Dynamic Modeling and Econometrics in
Economics and Finance 20

Marco Gallegati
Willi Semmler *Editors*

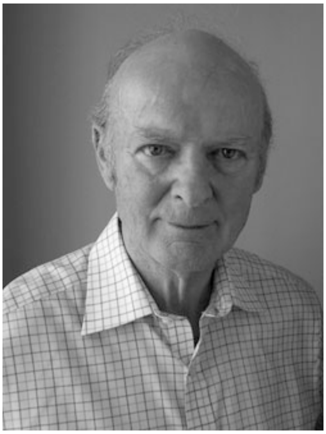
Wavelet Applications in Economics and Finance

Springer

Dynamic Modeling and Econometrics in Economics and Finance

Volume 20

For further volumes:
http://www.springer.com/series/5859



James B. Ramsey

Marco Gallegati • Willi Semmler
Editors

Wavelet Applications in Economics and Finance

Springer

Editors

Marco Gallegati
Faculty of Economics "G.Fuà"
Polytechnic University of Marche
Ancona
Italy

Willi Semmler
New School for Social Research
The New School University
New York
USA

ISSN 1566-0419 Dynamic Modeling and Econometrics in Economics and Finance
ISBN 978-3-319-07060-5 ISBN 978-3-319-07061-2 (eBook)
DOI 10.1007/978-3-319-07061-2
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014945649

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

Mater semper certa est, pater numquam ("The mother is always certain, the father is always uncertain") is a Roman-law principle which has the power of *praesumptio iuris et de iure*. This is certainly true for biology, but not for *wavelets* in economics which have a true father: James Ramsey.

The most useful property of wavelets is its ability to decompose a signal into its time scale components. Economics, like many other complex systems, include variables simultaneously interacting on different time scales so that relationships between variables can occur at different horizons. Hence, for example, we can find a stable relationship between durable consumption and income. And the literature is soaring: from money–income relationship to Phillips curve, from financial market fluctuations to forecasting. But this feature threatens to undermine the very foundations of the Walrasian construction. If variables move differently at different time scales (stock market prices in nanoseconds, wages in weeks, and investments in months), then also a linear system can produce chaotic effects and market self-regulation is lost. If validated, wavelet research becomes a silver bullet.

James is also an excellent sailor (in 2003 he sailed across the Atlantic to keep his boat from North America to Turkey), and his boat braves the streams with "nonchalance": by the way if you are able to manage wavelets, you are also ready for waves.

Ancona, Italy   Mauro Gallegati
March 2, 2014

Preface

James Bernard Ramsey received his B.A. in Mathematics and Economics from the University of British Columbia in 1963, and his M.A. and Ph.D. in Economics from the University of Wisconsin, Madison in 1968 with the thesis "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis". After being Assistant and Associate Professor at the Department of Economics of Michigan State University, he became Professor and Chair of Economics and Social Statistics at the University of Birmingham, England, from 1971 to 1973. He went back to the US as Full Professor at Michigan State University until 1976 and finally moved to New York University as Professor of Economics and Chair of the Economics Department between 1978 and 1987, where he remained for 37 years until his retirement in 2013. Fellow of the American Statistical Association, Visiting Fellow at the School of Mathematics (Institute for Advanced Study) at Princeton in 1992–1993, and ex-president of the Society for Nonlinear Dynamics and Econometrics, James Ramsey was also a jury member of the Econometric Game 2009. He has published 7 books and more than 60 articles on nonlinear dynamics, stochastic processes, time series, and wavelet analysis with special emphasis on the analysis of economic and financial data.

This book intends to honor James B. Ramsey and his contribution to economics on occasion of his recent retirement from academic activities at the NYU Department of Economics. This festschrift, as it is called in the German tradition, intends to honor an exceptional scholar whose fundamental contributions have influenced a wide range of disciplines, from statistics to econometrics and economics, and whose lifelong ideas have inspired more than a generation of researchers and students.

He is widely acclaimed for his pioneering work in the early part of his career on the general specifications test for the linear regression model, Ramsey's RESET test, which is part of any econometric software now. He is also well known for his contributions to the theory and empirics of chaotic and nonlinear dynamical systems. A significant part of his work has also been devoted to the development of genuine new ways of processing data, as for instance the application of functional data analysis or the use of wavelets in terms of nonparametric analysis.

Each year the Society for Nonlinear Dynamics and Econometrics, at its Annual Conference, awards two James Ramsey prizes for top graduate papers in econometrics. This year there will also be a set of special sessions dedicated to his research. One of these sessions will be devoted to wavelet analysis, an area where James work has had a great outstanding impact in the last twenty years. James Ramsey and his coauthors have provided early applications of wavelets in economics and finance by making use of discrete wavelet transform (DWT) in decomposing economic and financial data. These works paved the way to the application of wavelet analysis for empirical economics. The articles in this book are comprised of contributions by colleagues, former students, and researchers covering a wide range of wavelet applications in economics and finance and are linked to or inspired by the work of James Ramsey.

We have been working with James continuously over the last 10 years and have always been impressed by his competence, motivation, and enthusiasm. Our collaboration with James was extraordinarily productive and an inspiration to all of us. Working together we developed a true friendship strengthened by virtue of the pleasant meetings held periodically at James office on the 7th floor of the NYU Department of Economics, which became an important space for discussing ongoing as well as new and exciting research projects. As one of his students has recently written, rating James' Statistics class: "He is too smart to be teaching!" Sometimes our impression was that he could also have been too smart for us as coauthor. This book is a way to thank him for the privilege we have had to met and work with him.

Ancona, Italy — Marco Gallegati
New York, NY — Willi Semmler
March 2014

Introduction

Although widely used in many other disciplines like geophysics, engineering (sub-band coding), physics (normalization groups), mathematics (C-Z operators), signal analysis, and statistics (time series and threshold analysis), wavelets still remain largely unfamiliar to students of economics and finance. Nonetheless, in the past decade considerable progress has been made, especially in finance and one might say that wavelets are the "wave of the future". The early empirical results show that the separation by time scale decomposition analysis can be of great benefit for a deeper understanding of economic relationships that operate simultaneously at several time scales. The "short and the long run" can now be formally explored and studied.

The existence of time scales, or "planning horizons", is an essential aspect of economic analysis. Consider, for example, traders operating in the market for securities: some, the fundamentalists, may have a very long view and trade looking at market fundamentals and concentrate their attention on "long run variables" and average over short run fluctuations. Others, the chartists, may operate with a time horizon of only weeks, days, or even hours. What fundamentalists deem to be variables, the chartists deem constants. Another example is the distinction between short run adaptations to changes in market conditions; e.g., merely altering the length of the working day, and long run changes in which the firm makes strategic decisions and installs new equipment or introduces new technology.

A corollary of this assumption is that different planning horizons are likely to affect the structure of the relationships themselves, so that they might vary over different time horizons or hold at certain time scales, but not at others. Economic relationship might also show negative relationship over some time horizon, but a positive one over others. These different time scales of variation in the data may be expected to match the economic relationships more precisely than a single time scale using aggregated data. Hence, a more realistic assumption should be to separate out different time scales of variation in the data and analyze the relationships among variables at each scale level, not at the aggregate level. Although the concepts of the "short-run" and of the "long-run" are central for modeling economic and financial decisions, variations in those relationships across time scales are seldom discussed nor empirically studied in economics and finance.

The theoretical analysis of time, or “space series” split early on into the “continuous wavelet transform”, CWT, and into “discrete wavelet transform”. DWT. The latter is often more useful for applying to regular time series analysis with observations at discrete intervals. Wavelets provide a multi-resolution decomposition analysis of the data and can produce a synthesis of the economic relationships that is parameter preserving. The output of wavelet transforms enables one to decompose the data in ways that are potentially revealing relationships that are not visible using standard methods on “scale aggregated” data. Given their ability to isolate the bounds on the temporary frequency content of a process as a function of time, it is a great advantage of those transforms to be able to rely only on local stationarity that is induced by the system, although Gabor transforms provide a similar service for Fourier series and integrals.

The key lesson in synthesizing the wavelet transforms is to facilitate and develop the theoretical insight into the interdependence of economic and financial variables. New tools are most likely to generate new ways of looking at the data and new insights into the operation of the finance–real interaction.

The 11 articles collected in this volume, all strictly refereed, represent original up-to-date research papers that reflect some of the latest developments in the area of wavelet applications for economics and finance.

In the first chapter James provides a personal retrospective of a decade’s research that highlights the links between CWT, DWT wavelets and the more classical Fourier transforms and series. After stressing the importance of analyzing the various basis spaces, the exposition evaluates the alternative bases available to wavelet researchers and stresses the comparative advantage of wavelets relative to the alternatives considered. The appropriate choice of class of function, e.g., Haar, Morlet, Daubchies, etc., with rescaling and translation provide appropriate bases in the synthesis to yield parsimonious approximations to the original time or space series.

The remaining papers examine a wide variety of applications in economics and finance that reveal more complex relationships in economic and financial time series and help to shed light on various puzzles that emerged in the literature since long; on business cycles, traded assets, foreign exchange rates, credit markets, forecasting, and labor market research. Take, for example, the latter. Most economists agree that productivity increases welfare, but whether productivity also increases employment is still controversial. As economists have shown using data from the EU and the USA, productivity may rise, but employment may be de-linked from productivity increases. Recent work has shown that the analysis of the relationship between productivity and employment is one that can only properly be analyzed after decomposition by time scale. The variation in the short run is considerably different from the variation in the long run. In the chapter “Does Productivity Affect Unemployment? A Time-Frequency Analysis for the US”, Marco Gallegati, Mauro Gallegati, James B. Ramsey, and Willi Semmler, applying parametric and nonparametric approaches to US post-war data, conclude that productivity creates unemployment in the short and medium term, but employment in the long run.

The chapters "The Great Moderation Under the Microscope: Decomposition of Macroeconomic Cycles in US and UK Aggregate Demand" and "Nonlinear Dynamics and Wavelets for Business Cycle Analysis" contain articles using wavelets for business cycles analysis. In the paper by P.M. Crowley and A. Hughes Hallett the Great Moderation is analyzed employing both static and dynamic wavelet analysis using quarterly data for both the USA and the UK. Breaking the GDP components down into their frequency components they find that the "great moderation" shows up only at certain frequencies, and not in all components of real GDP. The article by P.M. Addo, M. Billio, and D. Guégan applies a signal modality analysis to detect the presence of determinism and nonlinearity in the US Industrial Production Index time series by using a complex Morlet wavelet.

The chapters "Measuring the Impact Intradaily Events Have on the Persistent Nature of Volatility" and "Wavelet Analysis and the Forward Premium Anomaly" deal with foreign exchange rates. In their paper M.J. Jensen and B. Whitcher measure the effect of intradaily events on the foreign exchange rates level of volatility and its well-documented long-memory behavior. Volatility exhibits the strong persistence of a long-memory process except for the brief period after a market surprise or unanticipated economic news announcement. M. Kiermeier studies the forward premium anomaly using the MODWT and estimate the relationship between forward and corresponding spot rates on foreign exchange markets on a scale-by-scale basis. The results show that the unbiasedness hypothesis cannot be rejected if the data is reconstructed using medium-term and long-term components.

Two papers analyzing the influence of several key traded assets on macroeconomics and portfolio behavior are included in the chapters "Oil Shocks and the Euro as an Optimum Currency Area" and "Wavelet-Based Correlation Analysis of the Key Traded Assets". L. Aguiar-Conraria, T.M. Rodrigues, and M.J. Joana Soares study the macroeconomic reaction of Euro countries to oil shocks after the adoption of the common currency. For some countries, e.g., Portugal, Ireland, and Belgium, the effects of an oil shock have become more asymmetric over the past decades. J. Baruník, E. Kočenda, and L. Vácha, in their paper, provide evidence for different dependence between gold, oil, and stocks at various investment horizons. Using wavelet-based correlation analysis they find a radical change in correlations after the 2007–2008 in terms of time-frequency behavior.

A surprising implication of the development of forecasting techniques to real and financial economic variables is the recognition that the results are strongly dependent on the analysis of scale. Only in the simplest of circumstances will forecasts based on traditional time series aggregates accurately reflect what is revealed by the time scale decomposition of the time series. The chapter "Forecasting via Wavelet Denoising: The Random Signal Case" by J. Bruzda presents a wavelet-based method of signal estimation for forecasting purposes based on wavelet shrinkage combined with the MODWT. The comparison of the random signal estimation with analogous methods relying on wavelet thresholding suggests that the proposed approach may be useful especially for short-term forecasting. Finally, the chapters

"Short and Long Term Growth Effects of Financial Crises" and "Measuring Risk Aversion Across Countries from the Consumption-CAPM: A Spectral Approach" contain two articles using the spectral approach. F.N.G. Andersson and P. Karpestam investigate to what extent financial crises can explain low growth rates in developing countries. Distinguishing between different sources of crises and separating short- and long-term growth effects of financial crises, they show that financial crises have reduced growth and that the policy decisions have caused them to be worsened and prolonged. In their paper E. Panopolou and S. Kalyvitis adopt a spectral approach to estimate the values of risk aversion over the frequency domain. Their findings suggest that at lower frequencies risk aversion falls substantially across countries, thus yielding in many cases reasonable values of the implied coefficient of risk aversion.

Contents

Contributors

Peter Martey Addo Université Paris 1 Panthéon-Sorbonne, Paris, France

Luís Aguiar-Conraria NIPE and Economics Department, University of Minho, Braga, Portugal

Fredrik N.G. Andersson Department of Economics, Lund University, Lund, Sweden

Jozef Baruník Institute of Economic Studies, Charles University, Prague, Czech Republic

Monica Billio Department of Economics, Università CaFoscari of Venice, Italy

Joanna Bruzda Department of Logistics Faculty of Economic Sciences and Management, Nicolaus Copernicus University, Torun, Poland

Patrick M. Crowley Economics Group, College of Business, Texas A&M University, Corpus Christi, TX, USA

Marco Gallegati Department of Economics and Social Sciences, Università Politecnica delle Marche, Ancona, Italy

Mauro Gallegati Department of Economics and Social Sciences, Università Politecnica delle Marche, Ancona, Italy

Dominique Guégan Université Paris 1 Panthéon-Sorbonne, Paris, France

Mark J. Jensen Federal Reserve Bank of Atlanta, Atlanta, GA, USA

Sarantis Kalyvitis Department of International and European Economic Studies, Athens University of Economics and Business, Athens, Greece

Peter Karpestam Department of Economics, Lund University, Lund, Sweden

Michaela M. Kiermeier Fachbereich Wirtschaft, Hochschule Darmstadt, Dieburg, Germany

Evžen Kočenda CERGE-EI, Charles University and the Czech Academy of Sciences, Prague, Czech Republic

Ekaterini Panopoulou Department of Statistics and Insurance Science, University of Piraeus, Athens, Greece and University of Kent, UK

James B. Ramsey Department of Economics, New York University, New York, NY, USA

Teresa Maria Rodrigues Economics Department, University of Minho, Braga, Portugal

Willi Semmler Department of Economics, New School for Social Research, New York, NY, USA

Maria Joana Soares NIPE and Department of Mathematics and Applications, University of Minho, Braga, Portugal

Lukáš Vácha Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague, Czech Republic

Brandon Whitcher Pfizer Worldwide Research & Development, Cambridge, MA, USA

Functional Representation, Approximation, Bases and Wavelets

James B. Ramsey

Abstract After stressing the importance of analyzing the various basis spaces, the exposition evaluates the alternative bases available to wavelet researchers. The next step is to demonstrate the impact of choice of basis for the representation or projection of the regressand. The similarity of formulating a basis is explored across a variety of alternative representations. This development is followed by a very brief overview of some articles using wavelet tools. The comparative advantage of wavelets relative to the alternatives considered is stressed.

1 Introduction

The paper begins with a review of the main features of wavelet analysis which are contrasted with other analytical procedures, mainly Fourier, splines, and linear regression analysis. A review of Crowley (2007), Percival and Walden (2000), Bruce and Gao (1996), the excellent review by Gençay et al. (2002), or the Palgrave entry for Wavelets by Ramsey (2010) before proceeding would be beneficial to the neophyte wavelet researcher.

The next section contains a non-rigorous development of the theory of wavelets and contains discussions of wavelet theory in contrast to the theory of Fourier series and splines. The third section discusses succinctly the practical use of wavelets and compares alternative bases; the last section concludes.

Before proceeding the reader should note that all the approximating systems are characterized by the functions that provide the basis vectors, e.g. Sin($k\omega_t$), Cos($k\omega_t$) for Fourier series, or "t" for the monomials, e^{λ_i} for the exponentials, etc.

J.B. Ramsey (✉)
Department of Economics, New York University, New York, NY, USA
e-mail: james.ramsey@nyu.edu

M. Gallegati and W. Semmler (eds.), *Wavelet Applications in Economics and Finance*, Dynamic Modeling and Econometrics in Economics and Finance 20,
DOI 10.1007/978-3-319-07061-2_1,

For a regular regression framework, the basis is the standard Euclidean space, E_N. For the Fourier projections we have the frequency scaled sine and cosine functions that produce a basis of infinite power, high resolution in the frequency domain, but no resolution in the time domain; e.g.

$$Re^{i2\pi ft} \quad \text{or alternatively expressed:}$$

$$1, sin(k\omega t), cos(k\omega t) \quad k = 1, 2, 3, \ldots$$

are highly differential, but are not suitable for analyzing signals with discrete changes and discontinuities.

The basis functions for splines are polynomials that are also differential and are defined over a grid determined by the knots; various choices for the differentiability at the knots determine the flexibility and smoothness of the spline approximation and the degree of curvature between knots.

Obviously, the analysis of any signal involves choosing both the approximating function and the appropriate basis vectors generated from the chosen function.

The concepts of "projection" and analysis of a function are distinguished; for the former one considers the optimal manner in which an N dimensional basis space can be projected onto a K dimensional subspace.

For a given level of approximation one seeks the smallest K for the transformed basis. Alternatively, a given function can be approximated by a series expansion, which implies that one is assuming that the function lies in a space defined in turn by a given class of functions, usually defined to be a Hilbert space. Projection and representation of a function are distinguished.

2 Functional Representation and Basis Spaces

2.1 An Overview of Bases in Regression Analysis

Relationships between economic variables are characterized by three universal components. Either the variable is a functional defined by an economic equation as a function of itself lagged to its past, i.e. is autoregressive; or is a function of time, i.e. is a "time series"; or it is a projection onto the space spanned by a set of functions, labeled, "regressors" each of which in turn may be autoregressive, or a vector of "time series". The projection of the regressand on the space spanned by the regressors provides a relationship between the variables, which is invariant to permutations of the indexing of the variables:

$$Y = X\beta + u$$

$$Y_{perm} = X_{perm}\beta + u_{perm}$$

where Y is the regressand, Y_{perm} the permuted values of Y, X_{perm}, represents a conformable permutation of the rows of X, and u_{perm}, a conformable permutation to Y_{perm}. However, if the formulation of the model involves an "ordering" of the variables over space, or over time, the model is then not invariant to permutation of the index of the ordering. It is known, but seldom recognized as a limitation of the projection approach, that least squares approximations are invariant to any permutation of the ordering. Consequently, the projection approach omits the information within the ordering in the space spanned by the residuals, which is, of course, the null space. Another distinguishing characteristic is that added to the functional development of the variable known as the "regressand" is an unobserved random variable, "u", which may be represented by a solitary pulse, or may have a more involved stochastic structure. In the former case, the regressand vector is contained in the space spanned by the regressors, where as in the latter case the regressand is projected onto the space spanned by the designated regressors.

The usual practice is to represent regressors and the regressand in terms of the standard Euclidean N dimensional space; i.e. the i_{th} component of the basis vector is "1", the remaining entries are zero; in this formulation, we can interpret the observed terms, x_i, y_i, $i = 1, 2, \ldots k$; as N dimensional vectors relative to the linear basis space, E_N.

The key question the analyst needs to resolve is to derive an appropriate procedure for determining reasonable values for the unknown parameters and coefficients of the system; i.e. estimation of coefficients and forecasting of declared regressands. Finally, if the postulated relationship is presumed to vary over space or time, special care will be needed to incorporate those changes in the relationship over time or over the sample space.

Consider as a first example, a simple non-linear differentiable function of a single variable x, $f(x|\theta)$, which can be approximated by a Taylor's series expansion about the point $a_1$in powers of x:

$$y = f(x|\theta) = f(a_1|\theta) + f^1(a_1|\theta)(x - a_1) + \frac{f^2(a_1|\theta)(x - a_1)^2}{2!} \tag{1}$$
$$+ \frac{f^3(a_1|\theta)(x - a_1)^3}{3!} + \frac{R(\xi)}{4!}$$

for some ξ value. This equation approximately represents the variation of y in terms of powers of x. Care must be taken in that the derived relationship is not exact, as the required value for ξ in the remainder term will vary for different values for a_1, x, and the highest derivative used in the expansion. Under the assumption that $R(\xi)$ is approximately zero, the parameters θ given the coefficient a_1 can be estimated by least squares using N independent drawings on the regressand's error term. Assuming the regressors are observed error free; one has:

$$\min_{\theta}\{\Sigma_1^N (y_i - f(x_i|\theta))^2\} \tag{2}$$

A single observation, i, on this simple system is:

$$y_i \ \{x_i, x_i^2 \ x_i^3\} \tag{3}$$

$$i = 1, 2, 3 \ldots . N. \tag{4}$$

This model is easily extended to differential functions which are themselves functions of multivariate regressors. The key aspect of the above formulation is that the estimators are obtained by a projection onto the space spanned by the regressors. Other, perhaps more suitable spaces, can be used instead. The optimal choice for a basis, as we shall see, is one that reduces significantly the required number of coefficients to represent the function y_i with respect to the chosen basis space. Different choices for the basis will yield different parameterizations; the research analyst is interested in minimizing the number of coefficients, actually the dimension of the supporting basis space.

2.2 *Mononomial Basis*

An alternative, ancient, procedure is provided by the monomials:

$$\{1, t^1, t^2, t^3, t^4, t^5, \ldots\} \tag{5}$$

that is, we consider the projection of a vector y on the space spanned by the monomials, $t^0, t^1, \ldots, t^k$, or as became popular as a calculation saving device, one considers the projection of y on the orthogonal components of the sequence in Eq. (5), see Kendall and Stuart (1961).

These first two procedures indicate that the underlying concept was that insight would be gained if the projections yielded approximations that could be specified in terms of very few estimated coefficients. Further very little structure was imposed on the model, either in terms of the statistical properties of the model or in terms of the restrictions implied by the underlying theory.

Two other simple basis spaces are the exponential

$$\{e^{\lambda_1 t}, \ e^{\lambda_2 t}, \ e^{\lambda_3 t} \ldots ..e^{\lambda_k t}\} \tag{6}$$

and the power base:

$$\{t^{\lambda_1}, \ t^{\lambda_2}, \ t^{\lambda_3}, \ \ldots \ldots .t^{\lambda_k}\}. \tag{7}$$

The former is most useful in modeling differential equations, the latter in modeling difference equations.

2.3 Spline Bases

A versatile basis class is defined by the spline functions. A standard definition of a version of the spline basis, the B-spline, $S_B(t)$, is:

$$S_B(t) = \sum_{k=1}^{m+L-1} c_k B_{k,(t,\tau)} \tag{8}$$

where $S_B(t)$ is the spline approximation, c_k, are the coefficients of the projection, $B_{k,(t,\tau)}$ is the B-spline function at position k, with knot structure, τ. The vector τ designates the number of knots, L, and their position which defines the subintervals that are modeled in terms of polynomials of degree m. At each knot the polynomials are constrained to be equal in value for polynomials of degree 1, agreement for the first derivative for polynomials of degree 2, etc. Consequently, adjacent spline polynomials line up smoothly.

B-Splines are one of the most flexible basis systems, so that it can easily fit locally complex functions. An important use of splines is to interpolate over the grid created by the knots in order to generate a differential function, or more generally, a differential surface. Smoothing is a local phenomenon.

2.4 Fourier Bases

The next procedure in terms of longevity of use is Fourier analysis. The basis for the space spanned by Fourier coefficients is given by:

$$1, \sin(k\omega t), \cos(k\omega t), \tag{9}$$

$$k = 1, 2, 3, \ldots$$

$$i.e. \exp(i\omega_k) \tag{10}$$

where ω is the fundamental frequency. The approximating sequences are given most simply by:

$$y = f(t) \cong \sum_{k=1}^{K} c_k \phi_k \tag{11}$$

where the sequence c_k specifies the coefficients chosen to minimize the squared errors between the observed sequence and the known functions shown in Eq. (11), ϕ_k is the basis function as used in Eq. (9), and the coefficients are given by

$$c_k = \int f(t)\phi_k(t)\, dt \tag{12}$$

The implied relationships between the basis function, ϕ, the basis space given by ϕ_k, $k = 1, 2, 3, ..$, and representation of the function $f(t)$, are given in abstract form in Eq. (12), in order to emphasize the similarities between the various basis spaces.

We note two important aspects of this equation. We gain in understanding if the number of coefficients are few in number; i.e. k is "small". We gain if the function "f" is restricted to functions of a class that can be described in terms of the superposition of the basis functions, e.g. trigonometric functions and their derivatives for Fourier analysis. The fit for functions that are continuous, but not every where differential, can only be approximated using many basis functions. The equations generating the basis functions, ϕ_k, based on the fundamental frequency, ω, are re-scaled versions of that fundamental frequency. The concept of re-scaling a "fundamental" function to provide a basis will occur in many guises.

Fourier series are useful in fitting global variation, but respond to local variation only at very high frequencies thereby substantially increasing the required number of Fourier coefficients to achieve a given level of approximation For example, consider fitting a Fourier basis to a "Box function", any reasonable degree of fit will require very many terms at high frequency at the points of discontinuity (see Bloomfield 1976; Korner 1988).

Economy of coefficients can be obtained for local fitting by using windows, that is, instead of

$$\hat{h}(\omega) = \frac{1}{2\pi} \sum_{-\infty}^{\infty} \hat{R}(s) \cos(s\omega)$$

where $\hat{R}(s)$ is the sample covariance at lag "s", we consider

$$\hat{h}(\omega) = \frac{1}{2\pi} \sum_{-M}^{M} \lambda(s) \hat{R}(s) \cos(s\omega) \tag{13}$$

where $\lambda(s)$ is the "window function" which has maximum effect at $s = 0, +/- 2k\pi$, $k = 1, 2, 3, \ldots$. Distant correlations are smoothed, the oscillations of local events are enhanced (see Bloomfield 1976).

A more precise formulation is provided by stating that for the function "f" defined for a mapping from the real line modulo 2π to *R,* the Fourier coefficients of "f" are given by:

$$\begin{aligned} \hat{f}(r) &= (2\pi)^{-1} \int_0^{2\pi} f(t) \exp(-irt)\, dt \\ &= (2\pi)^{-1} \int_T f(t) \exp(-irt)\, dt \end{aligned} \tag{14}$$

For simple functions we have the approximation (Korner 1988):

$$S_n(f,t) = \sum_{-n}^{n} \hat{f}(r) \exp irt \longrightarrow f(t) \tag{15}$$
$$as\, n \longrightarrow \infty$$

$T \longrightarrow C$ is continuous everywhere and has a continuous bounded derivative except at a finite number of points, then $S_n(f,\ ) \longrightarrow f$ uniformly (Korner 1988). The problem we have to face is the behavior of the function at points of discontinuity and to be aware of the difficulties imposed by even a finite number of discontinuities. For example, consider:

$$h(x) = x, \quad -\pi < x < \pi \tag{16}$$
$$h(\pi) = 0$$
$$\hat{h}(0) = 0$$

As pointed out by Korner (1988), the difficulty is due to the confusion between "the limit of the graphs and the graph of the limit of the sum". This insight was presented by Gibbs and illustrated practically by Michelson; that is $\mathrm{S}_n(h,t) \longrightarrow h(t)$ pointwise; that is the blips move towards the discontinuity but pointwise convergence of f_n to f does not imply that the graph of f_n starts to look like f for large N shown in (16). The important point to remember is that the difference is bounded from below in this instance by:

$$\frac{2}{\pi}\int_0^{\pi} \frac{\sin x}{x}\, dx > 1.17 \tag{17}$$

The main lesson here for the econometrician is that observed data may well contain apparently continuous functions that are not only sampled at discrete intervals, but that may in fact contain significant discontinuities. Indeed, one may well face the problem of estimating a continuous function that is nowhere differential,the so called "Weierstrass functions" (see, for example, Korner 1988).

It is useful to note that, whether we are examining wavelets (to be defined below), or sinusoids or Gabor functions, we are in fact approximating $f(t)$ by "atoms".[1] We seek to obtain the best M atoms for a given $f(t)$ out of a dictionary of P atoms. There are three standard methods for choosing the M atoms in this over sampled situation. The first is "matching pursuit" in which the M atoms are chosen one at a time; this procedure is referred to as greedy and sub-optimal (see Bruce and Gao 1996). An alternative method is the best basis algorithm which begins with a

[1] A collection of atoms is a "dictionary".

dictionary of bases. The third method, which will be discussed in the next section, is known as basis pursuit, where the dictionary is still over complete. The synthesis of $f(t)$ in terms of $\phi_i(t)$ is under-determined.

This brief discussion indicates that the essential objective is to choose a good basis. A good basis depends upon the resolution of two characteristics; *linear independence* and *completeness*. Independence ensures uniqueness of representation and completeness ensures that any $f(t)$ within a given class of functions can be represented in terms of the basis vectors. Adding vectors will destroy independence, removing vectors will destroy completeness. Every vector v or function $v(t)$ can be represented uniquely as:

$$v = \sum b_i v_i \quad (18)$$

$$or$$

$$v(t) = \sum b_i v_i(t)$$

provided the coefficients b_i satisfy:

$$A||v||^2 \leq \sum |b_i|^2 \leq B||v||^2 \; with \; A > 0. \quad (19)$$

This is the defining property of a Riesz basis (see, for example, Strang and Nguyen 1996).

If $0 < A < B$ and Eq. (19) holds and the basis generating functions are defined within a Hilbert space, then we have defined a frame and A, B are the frame bounds. If A equals B the bounds are said to be tight; if further the bounds are unity, i.e. $A = B = 1$, one has a orthonormal basis for the transformation. For example, consider a frame within a Hilbert space, C, given by: $e_1 = (0, 1)$, $e_2 = (-\frac{\sqrt{3}}{2} - \frac{1}{2})$, $e_3 = (-\frac{\sqrt{3}}{2}, \frac{1}{2})$. For any v in the Hilbert space we have:

$$\sum_{j=1}^{3} \left|\langle v, e_j \rangle\right|^2 = \frac{3}{2} \|v\|^2 \quad (20)$$

where the redundancy ratio is $3/2$, i.e. three vectors in a two dimensional space (Daubechies 1992).

2.5 *Wavelet Bases*

Much of the usefulness of wavelet analysis has to do with its flexibility in handling a variety of nonstationary signals. Indeed, as wavelets are constructed over finite intervals of time and are not necessarily homogeneous over time, they are localized in time and scale. The projection of the analizable signal onto the wavelet function by time scale and translation produces an orthonormal transformation matrix, W, such that the wavelet coefficients, w, are represented by:

$$w = Wx \tag{21}$$

where x is the analizable signal (see Eq. (21)). While theoretically this is a very useful relationship which clarifies the link between wavelet coefficients and the original data, it is decidedly not useful in reducing the complexity of the relationships and does not provide a suitable mechanism for evaluating the coefficients (Bruce and Gao 1996).

The experienced Waveletor knows also to consider the shape of the basis generating function and its properties at zero scale. This concern is an often missed aspect of wavelet analysis. Wavelet analysis, unlike Fourier analysis, can consider a wide array of generating functions. For example, if the function being examined is a linearly weighted sum of Gaussian functions, or of the second derivatives of Gaussian functions, then efficient results will be obtained by choosing the Gaussian function, or the second derivative of the Gaussian function in the latter case. This is a relatively under utilized aspect of wavelet analysis, which will be discussed more fully later.

Further any moderately experienced "Waveletor" knows to choose his wavelet generating function so as to maximize the "number of zero moments", to ascertain the number of continuous derivatives (as a measure of smoothness), and to worry about the symmetry of the underlying filters although one may consider models for which asymmetry in the wavelet generating function is appropriate. While many times the choice of wavelet generating function makes little or no difference there are times when such considerations are important for the analysis in hand. For example, the inappropriate use of the Haar function for resolving continuous smooth functions, or using smooth functions to represent samples of discontinuous paths. Wavelets provide a vast array of alternative wavelet generating functions, e.g. Gaussian, Gaussian first derivative, Mexican hat, the Daubechies series, the Mallat series, and so on. The key to the importance of the differences lies in choosing the appropriate degree and nature of the oscillation within the supports of the wavelet function. With the Gaussian, first, and second derivatives as exceptions, the generating functions are usually derived from applying a pair of filters to the data using subsampled data (Percival and Walden 2000).

I have previously stated that at each scale the essential operation is one of differencing using weighted sums; the alternative rescaleable wavelet functions provide an appropriate basis for such differences. Compare for example:

$$\begin{aligned} Haar \ : \ (h_0, \ h_1) &= (-\frac{1}{2}, \ \frac{1}{2}) \qquad (22) \\ Daubchies(D4) &= (h_0, \ h_1, \ h_2, \ h_3) \\ &= (\frac{1-\sqrt{3}}{4\sqrt{2}}, \ \frac{-3+\sqrt{3}}{4\sqrt{2}}, \ \frac{3+\sqrt{3}}{4\sqrt{2}}, \ \frac{-1-\sqrt{3}}{4\sqrt{2}}) \end{aligned}$$

The Haar transform is of width two, the Daubchies ($D4$) is of width 4. The Haar wavelet generates a sequence of paired differences at varying scales 2^j. In comparison, the Daubechies transform provides a "nonlinear differencing" over sets of four scaled elements, at scales 2^j.

Alternatively, wavelets can be generated by the conjunction of high and low pass filters, termed "filter banks" by Strang and Nguyen (1996) to produce pairs of functions $\Psi(t), \Phi(t)$ that with rescaling yield a basis for the analysis of a function f_t. Unlike the Fourier transform, which uses the sum of certain basis functions (sines and cosines) to represent a given function and may be seen as a decomposition on a frequency-by-frequency basis, the wavelet transform utilizes some elementary functions (father Φ and mother wavelets Ψ) that, being well-localized in both time and scale, provide a decomposition on a "scale-by-scale" basis as well as on a frequency basis. The inner product Φ with respect to f is essentially a low pass filter that produces a moving average; indeed we recognize the filter as a linear time-invariant operator. The corresponding wavelet filter is a high pass filter that produces moving differences (Strang and Nguyen 1996). Separately, the low pass and high pass filters are not invertible, but together they separate the signal into frequency bands, or octaves. Corresponding to the low pass filter there is a continuous time scaling function $\phi(t)$. Corresponding to the high pass filter is a wavelet $w(t)$.

For any set of filters that satisfy the following conditions

$$\sum_{l=0}^{L-1} h_l = 0 \tag{23}$$

$$\sum_{l=0}^{L-1} h_l^2 = 0 \tag{24}$$

$$\sum_{l=0}^{L-1} h_l h_{l+2n} = 0 \tag{25}$$

defines a wavelet function and so is both necessary and sufficient for the analysis of a function "f". However, this requirement is insufficient for defining the synthesis for a function "f". To achieve synthesis, one must add the constraint that:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\psi(\hat{\omega})|^2}{\omega} d\omega < \infty \tag{26}$$

see Chui (1992).

This gives wavelets a distinct advantage over a purely frequency domain analysis. Because Fourier analysis presumes that any sample is an independent drawing, Fourier analysis requires "covariance stationarity", whereas wavelet analysis may analyze both stationary and long term non-stationary signals. This approach provides a convenient way to represent complex signals. Expressed differently, spectral decomposition methods perform a global analysis, whereas wavelet methods act locally in both frequency and time. Fourier analysis can relax local non-stationarity by windowing the time series as was indicated above. The problem with this approach is that the efficacy of this approach depends critically on making the right choice of window and, more importantly, presuming its constancy over time.

Any pair of linear filters that meets the following criteria can represent a wavelet transformation (Percival and Walden 2000). Equation (23) gives the necessary conditions for an operator to be a wavelet: h_l denotes the high pass filter, and the corresponding low pass filter is given by:

$$\begin{aligned} g_l &= (-1)^{l+1} h_{L-l-1} \\ &\quad or \\ h_l &= (-1)^{l} g_{L-l-1} \end{aligned} \tag{27}$$

Equation (27) indicates that the filter bank depends on both the lowpass and high pass filters. Recall the high pass filter for the Daubechies $D(4)$, see Eq. (22), the corresponding low pass filter is:

$$g_0 = -h_3,\ g_1 = h_1,\ g_2 = -h_1,\ g_3 = h_0 \tag{28}$$

For wavelet analysis however, as we have observed, there are two basic wavelet functions, father and mother wavelets, $\phi(t)$ and $\psi(t)$. The former integrates to 1 and reconstructs the smooth part of the signal (low frequency), while the latter integrates to 0 and can capture all deviations from the trend The mother wavelets, as said above, play a role similar to sines and cosines in the Fourier decomposition. They are compressed or dilated, in the time domain, to generate cycles to fit actual data. The approximating wavelet functions $\phi_{J,k}(t)$ and $\psi_{J,k}(t)$ are generated from father and mother wavelets through scaling and translation as follows:

$$\phi_{J,k}(t) = 2^{-\frac{J}{2}} \phi\left(\frac{t - 2^J k}{2^J}\right) \tag{29}$$

and

$$\psi_{J,k}(t) = 2^{-\frac{J}{2}} \psi\left(\frac{t - 2^J k}{2^J}\right) \tag{30}$$

where j indexes the scale, so that 2^j is a measure of the scale, or width, of the functions (scale or dilation factor), and k indexes the translation, so that $2^j k$ is the translation parameter.

Given a signal $f(t)$, the wavelet series coefficients, representing the projections of the time series onto the basis generated by the chosen family of wavelets, are given by the following integrals:

$$\begin{aligned} d_{j,k} &= \int \psi_{j,k}(t) f(t)\, dt \\ s_{J,k} &= \int \phi_{J,k}(t) f(t)\, dt \end{aligned} \tag{31}$$