Forum for Interdisciplinary Mathematics

Khalil Ahmad · Abdullah

Wavelet Packets and Their Statistical Applications

FORUM FOR INTERDISCIPLINARY MATHEMATICS

FIM

शास्त्राणां गणितम् मूर्धनि स्थितम्

MATHEMATICS IS AT THE ZENITH OF ALL DISCIPLINES

Springer

Forum for Interdisciplinary Mathematics

The *Forum for Interdisciplinary Mathematics* series publishes high-quality monographs and lecture notes in mathematics and interdisciplinary areas where mathematics has a fundamental role, such as statistics, operations research, computer science, financial mathematics, industrial mathematics, and bio-mathematics. It reflects the increasing demand of researchers working at the interface between mathematics and other scientific disciplines.

More information about this series at http://www.springer.com/series/13386

Khalil Ahmad · Abdullah

Wavelet Packets and Their Statistical Applications

Springer

Khalil Ahmad
Department of Mathematics
Al-Falah University
Faridabad
India

Abdullah
Department of Mathematics
Zakir Husain Delhi College
New Delhi
India

ISSN 2364-6748 ISSN 2364-6756 (electronic)
Forum for Interdisciplinary Mathematics
ISBN 978-981-13-0267-1 ISBN 978-981-13-0268-8 (eBook)
https://doi.org/10.1007/978-981-13-0268-8

Library of Congress Control Number: 2018939017

Printed on acid-free paper

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
part of Springer Nature
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

To

Prof. H. P. Dikshit

Foreword

It has been our endeavor, to motivate and encourage teaching and research in emerging areas of mathematics and its applications in our country. One of the important initiatives taken by some of us including Prof. Khalil Ahmad in this direction was to organize an International Conference on Advances in Computational Mathematics in 1993 at the Indira Gandhi National Open University, New Delhi, which was supported by the National Board of Higher Mathematics (NBHM), Department of Atomic Energy, Government of India, and the National Science Foundation, USA. Wavelet and its applications were some of the main areas covered in the Conference, proceedings of which were published with Prof. C. A. Micchelli of T. J. Watson Research Center of IBM at Yorktown Heights, USA, and the undersigned as the editors. Since then, a few good schools of teaching and research emerged in the areas of wavelet and its applications in our country, like the one led by Prof. Khalil Ahmad, former Dean Faculty of Natural Sciences, Jamia Millia Islamia University, Delhi. Thus, Prof. Khalil Ahmad and several research students motivated by him have significantly contributed to the important areas of wavelet and its applications. Besides a large number of research contributions to his credit, Prof. Khalil Ahmad in his typically lucid and clear style of expression published jointly with F. A. Shah, a beautiful and exhaustive book Introduction to Wavelet Analysis with Applications, Real World Education Publishers, New Delhi (2013), which was well received by students and researchers alike. The present monograph entitled *Wavelet Packets and Their Statistical Applications* are jointly written by him and his coauthor Abdullah is an important contribution to wavelet packets which are a simple but powerful extension of wavelets and multiresolution analysis. A distinctive feature of the monograph is that two separate chapters are devoted to applications of wavelets and wavelet packets. The wavelet packets allow more flexibility in adapting the basis to the frequency contents of a signal, and it is easy to develop a fast wavelet packet transform. The power of wavelet packets lies on the fact that we have much more freedom in deciding which basis function we should use to represent the given function.

Generally, the contributions in the areas of wavelets and their extensions lay varying degree of emphasis on the theory, application, and computation aspects, but this monograph is different as starting from the essential mathematical tools, to a fairly complete and clear development of the theory and its potential areas of applications to computational implementation, all have been treated equally well with clarity. I am confident that this approach will be especially useful for interdisciplinary research in a variety of fields including the computational harmonic analysis and its applications to physical, biological, and medical sciences.

Starting with some basic results in functional analysis, wavelet analysis, and thresholding, construction of wavelet packets and band-limited wavelet packets and some of their important properties are presented in Chap. 2. Chapter 3 deals with the pointwise convergence of wavelet packet series, convolution bounds and convergence of wavelet packet series. Chapter 4 is devoted to characterizations of certain Lebesgue spaces, Hardy space, and Sobolev function spaces by using wavelet packets.

The last two Chaps. 5 and 6 provide a comprehensive study of applications of wavelets and wavelet packets to the important areas of signal and image processing. Speech denoising methods based on wavelets and wavelet packet decompositions of speech signals have been given. The proposed method of wavelet decomposition of speech signals and Wiener filter as post-filtering gives better results in comparison with Donoho's thresholding method. Similarly, the proposed method of wavelet packet decomposition of speech signals gives better results in comparison with those presented in J. P. Areanas (Combining adaptive filtering and wavelet techniques for vibration signal enhancement, Acustica, Paper ID-99, (2004), 1–8). A novel wavelet packet denoising method based on optimal decomposition and global threshold value has been also proposed for speech denoising in this work. To check the performance of this method, the signal-to-noise ratio is computed for the denoised speech signal.

Applications of wavelets in biomedical signals related to cardiac problems have been presented with clarity, especially in reference to ECG. An ECG signal denoising method based on wavelet transform is proposed. An optimum threshold value is estimated by computing the minimum error between detailed coefficients of noisy ECG signal and the original noise-free ECG signal. In comparison with the method used in A. Mikhled and D. Khaled (ECG signal denoising by wavelet transform thresholding, American Journal of Applied Sciences, 5(3) (2008), 276–281), the proposed method gives better result. Similarly, an ECG signal denoising method based on wavelet packet decomposition is proposed which gives a better result in comparison with the method proposed in M. Chendeb, K. Mohamad, and D. Jacques (Methodology of wavelet packet selection for event detection, Signal Processing Archive, 86(12) (2006), 3826–3841). For correct estimation of baseline drift in ECG signal, wavelet packet transform has been used in the present work.

Simulation result shows that level 8 is best for correct estimation of baseline drift in ECG signal. Applications of wavelets and wavelet packets to image processing are given in the last chapter.

Jabalpur (M.P.), India

H. P. Dikshit
Ex Vice Chancellor IGNOU

Preface

The theory of wavelets is the latest comer to the world of signal processing (more than 30 years now). It grew and brewed in different areas of science. Harmonic analysts had developed powerful time–frequency tools, electrical engineers were busy with subband coding, and quantum physicists were trying to understand coherent states. They were not aware of each other's progress until the late 1980's when a synthesis of all these ideas came to be, and what is now called wavelet theory contains all those ideas and much more. Wavelet is not one magical transform that solves all problems. It is a library of bases that is appropriate for a large number of situations where the traditional tools, for example, Fourier analysis, are not so good. There are many other problems which cannot be optimally treated with either of the known tools; therefore, new ones have to be designed.

A simple, but a powerful extension of wavelets and multiresolution analysis is wavelet packets, pioneered by Coifman, Meyer, Wickerhauser, and other researchers. The wavelet transform is generalized to produce a library of an orthonormal basis of modulated wavelet packets, where each basis comes with a fast transform. By generalizing the method of multiresolution decomposition, it is possible to construct an orthonormal basis for $L^2(\mathbb{R})$. Discrete wavelet packets have been thoroughly studied by Wickerhauser, who has also developed computer programs and implemented them. The wavelet packets allow more flexibility in adapting the basis to the frequency contents of a signal, and it is easy to develop a fast wavelet packet transform. The power of wavelet packet lies on the fact that we have much more freedom in deciding which basis function we use to represent the given function.

Wavelet packet functions are generated by scaling and translating a family of basic function shapes, which include father wavelet ϕ and mother wavelet ψ. In addition to ϕ, and ψ there is a whole range of wavelet packet functions ω_n. These functions are parametrized by an oscillation or frequency index n. A father wavelet corresponds to $n = 0$, so $\phi = \omega_0$. A mother wavelet corresponds to $n = 1$, so $\psi = \omega_1$. Larger values of n correspond to wavelet packets with more oscillations and higher frequency. Wavelet packets are particular linear combinations or superpositions of wavelets. They form bases which retain many of the orthogonality, smoothness, and localization properties of their parent wavelets.

The text begins with an elementary chapter on preliminaries such as basic concepts of functional analysis, a short tour of the wavelet transform, Haar scaling functions and function space, Lebesgue spaces $L^p(\mathbb{R})$, Hardy space, Sobolev spaces, Besov spaces, wavelets, Symlets wavelets, and Coiflets wavelets and thresholding.

Chapters 2 and 3 are devoted to the construction of wavelet packets, certain results on wavelet packets, band-limited wavelet packets, characterizations of wavelet packets, MRA wavelet packets, pointwise convergence, the convergence of wavelet packet series, and convolution bounds. Characterizations of function spaces like Lebesgue spaces $L^p(\mathbb{R})$, Hardy space $\mathscr{H}^1(\mathbb{R})$ and Sobolev spaces $L^{p,s}(\mathbb{R})$ in terms of wavelet packets are given in Chap. 4.

A signal can be defined as a function that conveys information, generally about the state or behaviour of the physical system. In almost every area of science and technology, signals must be processed to assist the extraction of information. Thus, the development of signal processing techniques and systems is of great importance. The presence of noise in speech signal can significantly reduce the intelligibility of speech and degrade automatic speech recognition performance. These noises may be due to the background noise of the environment in which the speaker is speaking, or it may be introduced by the transmission media during transmission of the speech signal. It is often necessary to perform speech denoising as the presence of noise, which severely degrades the speech signal. Chapter 5 is devoted to applications of wavelets and wavelet packets in speech denoising and biomedical signals.

The growth of media communication industry and demand of the high quality of visual information in the modern age has an interest to researchers to develop various image denoising techniques. In recent years, there has been a plethora of work on using wavelet thresholding techniques for removing noise in both signal and image processing. Chapter 6 is devoted to applications of wavelets and wavelet packets in image denoising. An exhaustive list of references is given at the end of the monograph.

The present book is intended to serve as a reference book for those working in the area of wavelet packets and their applications in different branches of mathematics and engineering, in particular in signal and image processing. It is also useful for statisticians and to those working in the industrial sector.

New Delhi, India
January 2018

Khalil Ahmad
Abdullah

Contents

About the Authors

Khalil Ahmad has been a Professor at the Department of Mathematics, Al-Falah University, India, since 2015. He previously served as a Professor at the Department of Mathematics, Jamia Millia Islamia, New Delhi, from 1975 to 2014 and worked at Aligarh Muslim University and DAV College, Kanpur. He received his Ph.D. in Mathematics from the University of Delhi in 1982. His major fields of research are functional analysis and applicable mathematics. He is a life member of several academic societies such as the Indian Mathematical Society, Indian Society of Industrial and Applied Mathematics, Jammu Mathematical Society, and Gwalior Academy of Mathematical Sciences. He has over 50 research papers to his credit, published in national and international journals of repute. He is the author of 12 books and has delivered lectures at over 60 mathematical conferences.

Abdullah is an Assistant Professor at the Department of Mathematics, Zakir Husain Delhi College, University of Delhi, New Delhi, India. His main fields of research are wavelet analysis and frame theory. He received his Ph.D. in Mathematics from Jamia Millia Islamia, New Delhi, in 2007. He has over 25 research papers to his credit and has attended more than 20 mathematical conferences.

Chapter 1
Preliminaries

1.1 Introduction

In this chapter we enlist those concepts and results concerning functional analysis, wavelet analysis and thresholding which are already known in the literature and we require in the subsequent chapters.

1.2 Basic Concepts of Functional Analysis

Throughout, the functions f, g, φ, ψ, and ω_n will stand for $f(x)$, $g(x)$, $\varphi(x)$, $\psi(x)$, and $\omega_n(x)$, respectively.

Let $\mathbb{Z}$ and $\mathbb{R}$ denote the set of integers and real numbers, respectively, and $\mathbb{T}$ denote the unit circle in the complex plane which can be identified with the interval $[-\pi, \pi)$. The inner product of two functions $f \in L^2(\mathbb{R})$ and $g \in L^2(\mathbb{R})$ is denoted by $\langle f, g\rangle$ and is defined as

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)\, \overline{g(x)} dx.$$

The norm of $f \in L^2(\mathbb{R})$ is written as $\|f\|$. The Fourier transform of any function $f \in L^2(\mathbb{R})$ is denoted by $\hat{f}$ and is defined as

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx.$$

The inverse Fourier transform of any function $g \in L^2(\mathbb{R})$ is denoted by $\check{g}$ and is defined as

$$\check{g} = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(\xi) e^{i\xi x}\, d\xi$$

K. Ahmad, *Wavelet Packets and Their Statistical Applications*,
Forum for Interdisciplinary Mathematics,
https://doi.org/10.1007/978-981-13-0268-8_1

and if we apply it to $g = \hat{f}$ we obtain $\check{g} = f$, that is $(\hat{f})^{\vee} = f$. With this definition, the Plancherel theorem asserts that

$$\langle f,\ g\rangle = \frac{1}{2\pi}\langle \hat{f},\ \hat{g}\rangle.$$

The set $l^2(\mathbb{Z})$ is the vector space of square-summable sequences, i.e.,

$$l^2(\mathbb{Z}) = \left\{ \{h_k\}_{k\in\mathbb{Z}} : \sum_{k=-\infty}^{\infty} \mid h_k \mid^2 < \infty \right\}.$$

Throughout we shall denote $\mathbb{R}^0$, S for the regularity class and Schwartz class, respectively. For the dual of Schwartz class, we denote S'.

Lemma 1.2.1 *Let $\mathbb{H}$ be a Hilbert space and $\{e_j : j = 1, 2, \ldots\}$ be a family of elements of $\mathbb{H}$. Then*

(i) $\|f\|^2 = \sum_{j=1}^{\infty} |\langle f, e_j\rangle|^2$ *holds for all* $f \in \mathbb{H}$

if and only if

(ii) $f = \sum_{j=1}^{\infty} \langle f, e_j\rangle e_j$, *with convergence in* $\mathbb{H}$, *for all* $f \in \mathbb{H}$.

Lemma 1.2.2 *Suppose $\{e_j : j = 1, 2, \ldots\}$ be a family of elements in a Hilbert space $\mathbb{H}$ such that equality (i) in Lemma 1.2.1 holds for all f belonging to a dense subset D of $\mathbb{H}$, then the equality is valid for all $f \in \mathbb{H}$.*

Lemma 1.2.3 *Let C be a positive integer and let $\{v_j : j \geq 1\}$ be a family of vectors in a Hilbert space $\mathbb{H}$ such that*

(i) $\sum_{n=1}^{\infty} \|v_n\|^2 = C$ *and*

(ii) $v_n = \sum_{m=1}^{\infty} \langle v_n,\ v_m\rangle v_m$ *for all* $n \geq 1$.

Let $\mathbb{F} = \overline{span\{v_j : j \geq 1\}}$. Then, $dim\, \mathbb{F} = \sum_{j=1}^{\infty} \|v_j\|^2 = C$ (Number of basis elements of $\mathbb{F}$).

Definition 1.2.4 For a given function g defined on $\mathbb{R}$, we say that a bounded function $H : [0, \infty) \to \mathbb{R}^+$ is a radial decreasing L^1-majorant of g if $|g(x)| \leq H(|x|)$ and H satisfies the following conditions

$$\begin{cases} (i) H \in L^1[0, \infty) \\ (ii) H \text{ is decreasing} \\ (iii) H(0) < \infty \end{cases} \tag{1.2.1}$$

The set of all bounded radially decreasing functions is denoted by RB.

Lemma 1.2.5 *Let H be a function on $[0, \infty)$ satisfying condition (1.2.1). Then*

$$\sum_{k\in\mathbb{Z}} H\left(|x-k|\right) H\left(|y-k|\right) \leq CH\left[\frac{|x-y|}{2}\right], \quad \forall x, y \in \mathbb{R}$$

where C is a constant depending on H.

Definition 1.2.6 The point $x \in \mathbb{R}$ is said to be a Lebesgue point of a function f on $\mathbb{R}$ if f is integrable in some neighborhood of x and

$$\lim_{\varepsilon\to 0} \frac{1}{V(B_\varepsilon)} \int_{B_\varepsilon} |f(x) - f(x+y)|dy = 0$$

where B_ε denotes the ball of radius ε about the origin and V denotes volume.

Definition 1.2.7 For a function g defined on $\mathbb{R}$ and for a real number $\lambda > 0$, the maximal function is defined by

$$g_\lambda^*(x) = \sup_{y\in\mathbb{R}} \frac{|g(x-y)|}{(1+|y|)^\lambda}, \quad x \in \mathbb{R}. \tag{1.2.2}$$

Definition 1.2.8 Hardy–Littlewood maximal function, $\mathscr{M}f(x)$, is defined by

$$\mathscr{M}f(x) = \sup_{r>0} \frac{1}{2r} \int_{|y-x|\leq r} |f(y)|\,dy \tag{1.2.3}$$

for a locally integrable function f on $\mathbb{R}$.

It is well known that $\mathscr{M}$ is bounded on $L^p(\mathbb{R})$, $1 < p \leq \infty$. An important property of $\mathscr{M}$ that we shall need is the following vector-valued inequality:

Lemma 1.2.9 *Suppose $1 < p, q < \infty$; then there exists a constant $C_{p,q}$ such that*

$$\left\| \left\{ \sum_{i=1}^{\infty} (\mathscr{M} f_i)^q \right\}^{\frac{1}{q}} \right\|_{L^p(\mathbb{R})} \leq C_{p,q} \left\| \left\{ \sum_{i=1}^{\infty} |f_i|^q \right\}^{\frac{1}{q}} \right\|_{L^p(\mathbb{R})} \tag{1.2.4}$$

for any sequence $\{f_i : i = 1, 2, \ldots\}$ of locally integrable functions.

Lemma 1.2.10 *Let g be a band-limited function on $\mathbb{R}$ such that $g_\lambda^*(x) < \infty$ for all $x \in \mathbb{R}$. Then, there exists a constant C_λ such that*

$$g_\lambda^*(x) \leq C_\lambda \left\{ \mathscr{M}(|g|^{\frac{1}{\lambda}})(x) \right\}^\lambda, \quad x \in \mathbb{R}. \tag{1.2.5}$$

Lemma 1.2.11 *If g is a band-limited function, i.e., support of $\hat{g}$ is contained in a finite interval, defined on $\mathbb{R}$ such that $g \in L^p(\mathbb{R})$, $0 < p \leq \infty$, then we have $g_\lambda^*(x) < \infty$ for all $x \in \mathbb{R}$.*

Lemma 1.2.12 (Hörmander–Mihlin Multiplier Theorem). *Let $\mathbb{H}_0$ and $\mathbb{H}_1$ be two Hilbert spaces, and we denote by $\mathscr{L}(\mathbb{H}_0, \mathbb{H}_1)$ the set of all bounded linear operators T from $\mathbb{H}_0$ to $\mathbb{H}_1$. Assume that m is a function defined on $\mathbb{R}$ with values in $\mathscr{L}(\mathbb{H}_0, \mathbb{H}_1)$ such that*

$$\|(D^j m)(\xi)\|_{\mathscr{L}(\mathbb{H}_0,\ \mathbb{H}_1)} \leq B \frac{1}{|\xi|^j}, \quad j = 0, 1 \tag{1.2.6}$$

for some positive constant $B < \infty$. Then, the operator T_m given by

$$(T_m f)\hat{}(\xi) = m(\xi) \hat{f}(\xi) \quad \textit{for all} \quad f \in S(\mathbb{H}_0)$$

can be extended to a bounded linear operator from $L^p(\mathbb{R};\ \mathbb{H}_0)$ to $L^p(\mathbb{R};\ \mathbb{H}_1)$, $1 < p < \infty$. That is, there exists a constant C, $0 < C < \infty$ such that

$$\|T_m f\|_{L^p(\mathbb{R};\ \mathbb{H}_1)} \leq C \|f\|_{L^p(\mathbb{R};\ \mathbb{H}_0)} \qquad \textit{for all} \ \ f \in L^p(\mathbb{R};\ \mathbb{H}_0). \tag{1.2.7}$$

Lemma 1.2.13 *Given $\varepsilon > 0$ and $1 \leq r < 1 + \varepsilon$, there exists a constant C such that for all sequences $\left\{s_{l,k} : l, k \in \mathbb{Z}\right\}$ of complex numbers and all $x \in I_{l,k}$,*

(a) $$\sum_{k' \in \mathbb{Z}} \frac{|s_{l',k'}|}{\left(1 + 2^{l'} |2^{-l}k - 2^{-l'}k'|\right)^{1+\varepsilon}} \leq C \left[\mathscr{M} \left(\sum_{k' \in \mathbb{Z}} |s_{l',k'}|^{\frac{1}{r}} \chi_{I_{l',k'}} \right)(x) \right]^r \ \textit{if}\ l' \leq l$$

and

(b) $$\sum_{k' \in \mathbb{Z}} \frac{|s_{l',k'}|}{\left(1 + 2^{l} |2^{-l'}k' - 2^{-l}k|\right)^{1+\varepsilon}} \leq C 2^{(l'-l)r} \times \left[\mathscr{M} \left(\sum_{k' \in \mathbb{Z}} |s_{l',k'}|^{\frac{1}{r}} \chi_{I_{l',k'}} \right)(x) \right]^r \quad \textit{if}\ l' \geq l$$

where $\mathscr{M}$ is the Hardy–Littlewood maximal function defined in Definition 1.2.6 and $I_{l,k} = \left[2^{-l}k, 2^{-l}(k+1)\right]$.

Definition 1.2.14 We say that a function φ defined on $\mathbb{R}$ belongs to the regularity class $\mathbb{R}^0$ if there exist constants C_0, C_1, γ, and $\varepsilon > 0$ such that

(i) $\displaystyle\int_{\mathbb{R}} \varphi(x) dx = 0$

(ii) $|\varphi(x)| \leq \dfrac{C_0}{(1 + |\alpha|)^{2+\gamma}}$ for all $x \in \mathbb{R}$

(iii) $|\varphi'(x)| \leq \dfrac{C_1}{(1 + |\alpha|)^{1+\varepsilon}}$ for all $x \in \mathbb{R}$.

Lemma 1.2.15 *Let $\varepsilon > 0$. Suppose that g and h satisfy*

(a) $|g(x)| \le \dfrac{C_1}{(1+|x|)^{1+\varepsilon}}$ *for all* $x \in \mathbb{R}$ *and*

(b) $|h(x)| \le \dfrac{C_2}{(1+|x|)^{1+\varepsilon}}$ *for all* $x \in \mathbb{R}$

with C_1 and C_2 independent of $x \in \mathbb{R}$. Then, there exists a constant C such that for all $l, k, l', k' \in \mathbb{Z}$ and $l \le l'$, we have

$$|(g_{l,k} * h_{l',k'})(x)| \le \frac{C2^{\frac{1}{2}(l-l')}}{(1+2^l|x-2^{-l}k-2^{-l'}k'|)^{1+\varepsilon}} \quad \textit{for all } x \in \mathbb{R}.$$

Lemma 1.2.16 *Let $r \ge \varepsilon > 0$ and $N \in \mathbb{N}$. Suppose that g and h satisfy*

(a) $\left|\dfrac{d^n g}{dx^n}(x)\right| \le \dfrac{C_{n,1}}{(1+|x|)^{1+\varepsilon}}$ *for all* $x \in \mathbb{R}$ *and* $0 \le n \le N+1$

(b) $\displaystyle\int_{\mathbb{R}} x^n\, h(x)\, dx = 0$ *for all* $0 \le n \le N$

(c) $|h(x)| \le \dfrac{C_2}{(1+|x|)^{2+N+r}}$ *for all* $x \in \mathbb{R}$

with $C_{n,1},\ 0 \le n \le N+1$, and C_2 independent of $x \in \mathbb{R}$. Then, there exists a constant C such that for all $l, k, l', k' \in \mathbb{Z}$ and $l \le l'$, we have

$$|(g_{l,k} * h_{l',k'})(x)| \le \frac{C2^{(l-l')(\frac{1}{2}+N+1)}}{(1+2^l|x-2^{-l}k-2^{-l'}k'|)^{1+\varepsilon}} \quad \text{for all } x \in \mathbb{R}.$$

For $N \in \mathbb{N} \cup \{-1\}$, let $\mathscr{D}^N$ be the set of all functions f defined on $\mathbb{R}$ for which there exist constants $\varepsilon > 0$ and $C_n < \infty,\ n = 0, 1, \ldots, N+1$, such that

$$\left|D^n f(x)\right| \le \frac{C_n}{(1+|x|)^{1+\varepsilon}} \ \text{ for all } x \in \mathbb{R} \text{ and } 0 \le n \le N+1.$$

We write $\mathscr{M}^N$ for the set of all functions f defined on $\mathbb{R}$ for which there exist constants $\gamma > 0$ and $C < \infty$ such that

$$\int_{\mathbb{R}} x^n f(x)\, dx = 0 \quad \text{for } n = 0, 1, \ldots, N$$

and $$|f(x)| \le C \frac{1}{(1+|x|)^{2+N+\gamma}} \quad \text{for all } x \in \mathbb{R}.$$

Definition 1.2.17 For a nonnegative integer s, let $\mathbb{R}^s = \mathscr{D}^s \cap \mathscr{M}^s$; that is, $f \in \mathbb{R}^s$ if there exist constants $\varepsilon > 0,\ \gamma > 0,\ C < \infty,$ and $C_n < \infty,\ n = 1, 2, \ldots, s+1,$ such that

(i) $\int_{\mathbb{R}} x^n f(x)\,dx = 0 \quad$ for $n = 0, 1, \ldots, s$

(ii) $|f(x)| \le \dfrac{C}{(1+|x|)^{2+s+\gamma}} \quad$ for all $x \in \mathbb{R}$

(iii) $\left|D^n f(x)\right| \le \dfrac{C_n}{(1+|x|)^{1+\varepsilon}} \quad$ for all $x \in \mathbb{R},\ n = 1, 2, \ldots, s+1.$

Definition 1.2.18 The Schwartz space S is the subspace of C^∞ (the set of all bounded continuous functions) given

$$S = \bigcap_{N=1}^{\infty} \left\{ f \in C^\infty : \sum_{|\alpha| \le N} \sup_{x \in \mathbb{R}^n} < x >^N |\partial^\alpha f(x)| < \infty \right\}.$$

The topology of S is the weakest one for which the mapping $f \in\to p_N(f) \in \mathbb{R}$ is continuous for all $N \in \mathbb{N}$, where

$$p_N(f) = \sum_{|\alpha| \le N} \sup_{x \in \mathbb{R}^n} < x >^N |\partial^\alpha f(x)|, \quad < x >= \sqrt{1+|x|^2}.$$

Definition 1.2.19 One defines

$$S' = \{f : S \to \mathbb{C} : f \text{ is linear and continuous}\}$$

One equips S' with the weakest topology so that the mapping

$$f \in S' \to < f, \varphi > \in \mathbb{C}$$

is continuous for all $\varphi \in S$.

1.3 A Short Tour of Wavelet Transform

The need of simultaneous representation and localizations of both time and frequency for nonstationary signals (e.g., music, speech, images) led toward the advancement of wavelet transform from the popular Fourier transform. Different "time–frequency representations" (TFR) are very informative in understanding and modeling of wavelet transform [68, 124, 143].

1.3.1 Fourier Transform

Fourier transform is a well-known mathematical tool to transform time-domain signal to frequency-domain for efficient extraction of information and it is reversible also. For a signal $x(t)$, the Fourier transform is given by:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt$$

The Fourier transform has a great ability to detain signal's frequency content as long as $x(t)$ is composed of few stationary components (e.g., sine waves). However, any abrupt change in time for nonstationary signal $x(t)$ is spread out over the whole frequency axis in $X(f)$. Hence, the time-domain signal sampled with Dirac delta function is highly localized in time but spills over entire frequency band and vice versa. The limitation of Fourier transform is that it cannot offer both time and frequency localization of a signal at the same time.

To overcome the limitations of the standard Fourier transform, Gabor [117] introduced the initial concept of short-time Fourier transform (STFT). The advantage of STFT is that it uses an arbitrary but fixed-length window $g(t)$ for analysis, over which the actual nonstationary signal is assumed to be approximately stationary. The STFT decomposes such a pseudo-stationary signal $x(t)$ into a two-dimensional time–frequency representation $S(\tau, f)$ using that sliding window $g(t)$ at different times τ. Thus, the Fourier transform of windowed signal $x(t) * (t - \tau)$ yields STFT as:

$$\text{STFT}_x(\tau, f) = \int_{-\infty}^{\infty} x(t)g * (t - \tau)e^{-j2\pi ft}dt$$

Filter bank interpretation is an alternative way of seeing "windowing of the signal" viewpoint of STFT [19, 219]. With the modulated filter bank, a signal can be seen as passing through a bandpass filter centered at frequency f with an impulse response of the window function modulated to that frequency. From this dual interpretation, a possible drawback related to time–frequency resolution of STFT can be shown through "Heisenberg's uncertainty principle" [41, 256]. For a window $g(t)$ and its Fourier transform $G(f)$, both centered around the origin in time as well as in frequency, i.e., satisfying $\int t|g(t)|^2dt = 0$ and $\int f|G(f)|^2df = 0$. Then, the spreads in time and frequency are defined as:

$$\triangle_t^2 = \frac{\int_{-\infty}^{\infty} t^2|g(t)|^2dt}{\int_{-\infty}^{\infty} |g(t)|^2dt}, \quad \triangle_f^2 = \frac{\int_{-\infty}^{\infty} f^2|G(f)|^2df}{\int_{-\infty}^{\infty} |G(f)|^2df}$$

Thus, the time–frequency resolution for STFT is lower bounded by their product as:

$$\text{Time-Bandwidth product } \triangle_t\triangle_f \geq \frac{1}{4\pi}.$$

Once a window has been chosen for STFT, the time–frequency resolution is fixed over the entire time–frequency plane because the same window is used at all frequencies. There is always a trade-off between time resolution and frequency resolution in STFT.

1.3.2 Wavelet Transform

Fixed resolution limitation of STFT can be resolved by letting the resolution $\triangle_t$ and $\triangle_f$ vary in time–frequency plane in order to obtain multiresolution analysis. The wavelet transform (WT) in its continuous form; i.e., CWT provides a flexible time–frequency window, which narrows when observing high-frequency phenomena and widens when analyzing low-frequency behavior. Thus, time resolution becomes arbitrarily good at high frequencies, while the frequency resolution becomes arbitrarily good at low frequencies. This kind of analysis is suitable for signals composed of high-frequency components with short duration and low-frequency components with long duration, which is often the case in practical situations [230].

When analysis is viewed as a filter bank, the wavelet transform, generally termed as standard discrete wavelet transform (DWT), is seen as a composition of bandpass filters with constant relative bandwidth such that $\triangle_f/f$ is always constant. As $\triangle_f$ changes with frequencies, corresponding time resolution $\triangle_t$ also changes so as to satisfy the uncertainty condition. The frequency responses of bandpass filters are logarithmically spread over frequency.

1.3.3 Continuous Wavelet Transform

It is very clear that wavelet means "small wave," so wavelet analysis is about analyzing signal with short duration finite energy functions. Mathematically, wavelet can be represented as:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right)$$

with some wavelet admissibility conditions:

$$C_\psi = \int_0^\infty \frac{|\hat{\psi}(\omega)|}{\omega} d\omega < \infty$$

and

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1$$

Where "b" is location parameter, "a" is scaling parameter, $\hat{\psi}(\omega)$ is the Fourier transform, which ensures that $\hat{\psi}(\omega)$ goes to zero quickly as $\omega \to 0$. In fact to guarantee that $C_\psi < \infty$, we must impose $\hat{\psi}(0) = 0$. Wavelet transform is defined as:

$$W(a,b) = \int_t f(t)\frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) dt. \tag{1.3.1}$$

According to above equation, for every (a, b), we have a wavelet transform coefficient, representing how much the scaled wavelet is similar to the function at location $t = \frac{a}{b}$.

Now, we can say that continuous wavelet transform (CWT) is a function of two parameters and, therefore, contains a high amount of extra (redundant) information when analyzing a function. A critical sampling of the CWT

$$W(a, b) = \int_t f(t) \frac{1}{\sqrt{|a|}} \psi \left(\frac{t-b}{a} \right) dt$$

is obtained via $a = 2^{-j}$, where j and k (in next integral) are integers representing the set of discrete translations and discrete dilations. Upon substitution, Eq. (1.3.1) can become

$$\int f(t) 2^{j/2} \psi(2^j t - k) dt$$

which is function of j and k. We denote it by $W(j, k)$. In general, $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$ is wavelet for all integers j and k and produces an orthogonal basis. We call $\psi_{0,0}(t) = \psi(t)$ as mother wavelet. Other wavelets are produced by translation and dilation of the mother wavelet [60].

In earlier time, a phenomenon that is known as Heisenberg's uncertainly principle, which says that a signal cannot be simultaneously localized in time and frequency. Wavelets are an attempt to overcome this shortcoming. They provide a way to do time–frequency analysis. The idea is that one chooses a "mother wavelet," i.e., a function subject to some conditions like mean value by using two variable base (one for the amount to shift and one for the amount of dilation); we are able to introduce enough redundancy to maintain the local properties of the original function.

Overall, we can say that continuous wavelet transform is defined as the sum over all time of the signal multiplied by scaled, shifted version of the wavelet function ψ:

$$C(\text{scale, position}) = \int_{-\infty}^{\infty} f(t)\psi((\text{scale, position}))dt$$

The results of the CWT are many wavelet coefficients C, which are a function of scale and position. Multiplying each coefficient by the appropriately scaled and shifted wavelet yields the constituent wavelets of the original signal.

Now, we are very eager to know that what is continuous in CWT, because in any signal processing real-world data must be performed on a discrete signal. The speciality of CWT is that it can operate at every scale, from that of the original signal up to some maximum scale that we determine by trading off our need for detailed analysis. The CWT is also continuous in term of shifting: during computation, which makes continuous wavelet transform distinguishing from others.

1.3.4 Discrete Wavelet Transform

Unlike conventional methods, in wavelet transform, one can use a single function and its dilations and translations to generate a set of orthonormal basis functions to represent a signal. Numbers of such functions are infinite, and we can choose one that suits to the application. Unfortunately, most of the wavelets used in discrete wavelet transform are fractal in nature. They are expressed in terms of a recurrence relation so that to see them we must do several iterations. But fortunately, we have two special functions known as Haar wavelet functions and Haar scaling functions, which have explicit expression. To understand mathematical (geometry) approach insight wavelet transform, Haar function is the only hope. Scaling functions and wavelet functions are just like twins; corresponding to wavelet function there is a scaling function. The details about these will be discussed in next sections of this chapter [18, 257].

1.4 Haar Scaling Functions and Function Space

In discrete wavelet transform, we have to deal with basically two sets of functions—scaling and wavelet functions. Understanding the relation between these two functions, consider Haar scaling function $\varphi(t)$ defined in Eq. (1.4.1) and shown in Fig. 1.1

$$\varphi(t) = \begin{cases} 1, & 0 \le t \le 1 \\ 0 & \text{elsewhere} \end{cases} \tag{1.4.1}$$

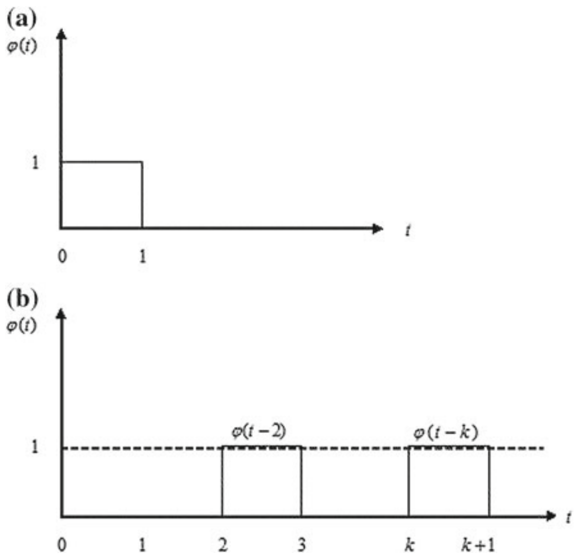

Fig. 1.1 **a** Haar scaling function and **b** translation of Haar scaling function