefore from (5.50) and (5.52) it follows that $N_0 = L-2$, while $Q(y) = 1+(L-2)y$. Consequently $B = \sqrt{L-1}$ and Daubechies' criterion immediately proves that the derived wavelet is at least $L - 3 - \frac{1}{2}\log_2 L$ times continuously differentiable. Note that wavelets with a single or two vanishing moments generated from orthogonal filter banks cannot have such a high regularity. The reason is that in the orthogonal design

the multiplicity of the zero of the low-pass filter at $z = -1$ has to be equal to the multiplicity of the zero of the high-pass filter at $z = 1$, so that the requirements for a small number of vanishing moments and a high regularity are contradictory.

To achieve a sufficient regularity, the low-pass filter $H_0(z)$ does not necessarily have to be maximally flat at $\omega = \pi$. This brings additional freedom which can be used to meet other design specifications. We illustrate this point by the following design example.

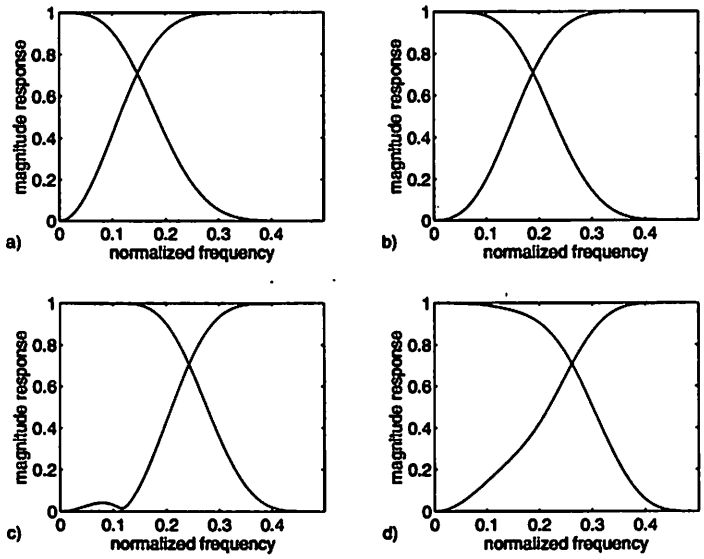

Figure 5.3: *Magnitude responses of power complementary filters for generating wavelets with two vanishing moments. The length of the filters is $L = 9$ and the high-pass filters have a zero of multiplicity $N_1 = 2$ at $z = 1$ in all cases. a) Maximally flat filters: low-pass filter has $N_0 = 7$ zeros at $z = -1$. b) Power complementary pair: low-pass filter with $N_0 = 6$ zeros at $z = -1$. c) Power complementary pair: low-pass filter with $N_0 = 5$ zeros at $z = -1$. d) Power complementary pair: low-pass filter with $N_0 = 4$ zeros at $z = -1$.*

Example 5 *Figure 5.3 shows magnitude responses of several power complementary filters for generating wavelets with two vanishing moments. All filters are of length $L = 9$ and differ in the multiplicity of the zero of the low-pass filter at $\omega = \pi$. The maximally flat pair is shown in Figure 5.3a. The design flexibility obtained by relaxing the maximally flat constraint can be used for attaining different bandwidths,*

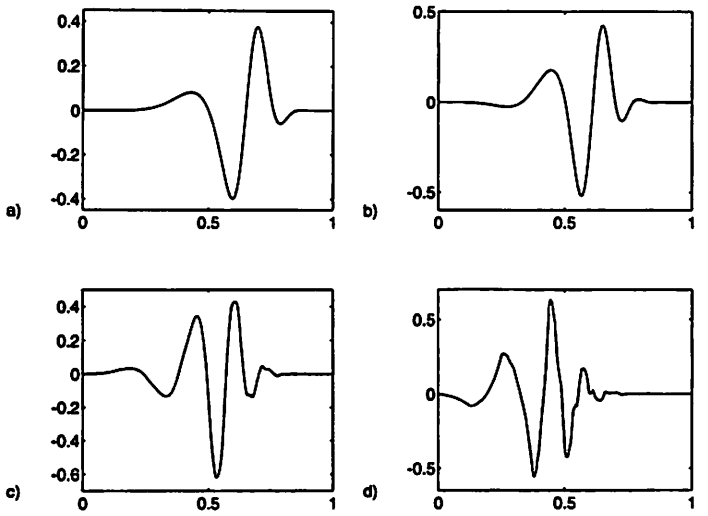

Figure 5.4: *Wavelets derived from the filters represented in Figure 5.3. Wavelets in figures a, b, c and d are derived from filters in figures 5.3a, 5.3b, 5.3c and 5.3d, respectively.*

as illustrated by the three other examples in the same figure. Figure 5.4 shows wavelets which are derived from these filters, in the manner described by equations (3.43) and (3.44). The coefficients of these filters are given in Table 1. □

5.3.4 Consistent Reconstruction

The wavelet extrema and zero-crossings representations of a signal define a number of convex constraints which the signal obeys. In general, an infinite number of signals satisfy the same set of constraints. The set of all such signals is called the *reconstruction set.* A consistent reconstruction strategy, which is our goal, means finding a signal in the reconstruction set since it satisfies all of the constraints and can not be distinguished from the original based on the representation. This notion is also important when quantization of extrema values (or integral values between zero-crossings) is used, as would be the case in coding. Recall that we used the notions of reconstruction set and consistent reconstruction with the same meaning in the context of quantization of overcomplete expansions in Chapter 4.

The reconstruction procedures described here actually recover the wavelet trans-

a		b	
H_0	H_1	H_0	H_1
0.37690273	0.53402011	0.03078682	0.22376086
-0.65503275	-0.61696397	-0.17864344	-0.53553401
0.09170122	-0.13570812	0.38821680	0.38072858
0.18384469	0.06867234	-0.47681949	0.05794996
0.05438736	0.08732849	0.12287851	-0.14190924
-0.02189883	0.04577467	0.22906289	-0.04590050
-0.02218162	0.01415950	-0.02304830	0.03319186
-0.00691311	0.00251695	-0.07359995	0.02348454
- 0.00080969	0.00020001	-0.01883383	0.00422792
c		d	
H_0	H_1	H_0	H_1
0.04095616	0.01495479	0.18430940	0.10430306
0.21295453	0.09754126	0.53853041	0.40558545
0.42509906	0.26405461	0.41251639	0.51532830
0.37209430	0.37342961	-0.07678751	0.13439669
0.07046279	0.27343750	-0.10938868	-0.16366452
-0.04396930	-0.04530460	0.00941692	0.05310331
-0.04396930	-0.04530460	0.00941692	0.05310331
0.01192523	-0.03504126	-0.01515299	0.00659616
0.00745127	-0.00714229	0.00314598	-0.00907015

Table 5.1: *Coefficients of the power complementary filters whose magnitude responses are plotted in Figure 5.3. The high-pass filter, $H_1(z)$, has a zero of multiplicity 2 at $z = 1$. a) Maximally flat filters. b) Low-pass filter, $H_0(z)$, with $N_0 = 6$ zeros at $z = -1$. c) Low-pass filter, $H_0(z)$, with $N_0 = 5$ zeros at $z = -1$. d) Low-pass filter, $H_0(z)$, with $N_0 = 4$ zeros at $z = -1$.*

form $F_r = \mathbf{W}f_r$ of a signal f_r in the reconstruction set, which is then itself obtained using the inverse wavelet transform. The closure of the image of the reconstruction set of an f, $\Phi_e^c(F)$[3], under the wavelet transform can be represented as the intersection

$$\Phi_e^c(F) = \mathcal{V} \bigcap \mathcal{E} \bigcap \left(\bigcap_{i,j} \mathcal{C}_{i,j} \right) \tag{5.37}$$

of the following sets:

- $\mathcal{V}$ - the range of the wavelet transform

$$\mathcal{V} = \{G : G = \mathbf{W}g \text{ for some } g \in \ell^2(\mathbf{Z})\}; \tag{5.38}$$

- $\mathcal{E}$ - the set of all $G \in \ell^2(\mathbf{I})$, such that $G^i(k) = F^i(k)$ for all k which are local extrema of F^i, across all the scales $i = 1, 2, ..., J+1$;

- $\mathcal{C}_{i,j}$ - the set determined by the requirement that the component G^i of $G \in \Phi_e^c(F)$ has to be nonincreasing/nondecreasing at the point j if F^i is decreasing/increasing at the same point. Note that the sets $\mathcal{C}_{i,j}$ are associated only to those points where F is strictly increasing or decreasing, i.e. only for those indices (i,j) such that $F^i(j) \neq F^i(j+1)$.

Obviously, $\mathcal{V}$ is a subspace of $\ell^2(\mathbf{I})$ and $\mathcal{E}$ and $\mathcal{C}_{i,j}$'s are closed convex sets, therefore alternating projections [55] of any initial point $F_0 \in \ell^2(\mathbf{I})$ onto $\mathcal{V}$, $\mathcal{E}$ and all the $\mathcal{C}_{i,j}$'s will converge to a point in their intersection, $\Phi_e^c(F)$.

The projection $G_{\mathcal{E}}$ of some G onto $\mathcal{E}$ is obtained by assigning extrema values of F to the corresponding points of G (see Figure 5.5):

$$G_{\mathcal{E}}^i(k) = \begin{cases} F^i(k), & k \text{ is an extremum of } F^i, \\ G^i(k), & \text{otherwise} \end{cases} . \tag{5.39}$$

The projection, $G_{\mathcal{C}_{i,j}}$, of some $G \in \ell^2(\mathbf{I})$ onto $\mathcal{C}_{i,j}$ is equal to G, except possibly at the points j and $j+1$ of G^i, if the monotonicity condition imposed by the set $\mathcal{C}_{i,j}$ is

[3]Recall that F denotes $\mathbf{W}f$.

violated. In that case

$$G^i_{\mathcal{C}_{i,j}}(k) = \begin{cases} \frac{1}{2} \cdot (G^i(j) + G^i(j+1)) & k = j, j+1 \\ G^i(k) & \text{otherwise} \end{cases}, \tag{5.40}$$

as illustrated by Figure 5.5c.

Hence, finding successive projections of some G onto $\mathcal{E}$ and all $\mathcal{C}_{i,j}$'s consists of assigning the arithmetic mean to the pairs of points of G which do not obey the required monotonicity, and assigning the local extrema values of F at the corresponding points. This requires $O(JN)$ additions and $O(JN)$ divisions by two for a length N signal. If the conditions for an FIR implementation of $P_{\mathcal{V}}$ are met, the numerical complexity of the $P_{\mathcal{V}}$ operator is $O(JLN)$ additions and $O(JLN)$ multiplications, where L is the filters' length.

In the wavelet zero-crossings case, the image of the reconstruction set of f under the $\mathbf{W}$ operator is the intersection

$$\Phi^c_z(F) = \mathcal{V} \bigcap \mathcal{U} \bigcap \left(\bigcap_{i,j} \mathcal{Z}_{i,j} \right), \tag{5.41}$$

where $\mathcal{U}$ and $\mathcal{Z}_{i,j}$ are defined as the following:

- $\mathcal{U}$ - the set of all sequences $G \in \ell^2(\mathbf{I})$ such that for all scales $i = 1, ..., J+1$, F^i and G^i have the same integral values between any two adjacent zero-crossings of F^i. If z^i_k denotes the kth zero-crossing of F^i, then $\mathcal{U}$ can be written as:

$$\mathcal{U} = \left\{ \begin{array}{l} G : G \in \ell^2(\mathbf{I}), \sum_{j=z^i_{k-1}}^{z^i_k - 1} G^i(j) = \sum_{j=z^i_{k-1}}^{z^i_k - 1} F^i(j) \\ k = 1, 2, ..., |ZF^i| + 1, \; i = 1, 2, ..., J+1 \end{array} \right\}; \tag{5.42}$$

- $\mathcal{Z}_{i,j}$ - the set of all sequences $G \in \ell^2(\mathbf{I})$ such that G^i has the same sign as F^i at point j. The $\mathcal{Z}_{i,j}$ sets are defined only for the nonzero points of F, only for those indices (i,j) satisfying $F^i(j) \neq 0$.

Since the sets $\mathcal{Z}_{i,j}$ and $\mathcal{U}$ are also closed and convex, a point in the reconstruction set $\Phi^c_z(F)$ can be reached as the limit of a sequence of alternating projections of an

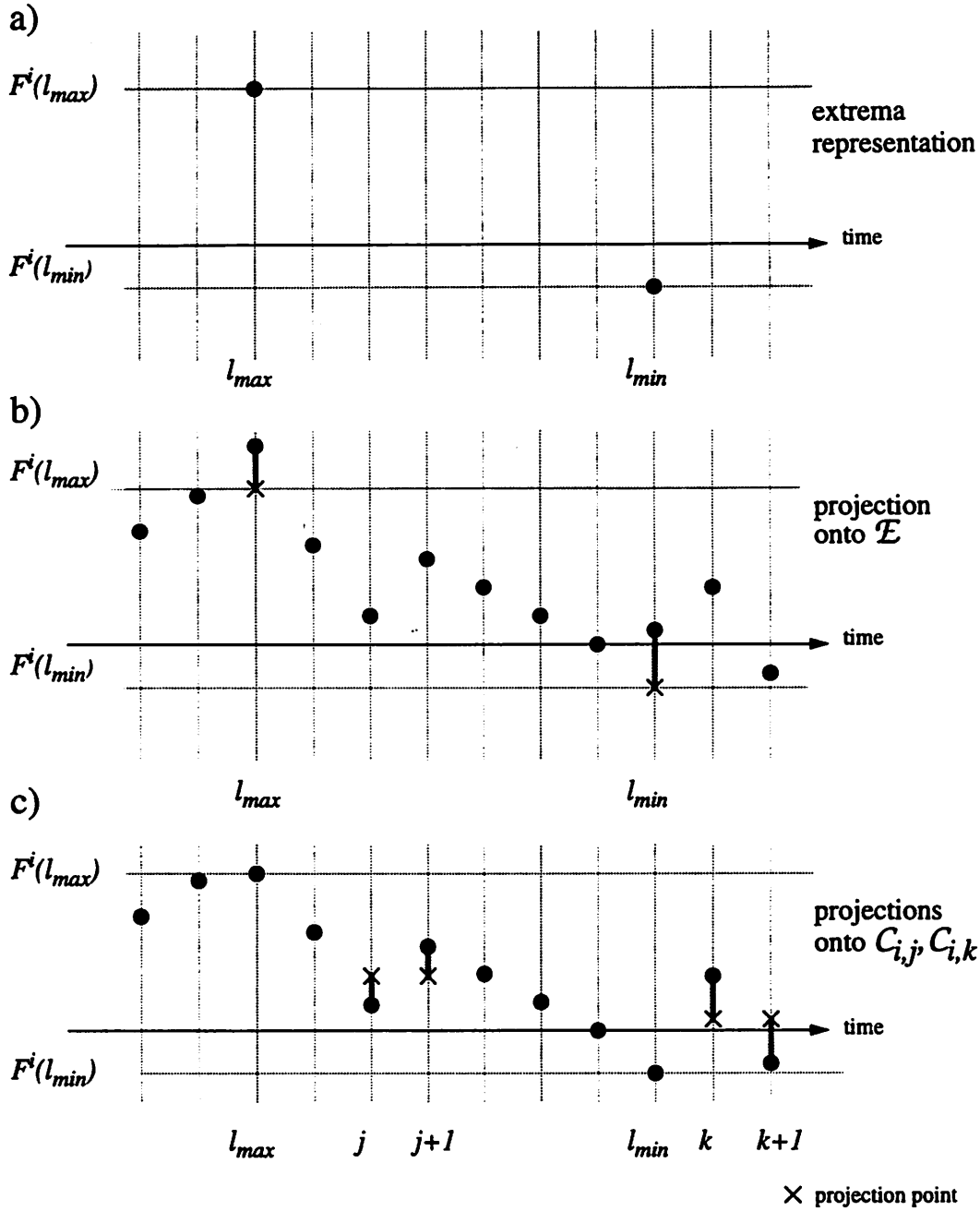

Figure 5.5: *Reconstruction from the wavelet extrema representation: implementation of projection operators. a) A segment of the extrema representation of the sequence F^i, with the local maximum and minimum occurring at the points l_{max} and l_{min}, respectively. b) A segment of the signal G^i (bold dots) and its projection onto $\mathcal{E}$ is obtained by assigning to it values of F^i at the points which are local extrema of F^i (crosses represent the new values of the altered points). c) Projection of G^i onto $\mathcal{E}$ is now represented by the bold dots. It is increasing at the point j, and therefore is not in $C_{i,j}$. The projection is obtained by assigning to the points j and $j+1$ their arithmetic mean.*

arbitrary starting point $F_0 \in \ell^2(\mathbf{I})$ onto $\mathcal{V}$, $\mathcal{U}$ and all $\mathcal{Z}_{i,j}$'s. The projection $G_{\mathcal{Z}_{i,j}}$ of any vector $G \in \ell^2(\mathbf{I})$ onto $\mathcal{Z}_{i,j}$, is obtained by assigning a zero value to the point j of G^i if F^i and G^i don't have the same sign at that point, as shown in Figure 5.6a and Figure 5.6b. The projection operator onto $\mathcal{U}$ operates on some G in the following way. For each sequence G^i, we consider every interval between consecutive zero-crossings of the corresponding F^i. To each point of an interval we add the average difference of G^i and F^i on that interval (see Figure 5.6c). Thus the projection $G_{\mathcal{U}}$ of a G onto $\mathcal{U}$ is given by

$$G^i_{\mathcal{U}}(k) = G^i(k) + \frac{1}{z^i_n - z^i_{n-1}} \sum_{j=z^i_{n-1}}^{z^i_n - 1} (F(j) - G(j)), \tag{5.43}$$

$$z^i_{n-1} \leq k < z^i_n,\ n = 1, ..., |\mathcal{Z}F^i| + 1,\ i = 1, ..., J + 1.$$

Therefore, $G_{\mathcal{U}}$ has the same integral values as F on each interval between consecutive zero-crossing points of F.

This algorithm actually iterates alternatingly between the operator $P_{\mathcal{V}}$, the projection operator onto $\mathcal{U}$, $P_{\mathcal{U}}$, and the projection operator onto $\mathcal{Z}$, $P_{\mathcal{Z}}$, where $\mathcal{Z}$ is the set of all vectors in $\ell^2(\mathbf{I})$ which have the prespecified zero-crossings. A similar algorithm for consistent reconstruction from wavelet transform zero-crossings is proposed by Mallat [25]. Mallat's algorithm iterates between $P_{\mathcal{V}}$ and the projection operator onto $\Gamma = \mathcal{U} \bigcap \mathcal{Z}$, P_Γ. Numerical complexity of P_Γ is $O(JNlogN)$ additions and $O(JN)$ divisions by integers. On the other hand the composition of $P_{\mathcal{Z}}$ and $P_{\mathcal{U}}$ requires $O(JN)$ additions and $O(JN)$ divisions by integers and reduced number of loops with respect to P_Γ. It may appear that the algorithm described here has a slower convergence since the constraints in Γ are split between $\mathcal{U}$ and $\mathcal{Z}$, and successive projections onto $\mathcal{U}$ and $\mathcal{Z}$ in general do not yield projection onto Γ. However, the situation is not so simple. It is possible to come up with examples where such splitting strategy can even improve speed of convergence or give exactly the same results at a reduced cost of implementation. For instance, if we start the reconstruction with a point inside Γ and the reconstructed signals stayed in $\mathcal{Z}$ throughout the reconstruction procedure, the two algorithms would give the same result. Experiments showed that during the

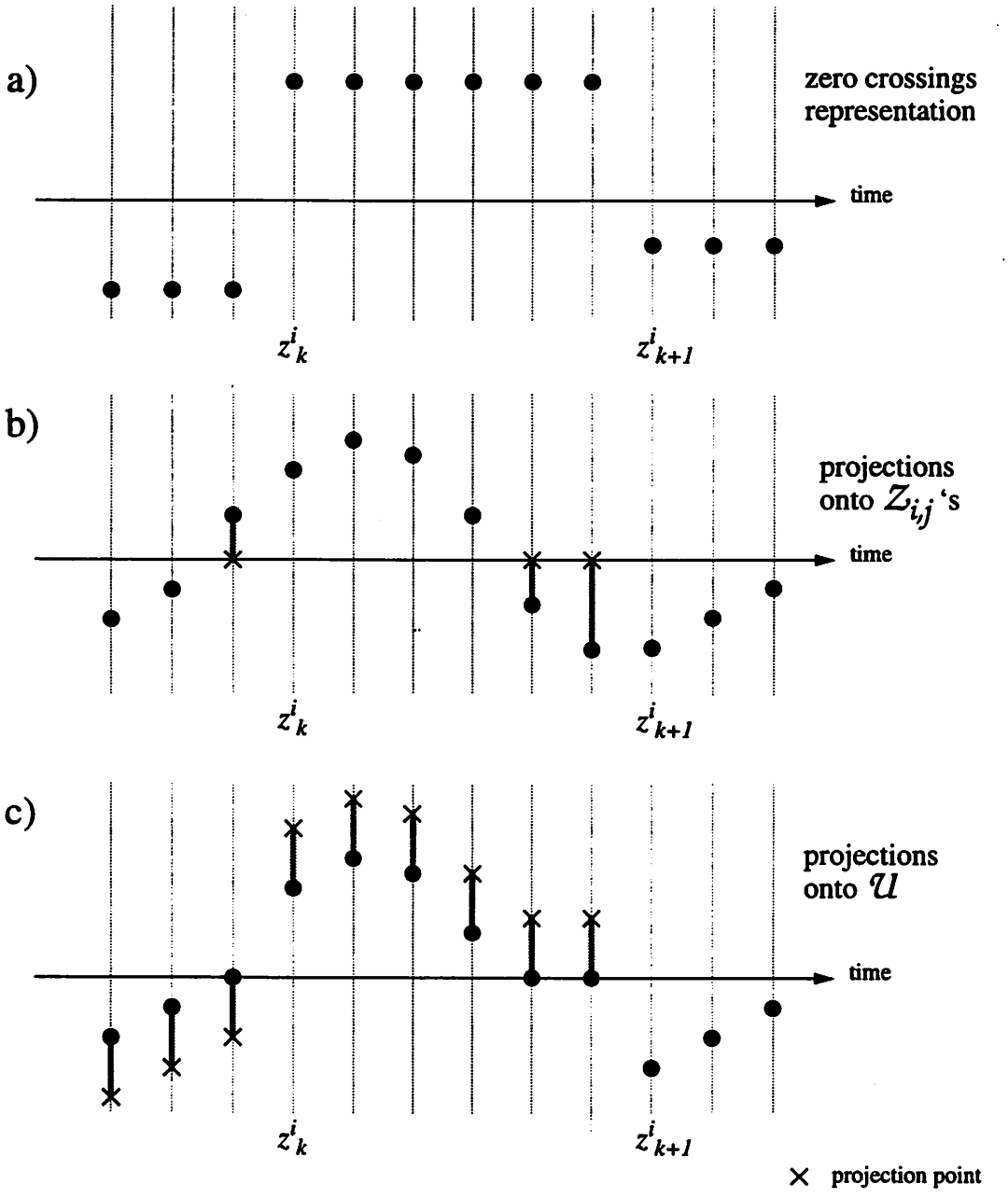

Figure 5.6: *Reconstruction from the wavelet zero-crossings representation: implementation of the projection operators. a) A segment of the zero-crossings representation of F^i with zero crossings occurring at the points z^i_k and z^i_{k+1}; b) A segment of the sequence G^i, represented by the bold dots. Its projection onto each $\mathcal{Z}_{i,j}$ is obtained by assigning zero values to those points which do not have the required sign. c) The projection of G^i onto $\mathcal{Z}_{i,j}$ is represented by the dots. Its projection onto $\mathcal{U}$ is found by adding the same value to each point of a segment between zero crossings of F^i, so that the required integral values are achieved.*

reconstruction process all intermediate solutions are very close to $\mathcal{Z}$, if not inside it, and hence that the composition of operators $P_{\mathcal{U}}$ and $P_{\mathcal{Z}}$ yields points close to those obtained by the $P_{\mathcal{T}}$ operator. In most cases we observed that the reconstructed signals obtained by Mallat's and the new algorithm, starting at a same point, were very close and differed in the reconstruction error typically by less than $0.1dB$ in each iteration. It is important to note that with each iteration of the reconstruction algorithms the distance between the original signal and its estimate is decreased, which is a consequence of the fact that projectors onto convex sets are non-expansive operators. For alternative algorithms for consistent reconstruction, based on the gradient descent algorithm, the reader is referred to work by Berman et al. [4].

Extension of the wavelet modulus maxima representation framework to two-dimensional signals and its applications in image processing are considered by Mallat, *et al.* in [26, 28]. We introduce wavelet zero-crossings representation of images as a multiscale edge representation. An advantage of the wavelet zero-crossings representation over the wavelet modulus maxima representation is its convexity which makes consistent reconstruction simple.

The discrete wavelet transform operator of two-dimensional signals for wavelet modulus maxima [26, 28] and wavelet zero-crossings representations is the linear operator $\mathbf{W} : \ell^2(\mathbf{Z}^2) \to \ell^2(\{1,2,\ldots,2J+1\} \times \mathbf{Z}^2)$ consisting of $2J+1$ linear operators $W_{i,j} : \ell^2(\mathbf{Z}^2) \to \ell^2(\mathbf{Z}^2)$ $i = 1,2$ $j = 1,2,\ldots,J$ and $W_{J+1} : \ell^2(\mathbf{Z}^2) \to \ell^2(\mathbf{Z}^2)$. The operators $W_{1,j}$, $W_{2,j}$ and W_{J+1} denote, respectively, separable filtering with the filters

$$\begin{aligned} V_{1,j}(z_x, z_y) &= H_0(z_x)H_0(z_y)\cdots H_0(z_x^{2^{j-1}})H_0(z_y^{2^{j-1}})H_1(z_x^{2^j}), \\ V_{2,j}(z_x, z_y) &= H_0(z_x)H_0(z_y)\cdots H_0(z_x^{2^{j-1}})H_0(z_y^{2^{j-1}})H_1(z_y^{2^j}), \\ V_{J+1}(z_x, z_y) &= H_0(z_x)H_0(z_y)\cdots H_0(z_x^{2^J})H_0(z_y^{2^J}). \end{aligned} \tag{5.44}$$

Defined this way, the wavelet transform operator $\mathbf{W}$ can be implemented using the filter bank based on the prototype filters $H_0(z)$ and $H_1(z)$ as shown in Figure 5.7a. Perfect reconstruction is possible provided that there exist two filters $G_0(z)$ and $G_1(z)$ satisfying (5.27). An inverse operator $\mathbf{W}^{-1}$ implemented again as a filter bank, based on the filters $G_0(z)$, $G_1(z)$ and $L(z) = \frac{1}{2}(1 + H_0(z)G_0(z))$, is illustrated by Figure 5.7b. For the analysis filters of which $H_0(z)$ is low-pass and $H_1(z)$ has

exactly two zeros at $z = 1$, zero-crossings of $W_{1,j}f$ and $W_{2,j}f$ are related to the sharp variations of f along x and y coordinates, respectively. This explains the reason for using this particular type of wavelet transform. For the details on this issue the reader is referred to the work by Mallat and Zhong [28].

The wavelet zero-crossings representation of a two-dimensional signal will be defined again using two operators, denoted as in the one-dimensional case by Z and S, with similar meaning. The zero-crossing operator Z in $\ell^2(\mathbf{Z}^2)$ is defined as

$$Zf = \{(k,l) : f(k,l)f(k-1,l) \leq 0 \text{ or } f(k,l)f(k,l-1) \leq 0\}. \tag{5.45}$$

Zero-crossings of a two-dimensional signal define a number of connected areas of points sharing the same sign, which will henceforth be referred to simply as *areas*. The S operator provides information on integral values (the sum of points) of the signal in each of these areas:

$$Sf = \left\{ \begin{array}{c} Sf(k): \text{ the sum of points inside the area } k, \\ k = 1,2,\ldots, \text{ number of areas} \end{array} \right\}. \tag{5.46}$$

Note that according to the definition (5.45) of zero-crossings, all the points of f, where f assumes zero value, are declared as zero-crossings. In implementation these points can be associated with any of the contiguous areas. Analogous to the definition (5.20) in $\ell^2(\mathbf{Z})$, the wavelet zero-crossings representation for two-dimensional signals is defined as

$$E_z f = \{ZW_{i,j}f, SW_{i,j}f, \;\; i = 1,2 \; j = 1,\ldots,J+1, \; ZW_{J+1}f, SW_{J+1}f\}. \tag{5.47}$$

The reconstruction algorithm from this wavelet zero-crossings representation of a two-dimensional signal, is a straightforward extension of the reconstruction algorithm in the case of one-dimensional signals, and details are not given here. However, it can be shown that for the wavelet transform of a two-dimensional signal as defined in (5.44), the operator for orthogonal projection onto the range of the wavelet transform can not have an FIR implementation. For experimentation we use a $\mathbf{WW}^{-1}$ which has an FIR implementation. The FIR synthesis filters $G_0(z)$ and $G_1(z)$ are not

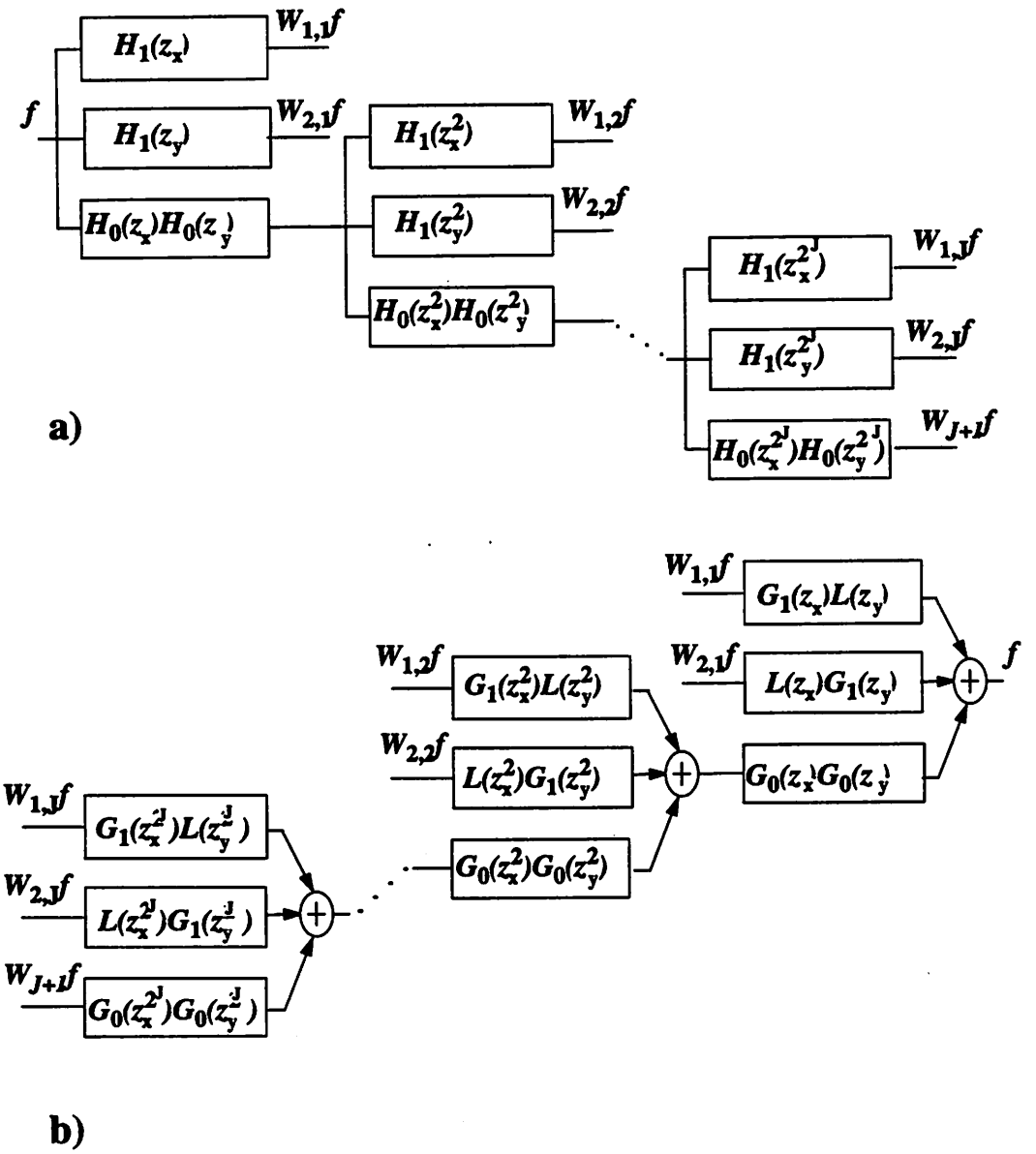

Figure 5.7: *The wavelet transform used in the two-dimensional wavelet zero-crossings representation. a) Filter bank implementation of the two-dimensional wavelet transform operator* $\mathbf{W}$*; b) Filter bank implementation of an inverse of the two dimensional wavelet transform operator.*

unique and should be designed with caution, because some choices may even lead to divergence of the reconstruction algorithm. For instance, a pair of FIR synthesis filters $G_0(z)$ and $G_1(z)$ which satisfy the perfect reconstruction condition can be obtained from $H_0(z)$ and $H_1(z)$ using Euclid's algorithm which gives a high-pass filter $G_0(z)$ and a low-pass filter $G_1(z)$ for $H_0(z)$ and $H_1(z)$ which are low-pass and high-pass, respectively. Such synthesis filters are certainly a bad choices and in this case we observe divergence of the reconstruction algorithm. The design of synthesis filters which would ensure convergence of the reconstruction algorithm in the case when $\mathbf{W}^{-1}$ is not implemented as the Hilbert adjoint of the dual of $\mathbf{W}$ is still an open problem. In experiments with power complementary filters $H_0(z)$ and $H_1(z)$ the reconstruction algorithm always converges for $G_0(z) = H_0(z^{-1})$ and $G_1(z) = H_1(z^{-1})$. Also, we observe the convergence and very good reconstruction results for $G_0(z)$ and $G_1(z)$ whose magnitude responses are close to those of $H_0(z)$ and $H_1(z)$, respectively. The numerical complexity of the $\mathbf{W}\mathbf{W}^{-1}$ operator is $O(JLN^2)$ additions and $O(JLN^2)$ multiplications if it is implemented using FIR filters of length L. The rest of the operators used in the reconstruction are analogous to their counter-parts in the one-dimensional case and their numerical complexity is $O(JN^2)$ additions and $O(JN^2)$ divisions by integers.

5.4 Experimental Results

Generally, signals cannot be reconstructed with arbitrary high quality from their wavelet extrema or zero-crossings representation since the representations are non-unique. It can be shown that for a finite length signal, the closures of the reconstruction set is the convex hull of finitely many vertices [4]. The size of the reconstruction set, determined by the distances between these vertices, directly influences the quality of the reconstructed signal. Each wavelet transform extremum or zero-crossing represents a linear constraint which defines a hyperplane in the signal space bounding the reconstruction set. It can be expected that signals producing more extrema (zero-crossings) have reconstruction sets of smaller sizes, and consequently yield bet-

ter reconstruction results. However, the price paid is the larger number of points in the representation. Experiments generally confirm this rather heuristic argument. Signals with higher frequency content, which results in larger number of extrema or zero-crossings usually have faster convergence and better reconstruction.

The regularity of the wavelet used for the representation plays an important role. Recall that the dyadic wavelet transform is a sequence of signals $f_{2^j} = \psi_{2^j} * f_c$, $j = 1, 2, \ldots$, which are results of filtering a signal $f_c \in L^2(\mathbf{R})$ by dilated versions of the wavelet ψ. Obviously, in generating signals f_{2^j}, the signal f_c and the dilated wavelet ψ_{2^j} play interchangeable roles. Therefore, the extrema or zero-crossings of the dyadic wavelet transform can result from sharp transitions of either the signal or the wavelet. If the aim is to suppress those local extrema (zero-crossings) which are wavelet rather than signal related, it is advisable to use smoother wavelets. However, more regular wavelets generally produce reduced number of wavelet transform extrema (zero-crossings), and consequently yield poorer reconstruction results. As an illustration of the above discussion, Figure 5.8a represents results from reconstructing randomly generated signals from the wavelet transform zero-crossings for wavelets with different regularity properties. Experimental results of the reconstruction from wavelet extrema representation are represented in Figure 5.8b. For a comparison with the zero-crossings reconstruction, one of the curves from Figure 5.8a is plotted again on the same graph. Note that in both cases shown in Figure 5.8b, each curve represents an average for the same set of random signals, and that the total number of the zero-crossings was on average around 6.5% smaller than the number of extrema. In the experiments with wavelet transform extrema representation, we used the same filters as Mallat, *et al.* for their wavelet modulus maxima scheme [28].

Examples of images reconstructed from the wavelet zero-crossings representation are shown in Figure 5.9. The size of the originals (the left column) is 256×256 pixels, and the reconstructed images (the right column) are obtained after 10 iterations of the algorithm. In these experiments the wavelet transform is performed across four scales and the representation includes the entire signal $W_{J+1}f$ (there is no significant difference in the reconstruction error with respect to the case when only the zero-

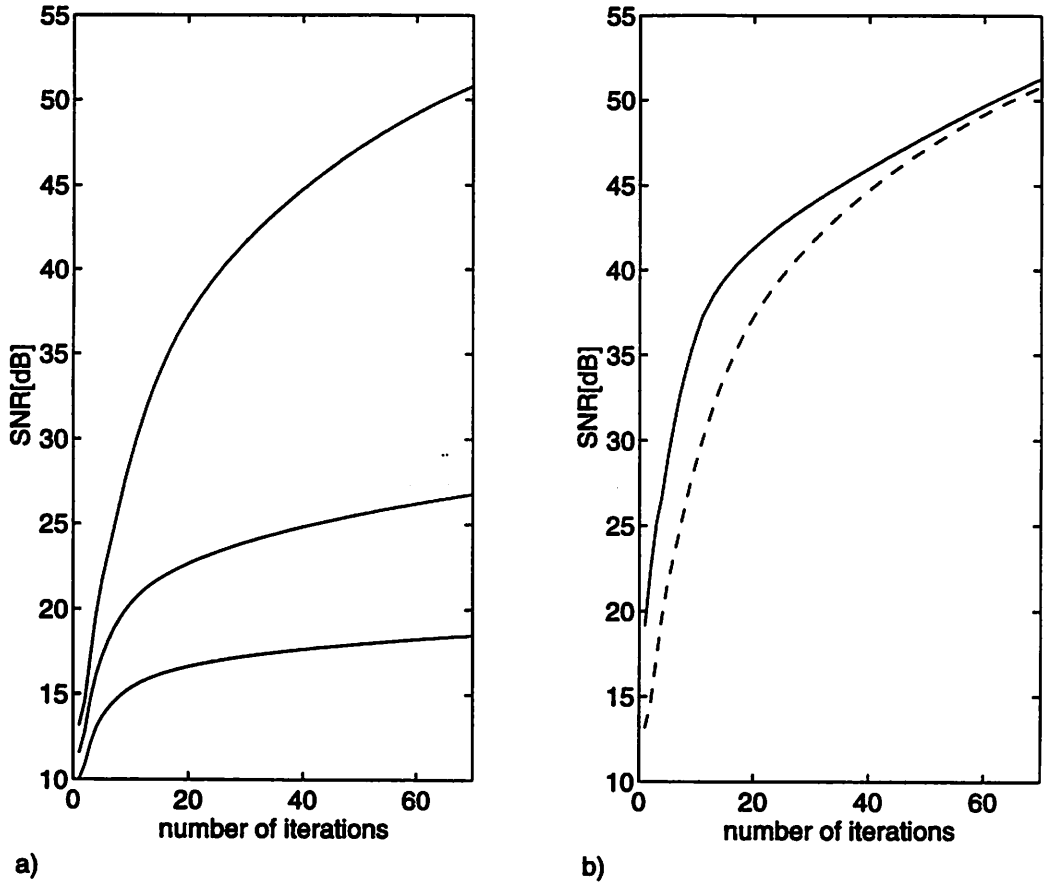

Figure 5.8: *Results from reconstruction of randomly generated one-dimensional signals. a) SNR of the reconstruction from the wavelet zero-crossings representation for wavelets with different regularities. Curves on the top, middle and bottom correspond to wavelets on figures 5.3c, 5.3b and 5.3a respectively. b) SNR of the reconstruction from the wavelet extrema representation (solid line), and SNR of the reconstruction from the wavelet zero-crossings representation for the wavelet in figure 5.3c (dashed line).*

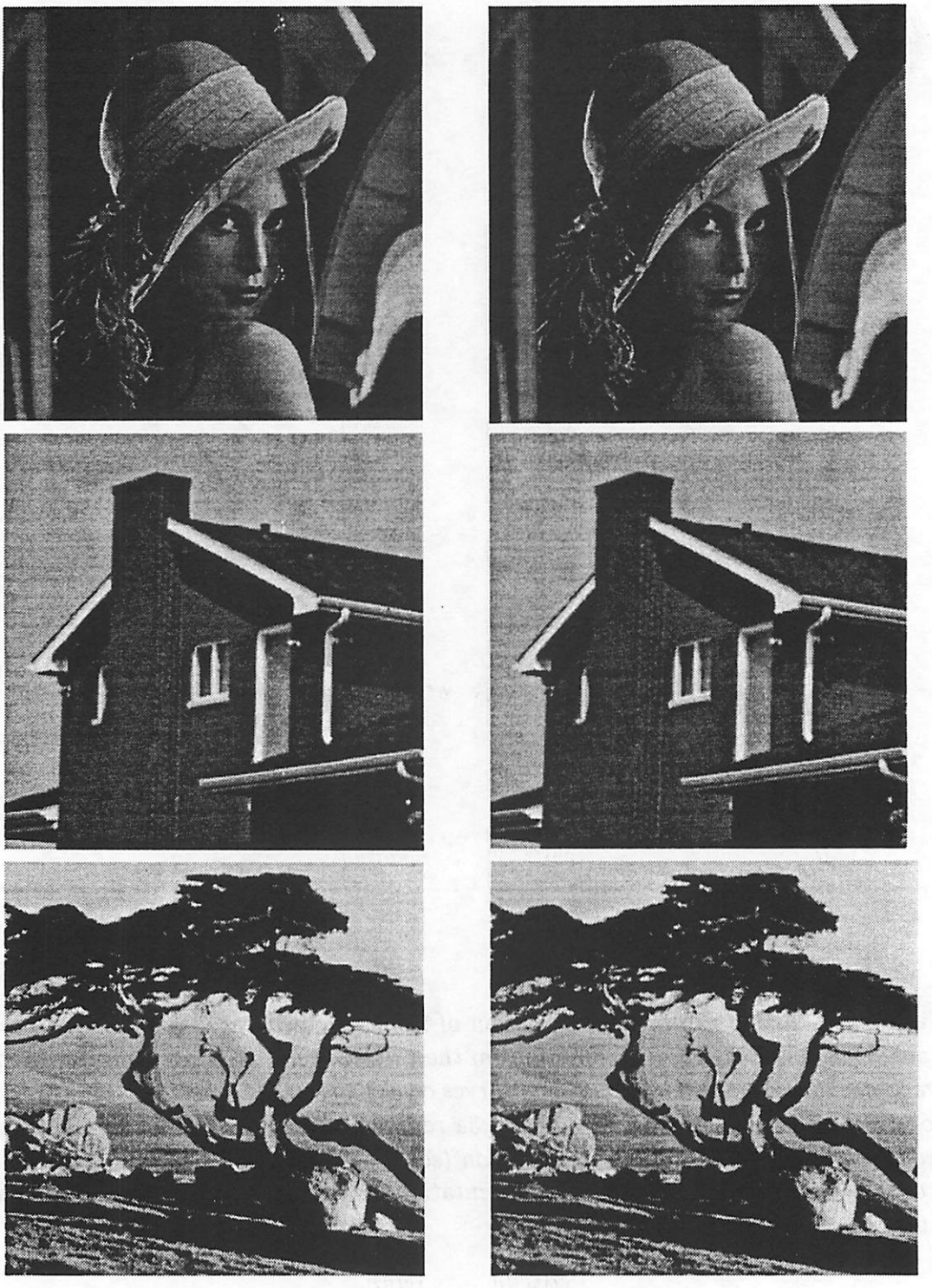

Figure 5.9: *Examples of images reconstructed from the wavelet zero-crossings representation. Left column: the originals,* 256×256 *pixels. Right column: reconstructed images, obtained after 10 iterations of the algorithm. The SNR's are 36.1dB, 40.3dB and 33.6dB for "Lenna", "House" and "Tree" images, respectively.*

Figure 5.10: *The "Lenna" image reconstructed from partial wavelet zero-crossings representation. Bottom left:* 256×256 *original. Bottom right: the image obtained after 10 iterations of the reconstruction algorithm, with 20.8dB SNR (PSNR is 28.0dB). Bilevel images: black regions mark selected areas across four scales of the wavelet transform; the number of selected areas is 5146. Top left, top right, middle left and middle right images represent scales 1, 2, 3 and 4, respectively.*

crossings information on $W_{J+1}f$ is used). As in the case of one-dimensional signals we observe that using more regular filters reduces the amount of data to be recorded but increases the mean-squared reconstruction error. In all experiments with images, we use linear phase filters: $H_0 = \frac{1}{4}[1 \quad 2 \quad 1]$, $H_1 = \frac{1}{4}[1 \quad -2 \quad 1]$, $G_0 = \frac{1}{8}[-1 \quad 2 \quad 6 \quad 2 \quad -1]$ and $G_1 = \frac{1}{8}[1 \quad 2 \quad -6 \quad 2 \quad 1]$.

Some modified versions of the original wavelet zero-crossings scheme can be more convenient. The definition of the two-dimensional wavelet zero-crossings representation, as introduced in the preceding section, can be generalized in the following manner. Some of the areas of the same sign can be partitioned into several subareas, and the information on these subareas, namely the locations and integral values, is extracted separately. This increases the overhead, but generally decreases distortion of the representation and may facilitate some signal processing tasks, such as the selection of important edges. On the other hand, only partial information on the wavelet transform zero-crossings can be kept, which has the opposite effect on rate-distortion properties of the scheme. The reconstruction algorithm can be modified in a straightforward manner to accommodate these variants of the representation.

Figure 5.10 illustrates the reconstruction of an image from the representation obtained by combining the two modifications. Black regions in the two-level images represent the subareas which are included in the representation, across four scales of the wavelet transform. The selected subareas are those with positive integral values and average intensity above a given threshold. The original (bottom left) is the 256×256 image, and the reconstructed signal (bottom right) is obtained after 10 iterations of the algorithm, with 20.8dB SNR. The number of subareas used in the representation is 5146. It can be noticed that even with this rather naive selection process, wavelet zero-crossings representation provides information on important multiscale edges and yields good reconstruction.

Applications of the wavelet zero-crossings or wavelet modulus maxima schemes to real coding systems require further research. One of the main issues is to find efficient algorithms for selecting the most relevant information contained in modulus maxima or zero-crossings and also for exploiting their interdependencies across the

scales for further lossless encoding of the selected information. An advantage of the wavelet modulus maxima over the zero-crossings scheme is that it discriminates between different types of singularities and therefore facilitates the selection process. On the other hand, extracting sign information, which is required for the wavelet zero-crossings representation, seems to be computationally less intensive than extracting information on local modulus maxima as proposed in [28]. Also, as we pointed out earlier, the convexity of the zero-crossings representation makes the reconstruction from partial information simple and more reliable. Compression results reported by Mallat and Zhong [28], based on the wavelet modulus maxima representation, were encouraging and we believe that the full potential of these schemes is still to be reached.

5.5 Image Interpolation

5.5.1 Traditional Linear Interpolating Schemes

Digital images obtained from data acquisition systems, such as medical or satellite images, are often required to be magnified for subsequent analysis which involves a human observer. The magnification problem amounts to estimating unknown pixel amplitudes from their known neighbors. A straightforward approach would be to assign an unknown pixel the amplitude of one of its neighbors, a scheme denoted as *pixel replication*. A problem with pixel replication are the jaggy line artifacts (see Figure 5.11). These artifacts can be alleviated if the estimation is based on local averages. A commonly used technique is the bilinear interpolation, which consists of linear interpolating between the known pixels along each row or column, followed by linear interpolating in the perpendicular direction. Various other interpolation functions based on different local averages are also used, such as bell and bicubic spline interpolation [20]. These methods though eliminating the artifacts tend to cause blurring of images, so that they are inevitably linked with the tradeoff between the interpolation error (artifacts) and the clarity degradation, as illustrated in Figure

5.11. The image clarity can subsequently be enhanced using a high-pass filtering. This is commonly used in the printing industry for edge crispening as the so called *unsharp masking* technique [20]. Examples of interpolated images enhanced using the unsharp masking are also illustrated in Figure 5.11.

These local averaging techniques are linear schemes which consist of linear filtering of an image which has been previously interleaved with zeros. The filter used is a cascade of a low-pass filter (averaging) and a high-pass filter (unsharp masking). The linearity implies that the same interpolating function is applied across the image regardless of the local characteristics. The interpolation algorithm we propose is a nonlinear locally adaptive scheme. It is based on the wavelet analysis of local behavior in the image in order to apply appropriate interpolation. Other locally adaptive methods include modeling of edges, nonlinear filtering for amplification of the high frequency components [17] or MAP estimation [39].

5.5.2 Locally Adaptive Image Interpolation

The perceptually most important image features are edges and textures. However, they are of different nature and require different approaches when interpolated. The algorithm described here is aimed at good rendition of edges only. For the sake of simplicity the idea is explained in more detail for one-dimensional signals. Images are treated as a separable extension of the one-dimensional case. Figure 5.12 illustrates the problem model we use. The available waveform f is considered to be obtained from the high resolution signal f_0, which we want to recover, by lowpass filtering followed by downsampling by 2. Let $H_0(z)$ be a lowpass filter and $H_1(z)$ be a highpass filter such that the two filters constitute a two channel perfect reconstruction filter bank. Let $G_0(z)$ and $G_1(z)$ be the corresponding synthesis filters. The filter bank in this model is arbitrary, but we conjecture that as long as it is reasonable (i.e. a good lowpass/highpass pair of filters), performance of the algorithm will not depend strongly on the filter bank. In order to perfectly restore the high resolution signal, we need to know both its highpass component g_s and its lowpass component f_s. However, only f, the downsampled version of f_s, is available. A standard approach

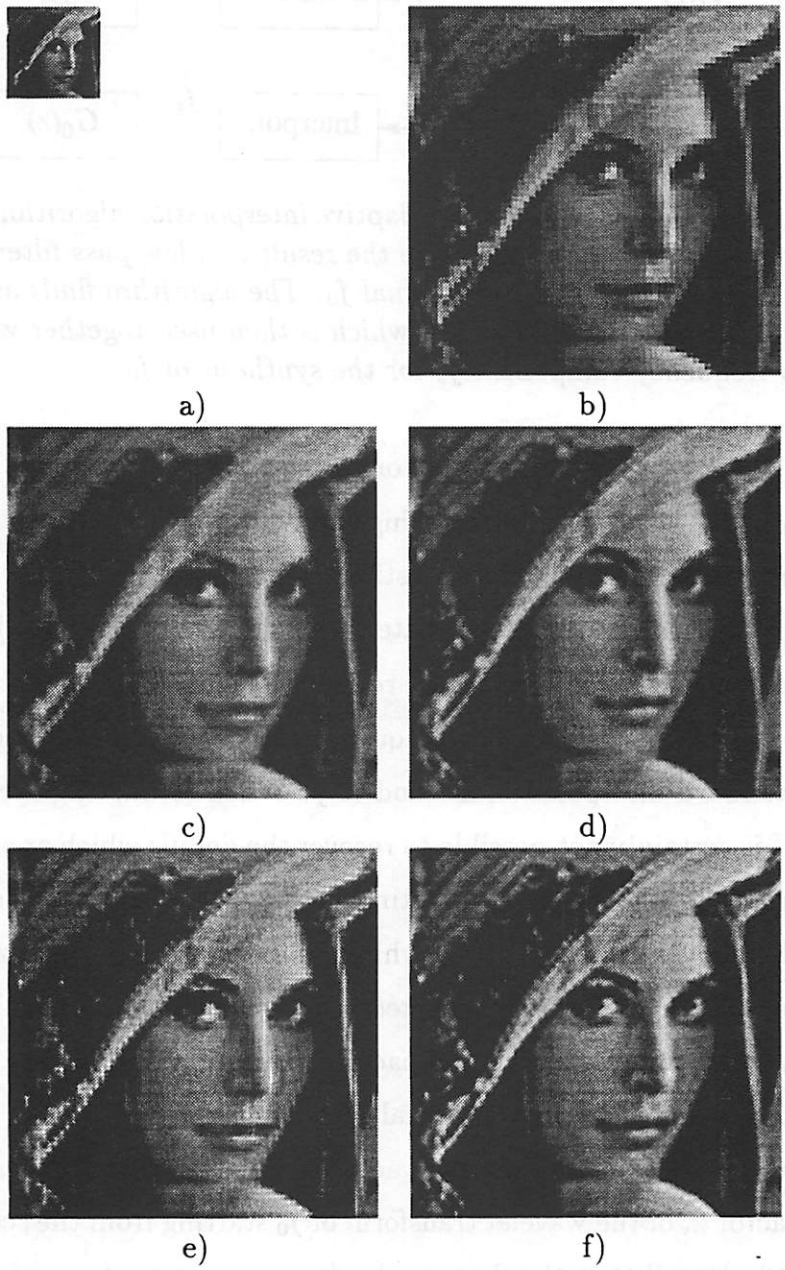

Figure 5.11: *Examples of an image magnified 4 times along each of the coordinates using linear interpolation schemes. a) Original low resolution image, 64 × 64 pixels. b) Pixel replication. c) Bilinear interpolation. d) Bicubic spline interpolation. e) Bilinear interpolation followed by unsharp masking. f) Bicubic spline interpolation followed by unsharp masking.*

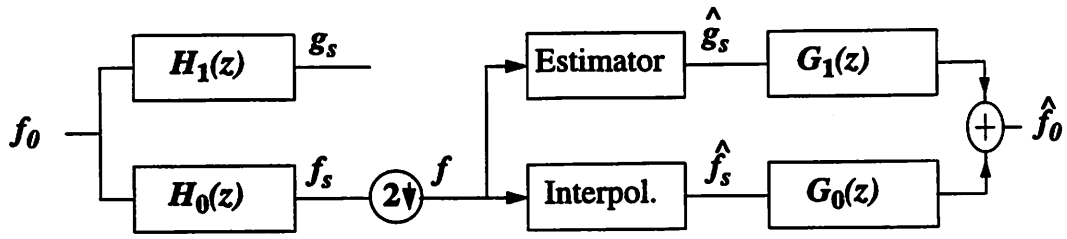

Figure 5.12: *The model for the locally adaptive interpolation algorithm. The available low resolution signal f is assumed to be the result of a low-pass filtering followed by subsampling of the higher resolution signal f_0. The algorithm finds an estimate $\hat{g}_s$ of the high frequency component g_s of f_0 which is then used together with an estimate $\hat{f}_s$ of the low-frequency component f_s for the synthesis of f_0.*

would be to interpolate f using linear or spline interpolation, possibly followed by some enhancement algorithm such as highpass filtering to deblur the result. The algorithm proposed here is based on estimating the high frequency component g_s, which is then combined with an estimate of f_s, through the synthesis filter bank to give a reconstructed version of the high resolution signal.

The initial estimate $\hat{f}_s$ of the low frequency component f_s can be obtained by applying any standard interpolation method to f. When estimating the high frequency component it is certainly not possible to recover the details which are irreversibly lost by the downsampling. However, the estimate of g_s should be based on the information on local smoothness of the signal which can be extracted from its low resolution version. Since this information is reflected in the propagation of local modulus maxima in the wavelet transform, the approach to the estimation of g_s is to find its local extrema by analyzing the available signal f.

It can be shown that the discrete dyadic wavelet transform of f is the decimated version, by factor 2, of the wavelet transform of f_0 starting from the scale 2^2, as shown in Figure 5.13. Recall that the discrete dyadic wavelet transform of f_0 is a sampled version of the continuous wavelet transform at scales 2, 2^2, 2^3, ... The high-pass component g_s is the wavelet transform of the high resolution signal at the scale 2. The idea of our algorithm is to extrapolate the wavelet transform of f_0 at the scale 2 from the subsampled versions of its wavelet transform at scales 2^2, 2^3, ... ,2^J. The extrapolation is based on the propagation of local extrema of the wavelet transform

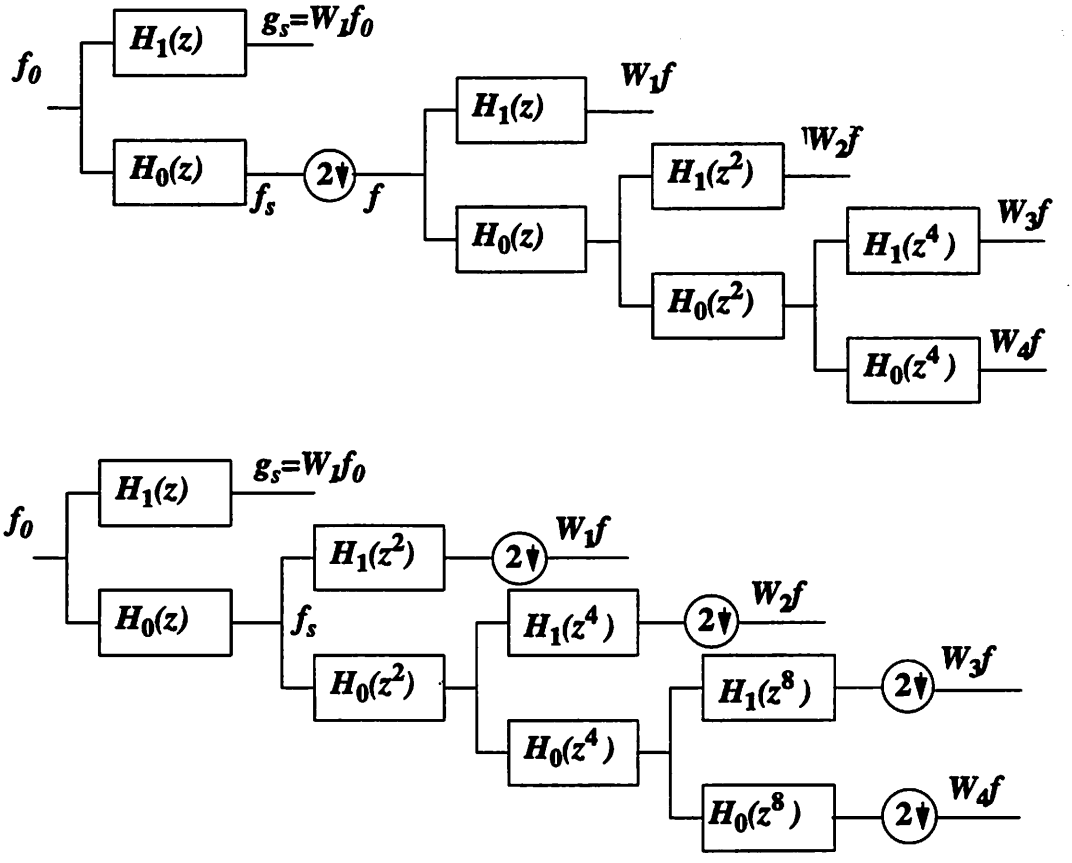

Figure 5.13: *Estimation of the high frequency component g_s amounts to extrapolation of the wavelet transform of f_0 across the scales. The extrapolation is based on the analysis of the wavelet transform of the low-resolution signal f, which is a subsampled version of the wavelet transform of the high resolution signal f_0 at the scales $2^2, 2^3, 2^4$....*

of the low resolution signal f across the scales, following the exponential propagation rule. This enables estimation of the locations and values of important local extrema of g_s. The estimate $\hat{g}_s$ is then obtained by interpolating between these extrema using, for instance, linear or spline interpolation. A higher resolution signal, which is our final goal, is generated from components $\hat{f}_s$ and $\hat{g}_s$ using the synthesis filter bank (see Figure 5.12).

The interpolated signal, obtained this way, can be enhanced by recognizing that the initial estimates $\hat{f}_s$ and $\hat{g}_s$ can be further improved. The enhancement is based on alternating projections on the sets of convex constraints which $\hat{f}_s$ and $\hat{g}_s$ should

obey. We can identify the following constraints:

1. The waveforms $(\hat{f}_s, \hat{g}_s)$ must be in the subspace $\mathcal{V}$ of $l^2(\{1,2\} \times \mathbf{Z})$, which is the space of all pairs of square-summable sequences which can be obtained at the outputs of the filter bank $[H_0(z)\ H_1(z)]$ for an excitation in $\ell^2(\mathbf{Z})$.

2. The decimated version (by factor 2) of $\hat{f}_s$ should be equal to f, which is the available low resolution signal.

3. The local extrema of $\hat{g}_s$ should reflect sharp variations in f_0, i.e. their values and locations are determined by singularities in f_0.

Let $\mathcal{V}$, $\mathcal{S}$, and $\mathcal{E}$ be sets in $l^2(\{1,2\} \times \mathbf{Z})$ denoting, respectively, the sets of points which satisfy these three constraints. The pair of estimates $(\hat{f}_s, \hat{g}_s)$ should belong to the sets $\mathcal{V}$ and $\mathcal{S}$, while projecting it onto $\mathcal{E}$ improves the signal clarity.

The projection operator onto $\mathcal{V}$ is discussed in the previous section. Projection of $(\hat{f}_s, \hat{g}_s)$ onto $\mathcal{S}$ amounts to assigning values of f to even samples of $\hat{f}_s$. The subspace $\mathcal{V}$ and the convex set $\mathcal{S}$ are well-defined, but the set $\mathcal{E}$ depends on our knowledge of the singularities of f_0 which is based on the analysis of the low resolution signal f, so there is a certain arbitrariness in its definition.

One approach would be to define $\mathcal{E}$ as the set of all pairs of waveforms $(\hat{f}_s, \hat{g}_s)$ such that the locations and amplitudes of local extrema of $\hat{g}_s$ match exactly values assigned based on the analysis of f. This is an extreme approach, with a certain controversy. The problem of this approach is that it insists on precisely assigned characteristics of local extrema of g_s which are, on the other hand, known with some uncertainty. There are the two reasons for systematic errors in the measurements of the local extrema. The first is that the discrete dyadic wavelet transform represents a sampled version of the continuous wavelet transform provided that the discrete transform is applied on the appropriate discrete-time version of the continuous-time signal, as pointed out in Section 5.3.2. Although a reasonable low-pass filtering followed by sampling would provide a good approximation of the required discretization, it still produces some error. The other source of error is the limited time resolution of the discrete wavelet transform. This is particularly pronounced at the fine scales of the wavelet

transform where the underlying continuous-time signal can assume a considerable variation between two points of the discrete wavelet transform. Consequently, the values of local modulus maxima based on which we determine parameters of the exponential propagation law, Ks^{α}, and estimate the local extrema of g_s, are not measured accurately. Also, there are always some errors in associating those wavelet extrema across the scales which correspond to the same singularity.

Considering the possible errors in the process of estimating the local extrema of g_s, when enhancing the initial interpolated signal we may abandon any further constraints on local extrema of the estimates $\hat{g}_s$ and project $\{\hat{f}_s, \hat{g}_s\}$ alternatingly onto $\mathcal{V}$ and $\mathcal{S}$ only. That gives the other extreme approach in defining $\mathcal{E}$. As a moderation of the two extreme cases, local extrema of $\hat{g}_s$ can be assigned with certain tolerance in both positions and amplitudes. These constraints can be defined so that $\mathcal{E}$ is still convex. The algorithm for signal reconstruction from the wavelet extrema representation, described in Section 5.3.4, can be easily modified to an algorithm for projection onto $\mathcal{E}$ in the case when the local extrema are constrained to be in the tolerance bounds.

Images are treated using a separable extension of the scheme. In general, analyzing two-dimensional signals by treating the two coordinates independently is not an optimal approach. However, we chose a separable scheme in order to reduce computational complexity. The discrete time wavelet transform of images for the interpolation algorithm is the one used for the wavelet extrema representation. The filter bank which implements this transform is shown in Figure 5.7. The wavelet transform of an image f consists of components $\{W_{1,j}f\}_{j=1,\ldots,J}$ and components $\{W_{2,j}f\}_{j=1,\ldots,J}$ which are generated by processing f, respectively, along rows and along columns. The one-dimensional extrapolation algorithm is then applied to rows of the $W_{1,j}$ components and columns of the $W_{2,j}$ components.

The discussion in this section gives only a global idea of the interpolation algorithm. Readers interested in more detail are referred to [6].

5.5.3 Experimental Results

The experiments confirm the expectation that the moderate approach in constraining local extrema of the high-pass component $\hat{g}_s$ gives better results than the two extreme cases. For the algorithm variant that does not constrain the extrema of $\hat{g}_s$, ringing effects tend to occur around the edges. On the other hand, when constraining the values to be the initial estimates, the resulting image is of lower quality compared to the case when some tolerances are allowed.

The images obtained after 1-2 iterations of the enhancement algorithm are already acceptable. After 5-10 iterations the images are of good quality and with more iterations they do not change discernibly or some artifacts appear, depending on algorithm parameters. A direct magnification of four times was also experimented, but the result was not as good as performing the 2-time magnification algorithm iteratively.

Figure 5.14b shows a 256×256 images obtained by performing the interpolation algorithm twice iteratively. For comparison, interpolated images using the bilinear or bicubic spline interpolation followed by unsharp masking are shown on the same figure. The original low-resolution 64×64 image is obtained from a 256×256 high resolution image by performing twice the process of low-pass filtering and subsampling by 2. The lowpass filter used in obtaining the test image is a separable filter $F(z_x, z_y) = F_1(z_x)F_1(z_y)$, where the impulse response of $F_1(z_x)$ is $[-1, 0, 9, 16, 9, 0, -1]/32$. The filters for the wavelet transform in this experiment were

$H_0 = [0.125, 0.375, 0375, 0.125]$,

$H_1 = [-2.0, 2.0]$,

$G_0 = [0.125, 0.375, 0.375, 0.125]$,

$G_1 = [0.0078125, 0.046875, 0.1171875, 0.65625, 0.1171875, 0.046875, 0.0078125]$.

This algorithm is still subject to experimental investigation aimed at fine tuning of the parameters and reduction of the computational complexity. One issue under investigation is the impact of filter regularity on the performance of the algorithm. Preliminary results indicate that using more regular filters yields more pleasing results. Also, as we have already pointed out, regular filters should provide a clearer

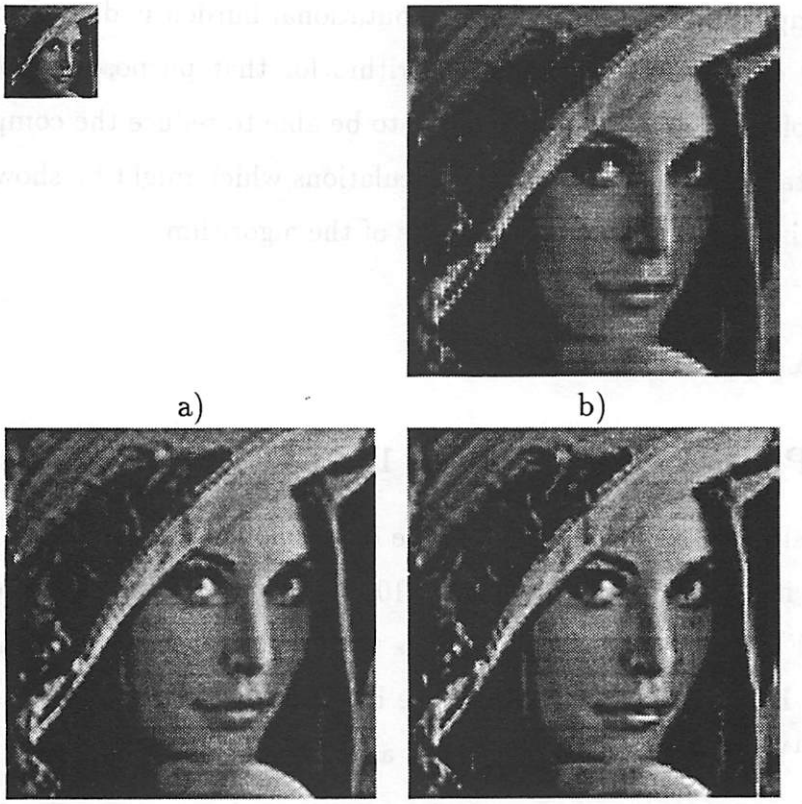

Figure 5.14: *Example of a* 64×64 *image magnified* 4 *times along each of the coordinates using the adaptive algorithm. The available small size image is shown in figure a), and its magnified version in figure d). For a comparison, images magnified using bilinear and bicubic spline interpolation followed by the unsharp masking are also shown in figures b) and c), respectively.*

picture of important local extrema and thus facilitate the estimation of the high frequency component. The tolerances of the constraints on the local extrema of the high frequency component constitute another set of parameters which require careful adjustment based on experimental results, since they constitute a tradeoff between the sharpness of interpolated images and the appearance of undesirable artifacts. As for the complexity, a considerable computational burden is due to extraction of relevant local extrema. An efficient algorithm for that purpose is one of the further directions of this project. We also hope to be able to reduce the complexity based on experimental results by eliminating calculations which might be shown not to have a significant influence on the performance of the algorithm.

5.6 Appendix

5.6.1 Proof of Proposition 12

The design procedure described here follows along the same line as the design of maximally flat orthogonal filter banks [10, 47]. There is a bijective mapping between the set of FIR autocorrelation functions with real coefficients and the set of polynomials over $\mathbf{R}$ which are positive on the interval [0,1]. The autocorrelation function $H(z)H(z^{-1})$ of a real FIR filter defines a real coefficient polynomial of $sin^2\frac{\omega}{2}$ on the unit circle:

$$H(e^{j\omega})H(e^{-j\omega}) = P\left(sin^2\frac{\omega}{2}\right). \tag{5.48}$$

On the other hand, any polynomial $P(y)$, positive on [0,1], with $y = \left(\frac{1-z}{2}\right)\left(\frac{1-z^{-1}}{2}\right)$ defines an FIR autocorrelation function

$$P\left(\left(\frac{1-z}{2}\right)\left(\frac{1-z^{-1}}{2}\right)\right) = H(z)H(z^{-1}). \tag{5.49}$$

According to this, in order to obtain power complementary filters of length L having flatness of order $N_0 - 1$ at $\omega = \pi$ and flatness $N_1 - 1$ at $\omega = 0$, it is sufficient to find

$$P(y) = (1-y)^{N_0}Q(y) \tag{5.50}$$

such that $Q(y)$ has no zeros at $y = 1$, and $P'(y)$, the formal derivative of $P(y)$, has a zero of multiplicity $N_1 - 1$ at $y = 0$. For the case of maximally flat filters, the coefficients of the polynomial

$$Q(y) = \sum_{l=0}^{L-1-N_0} q_l y^l \tag{5.51}$$

are obtained from the requirement that $P'(y)$ has a $(N_1 - 1)$th order zero at $y = 0$ and that $P(0) = 1$. This gives

$$q_l = \begin{pmatrix} N_0 + l - 1 \\ l \end{pmatrix}, \quad l = 0, 1, ..., L - 1 - N_0. \tag{5.52}$$

$H_0(z)$ is then obtained from (5.49), while $H_1(z)$ is derived by factoring

$$1 - P\left(\left(\frac{1-z}{2}\right)\left(\frac{1-z^{-1}}{2}\right)\right) \tag{5.53}$$

as $H_1(z)H_1(z^{-1})$. That this is always possible to do is ensured by the fact that $P(y)$ has all of its extrema at $y = 0$ and $y = 1$, and therefore it is monotonically decreasing from 1 to 0 on the interval $[0, 1]$, making the expression (5.53) positive on the unit circle. This gives the class of maximally flat, $N_0 + N_1 = L$, power complementary filters.

For filters which are not maximally flat, $N_0 + N_1 < L$, $P(y)$ still has the form $(1-y)^{N_0}Q(y)$, and the coefficients q_l, $l = 0, 1, ..., N_1 - 1$, are determined as in (5.52). This gives the required flatness of the filters. The rest of coefficients of $Q(y)$ represent additional degrees of design freedom.

Several design examples are given in Table 2. This table contains coefficients of the autocorrelation functions of the low-pass filters, for the maximally flat power-complementary pairs. These filters are designed for several filter lengths, L, and with different multiplicities, N_1 of the zero of the high-pass filter at $z = 1$. For even filter lengths and $N_1 = L/2$, maximally flat power complementary design gives Daubechies' filters [10].

L	N_1	$H_0(z)H_0(z^{-1})$
4	1	$\frac{1}{2^6}[1,6,15,20]$
4	2	$\frac{1}{2^5}[-1,0,9,16]$
5	1	$\frac{1}{2^8}[1,8,28,56,70]$
5	2	$\frac{1}{2^8}[-3,-8,12,72,110]$
6	1	$\frac{1}{2^{10}}[1,10,45,120,210,252]$
6	2	$\frac{1}{2^8}[-1,-5,-5,20,70,98]$
6	3	$\frac{1}{2^9}[3,0,-25,0,150,256]$
7	1	$\frac{1}{2^{12}}[1,12,66,220,495,792,924]$
7	2	$\frac{1}{2^{12}}[-5,-36,-90,-20,405,1080,1428]$
7	3	$\frac{1}{2^{11}}[5,12,-30,-100,75,600,924]$
8	1	$\frac{1}{2^{14}}[1,4,91,364,1001,2002,3003,3432]$
8	2	$\frac{1}{2^{13}}[-3,-28,-105,-168,77,924,2079,2640]$
8	3	$\frac{1}{2^{14}}[15,70,21,-420,-665,1050,4725,6792]$
8	4	$\frac{1}{2^{12}}[-5,0,49,0,-245,0,1225,2048]$
9	1	$\frac{1}{2^{16}}[1,16,120,560,1820,4368,8008,11440,12870]$
9	2	$\frac{1}{2^{16}}[-7,-80,-392,-1008,-1092,1456,8008,16016,19734]$
9	3	$\frac{1}{2^{16}}[21,144,280,-336,-2100,-1904,5544,18480,25278]$
9	4	$\frac{1}{2^{16}}[-35,-80,280,784,-980,-3920,1960,19600,30318]$
10	1	$\frac{1}{2^{18}}[1,18,153,816,3060,8568,18564,31824,43758,48620]$
10	2	$\frac{1}{2^{16}}[-2,-27,-162,-552,-1080,-756,2184,8424,15444,18590]$
10	3	$\frac{1}{2^{16}}[7,63,207,168,-756,-2268,-1092,6552,18018,23738]$
10	4	$\frac{1}{2^{16}}[-14,-63,18,504,504,-1764,-3528,3528,19404,28358]$
10	5	$\frac{1}{2^{17}}[35,0,-405,0,2268,0,-8820,0,39690,65536]$

Table 5.2: *The autocorrelation function of various low-pass filters, $H_0(z)$, for power complementary pairs $(H_0(z), H_1(z))$ of perfect reconstruction nonsubsampled filter banks. The autocorrelation functions are given for several filter lengths, L, and with different multiplicities, N_1, of the zero at $z = 1$, of the high-pass filters, $H_1(z)$. The autoccorelation functions are symmtric so that only the irst L out of $2L-1$ coefficients are given in the table.*

Bibliography

[1] R. Balian. Un principe d'incertitude fort en théorie du signal ou en mécanique quantique. *C. R. Acad. Sc. Paris*, 292(Série 2), 1981.

[2] M.J. Bastiaans. Gabor's Signal Expansion and Degrees of Freedom of a Signal. *Proc. IEEE*, 68:538–539, 1980.

[3] W.R. Bennett. Spectra of Quantized Signals. *Bell System Technical Journ.*, 27:446–472, July 1948.

[4] Z. Berman and J.S. Baras. Properties of the Multiscale Maxima and Zero-Crossings Representations. *IEEE Trans. on Signal Processing, Special Issue on Wavelets and Signal Processing*, 41(12):3216–3231, December 1993.

[5] J. Canny. A Computational Approach to Edge Detection. *IEEE Trans. Patt. Recog. and Mach. Intell.*, 8:679–698, 1986.

[6] S.G. Chang. Image Interpolation Using Wavelet-Based Edge Enhancement and Texture Analysis. Master's thesis, Department of Elecrical Engineering and Computer Scinces, University of California at Berkeley, May 1995.

[7] S.G. Chang, Z. Cvetković, and M. Vetterli. Resolution Enhancement of Images Using Wavelet Transform Extrema Extrapolation. In *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, volume 4, pages 2379–2382, May 1995.

[8] A. Cohen, I. Daubechies, and J.C. Feauveau. Biorthogonal Bases of Compactly Supported Wavelets. *Commun. on Pure and Appl. Math.*, 45:485–560, 1992.

[9] Z. Cvetković and M. Vetterli. Discrete-Time Wavelet Extrema Representation: Properties and Consistent Reconstruction. *IEEE Trans. on Signal Processing*, 43(3):681–693, March 1995.

[10] I. Daubechies. *Ten Lectures on Wavelets.* SIAM, Philadelphia, Pennsylvania, 1992.

[11] I. Daubechies. Orthonormal Bases of Compactly Supported Wavelets. *Commun. on Pure and Appl. Math.*, 41:909–996, November 1988.

[12] I. Daubechies. The Wavelet Transform, Time-Frequency Localization and Signal Analysis. *IEEE Trans. on Information Theory*, 36(5):961–1005, September 1990.

[13] R.J. Duffin and A.C. Schaeffer. A Class of Nonharmonic Fourier Series. *Trans. Amer. Math. Soc.*, 72:341–366, March 1952.

[14] D. Gabor. Theory of Communications. *Journ. IEE*, 93(3):429–457, 1946.

[15] V.K. Goyal. Quantized Overcomplete Expansions: Analysis, Synthesis and Algorithms. Master's thesis, EECS Department, University of California at Berkeley, July 1995.

[16] V.K. Goyal, M. Vetterli, and N.T. Thao. Quantization of Overcomplete Expansions. In *Proc. Data Compression Conference*, pages 13–22, 1995.

[17] H. Greenspan and C. Anderson. Image Enhancement by Non-linear Extrapolation in Frequency Space. Personal communication, 1994.

[18] S. Hein and A. Zakhor. Reconstruction of Oversampled Band-Limited Signals from $\Sigma\Delta$ Encoded Binary Sequences. *IEEE Trans. on Signal Processing*, 42:799–811, April 1994.

[19] S. Hein and A. Zakhor. Reconstruction of Oversampled Band-Limited Signals from $\Sigma\Delta$ Encoded Binary Sequences. In *Proc. 25th Asilomar Conf. Signals Syst.*, pages 241–248, November 1991.

[20] A. Jain. *Fundamentals of Digital Image Processing.* Prentice-Hall, Englewood Cliffs, New Jersey, 1989.

[21] A.J.E.M. Janssen. Gabor Respresentation and Wigner Distribution of Signals. *Proc. IEEE*, pages 41.B.2.1–41.B.2.4, 1984.

[22] I.M. Kadec. The Exact Value of the Paley-Wiener Constant. *Sov. Math. Dokl.*, pages 559–561, 1964.

[23] F. Low. Complete Sets of Wave Packets. In *A Passion for Physics - Essays in Honor of Geoffrey Chew*, pages 17–22. World Scientific, Singapore, 1985.

[24] S. Mallat. A Theory of Multiresolution Signal Decomposition: The Wavelet Transform. *IEEE Trans. IEEE Trans. Patt. Recog. and Mach. Intell.*, 11(7):674–693, July 1989.

[25] S. Mallat. Zero Crossings of a Wavelet Transform. *IEEE Trans. Information Theory*, 37(4):1019–1033, July 1991.

[26] S. Mallat and W. L. Hwang. Singularity Detection and Processing with Wavelets. *IEEE Trans. on Information Theory*, 38(2):617–643, March 1992.

[27] S. Mallat and Z. Zhang. Matching Pursuits with Time-Frequency Dictionaries. *IEEE on Trans. Signal Processing, Special Issue on Wavelets and Signal Processing*, 41(12):3397–3415, December 1993.

[28] S. Mallat and S. Zhong. Characterization of Signals from Multiscale Edges. *IEEE Trans. Patt. Recog. and Mach. Intell.*, 14(7):710–732, July 1992.

[29] D. Marr and E. Hilderth. Theory of Edge Detection. *Proc. Royal Soc. London*, 207:187–217, 1980.

[30] Y. Meyer. Un contre-exemple à la conjecture de Marr et à celle de S. Mallat. Preprint, 1991.

[31] N.J. Munch. Noise Reduction in Weyl-Heisenberg Frames. *IEEE Trans. on Information Theory*, 38(3):608–616, March 1992.

[32] R. Neff and A. Zakhor. Matching Pursuit Video Coding at Very Low Bit Rates. In *Proc. Data Compression Conference*, pages 411–420, 1995.

[33] R.S. Orr. The Order of Computation for Finite Discrete Gabor Transforms. *IEEE Trans. on Signal Processing*, 41(1):122–130, January 1993.

[34] A. Papoulis. *Signal Analysis.* McGraw-Hill, 1972.

[35] H. Park. *A Computational Theory of Laurent Polynomial Rings and Multidimensional FIR Systems.* PhD thesis, Department of Mathematics, University of California at Berkeley, May 1995.

[36] S. Qiu and H. Feichtinger. Structural Properties of Gabor Transforms and Numerical Algorithms. *Signal Processing*, 1994. Submitted.

[37] S. Qiu and H. Feichtinger. Gabor-Type Matrices and Discrete Huge Gabor Transforms. In *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, volume 2, pages 1089–1092, May 1995.

[38] S. Quian and D. Chen. Discrete Gabor Transform. *IEEE Trans. on Signal Processing*, 41(7):2429–2438, April 1993.

[39] R.R. Schultz and R.L. Stevenson. A Bayesian Approach to Image Expansion for Improved Definition. *IEEE Trans. on Image Processing*, 3(3), May 1994.

[40] M. Shensa. The Discrete Wavelet Transform: Wedding the À Trous and Mallat Algorithms. *IEEE Trans. on Signal Processing*, 40(10):2464–2482, October 1992.

[41] E.P. Simoncelli, W.T. Freeman, E.H. Adelson, and D.J. Hegger. Shiftable Multiscale Transforms. *IEEE Trans. on Information Theory*, 38(2):587–608, March 1992.

[42] H.J.S. Smith. On Systems of Linear Indeterminate Equations and Congruences. *Philos. Trans. Royal Soc. London*, 151:293–326, 1861.

[43] N.T. Thao and M. Vetterli. Reduction of the MSE in R-times Oversampled A/D Conversion from $o(1/R)$ to $o(1/R^2)$. *IEEE Trans. on Signal Processing*, 42(1):200–203, January 1994.

[44] N.T. Thao and M. Vetterli. Oversampled A/D Conversion Using Alternate Projections. In *Proc. Conference on Information Sciences and Systems*, pages 241–248, March 1991.

[45] N.T. Thao and M. Vetterli. Deterministic Analysis of Oversampled A/D Conversion and Decoding Improvement Based on Consistent Estimates. *IEEE Trans. on Signal Processing*, 42(3):519–531, March 1994.

[46] M. Unser and A. Aldroubi. General Sampling Theory for Non-Ideal Acquisition Devices. Technical Report NCCR Rep. 16/92, National Institute of Health, Bethseda, Maryland, November 1992.

[47] P.P. Vaidyanathan. *Multirate Systems and Filter Banks.* Prentice-Hall, Englewood Cliffs, New Jersey, 1993.

[48] M. Vetterli. Filter Banks Allowing Perfect Reconstruction. *Signal Processing*, 10(3):219–244, April 1986.

[49] M. Vetterli and C. Herley. Wavelets and Filter Banks: Theory and Design. *IEEE Trans. on Signal Processing*, 40(9):2207–2232, September 1992.

[50] M. Vetterli and J. Kovačević. *Wavelets and Subband Coding.* Prentice-Hall, Englewood Cliffs, New Jersey, 1993.

[51] M.V. Wickerhauser. INRIA lectures on wavelet packet algorithms. Technical report, Department of Mathematics, Yale University, 1991.

[52] A. Witkin. Scale Space Filtering. In *Proc. Int. Joint Conf. Artificial Intell.*, 1983.

[53] J. Yao. Complete Gabor Transformation for Signal Representation. *IEEE Trans. on Image Processing*, 2(2):152–159, April 1993.

[54] D.C. Youla. On the Factorization of Rational Matrices. *IRE Trans. on Information Theory*, pages 172–189, 1961.

[55] D.C. Youla. Mathematical Theory of Image Restoration by the Method of Convex Projections. In H. Stark, editor, *Image Recovery - Theory and Applications.* Academic Press, New York, 1987.

[56] R.M. Young. *An Introduction to Nonharmonic Fourier Series.* Academic Press, New York, 1980.

Appendix A

Publications

[1] Z. Cvetković and M. Vetterli. Robustness of Overcomplete Expansions. In J. Zeevi and R. Coifman, editors, *Signal and Image Represantations in Combined Spaces.* Academic Press, 1996. To appear.

[2] Z. Cvetković and M. Vetterli. Error-Rate Characteristics of Oversampled A/D Conversion. *IEEE Trans. Information Theory*, November 1995. Submitted.

[3] Z. Cvetković and M. Vetterli. Error Analysis in Oversampled A/D Conversion and Quantization of Weyl-Heisenberg Frame Expansions. Technical Report No. UCB/ERL M95/48, UC Berkeley, May 1995.

[4] Z. Cvetković and M. Vetterli. Discrete-Time Wavelet Extrema Representation: Design and Consistent Reconstruction. *IEEE Trans. on Signal Processing*, 43(3):681-693, March 1995.

[5] Z. Cvetković and M. Vetterli. Oversampled Filter Banks. *IEEE Trans. Signal Processing*, December 1994. Submitted.

[6] Z. Cvetković. Analysis of Errors in Quantization of Weyl-Heisenberg Frame Expansions and Oversampled A/D Conversion. In *Proc. Int. Conf. Acoust., Speech and Signal Processing.* May 1996. To appear.

[7] M. Vetterli and Z. Cvetković. Oversampled FIR Filter Banks and Frames in $\ell^2(\mathbf{Z})$. In *Proc. Int. Conf. Acoust., Speech and Signal Processing.* May 1996. To appear.

[8] Z. Cvetković. Oversampled Modulated Filter Banks and Tight Gabor Frames in $\ell^2(\mathbf{Z})$. In *Proc. Int. Conf. Acoust., Speech and Signal Processing.* volume 2, pages 1456-1459, May 1995.

[9] S. G. Chang, Z. Cvetković and M. Vetterli. Resolution Enhancement of Images Using Wavelet Transform Extrema Extrapolation. In *Proc. Int. Conf. Acoust., Speech and Signal Processing.* volume 4, pages 2379-2382, May 1995.

[10] Z. Cvetković and M. Vetterli. Wavelet Extrema and Zero-Crossing Representations: Properties and Consistent Reconstruction. In *Proc. Int. Conf. Acoust., Speech and Signal Processing.* volume 3, pages 117-120, April 1994.