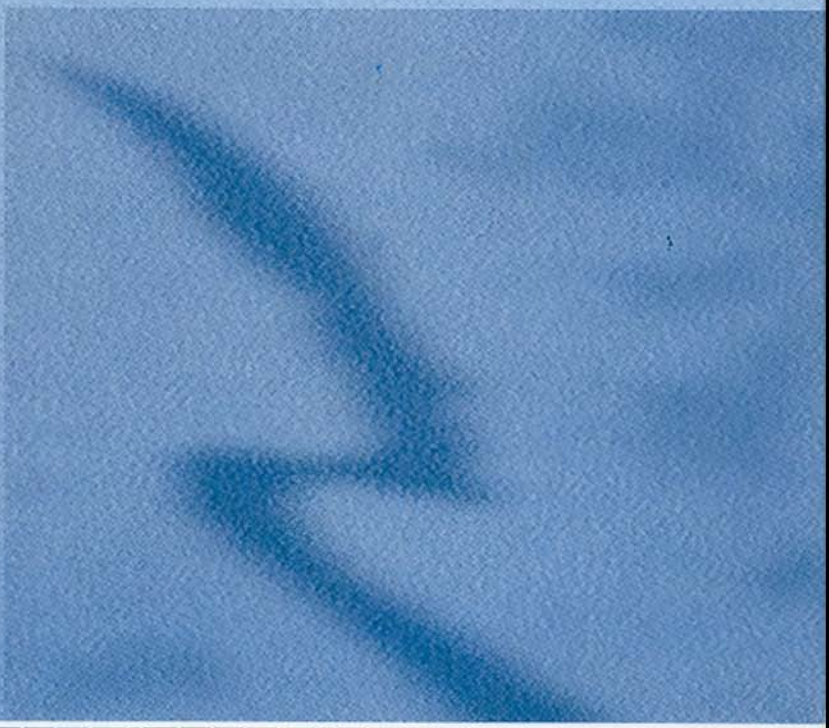
Financial Market Risk

Measurement and analysis

Cornelis A. Los

Routledge International Studies in Money and Banking

Routledge
Taylor & Francis Group



Financial Market Risk

What is financial market risk? How is it measured and analyzed? Is all financial market risk dangerous? If not, which risk is hedgeable?

These questions, and more, are answered in this comprehensive book written by **Cornelis A. Los**. The text covers such issues as:

- competing financial market hypotheses;
- degree of persistence of financial market risk;
- time–frequency and time–scale analysis of financial market risk;
- chaos and other nonunique equilibrium processes;
- consequences for term structure analysis.

This important book challenges the conventional statistical ergodicity paradigm of global financial market risk analysis. As such it will be of great interest to students, academics and researchers involved in financial economics, international finance and business. It will also appeal to professionals in international banking institutions.

Cornelis A. Los is Associate Professor of Finance at Kent State University, USA. In the past he has been a Senior Economist of the Federal Reserve Bank of New York and Nomura Research Institute (America), Inc., and Chief Economist of ING Bank, New York. He has also been a Professor of Finance at Nanyang Technological University in Singapore and at Adelaide and Deakin Universities in Australia. His PhD is from Columbia University in the City of New York.

Routledge International Studies in Money and Banking

1 Private Banking in Europe
Lynn Bicker

2 Bank Deregulation and Monetary Order
George Selgin

3 Money in Islam
A study in Islamic political economy
Masudul Alam Choudhury

4 The Future of European Financial Centres
Kirsten Bindemann

5 Payment Systems in Global Perspective
Maxwell J. Fry, Isaak Kilato, Sandra Roger, Krzysztof Senderowicz, David Sheppard, Francisco Soils and John Trundle

6 What is Money?
John Smithin

7 Finance
A characteristics approach
Edited by David Blake

8 Organisational Change and Retail Finance
An ethnographic perspective
Richard Harper, Dave Randall and Mark Rouncefield

9 The History of the Bundesbank
Lessons for the European Central Bank
Jakob de Haan

10 The Euro
A challenge and opportunity for financial markets
Published on behalf of *Société Universitaire Européenne de Recherches Financières (SUERF)*
Edited by Michael Artis, Axel Weber and Elizabeth Hennessy

11 Central Banking in Eastern Europe
Nigel Healey

12 Money, Credit and Price Stability
Paul Dalziel

13 Monetary Policy, Capital Flows and Exchange Rates
Essays in memory of Maxwell Fry
Edited by William Allen and David Dickinson

water flow changes like white noise (Whitcher et al.*, 2002). In other words, the Rhine river tends to produce no major catastrophic floods.*

The ACF of the fractionally-differenced time series can now be written in terms of the H-exponent, since we can now substitute $d = H - 0.5$ into the previously defined ACF to get:

$$\begin{aligned}\gamma(\tau) &= \frac{\sigma_\varepsilon^2 \Gamma(2-2H)\Gamma(\tau + H - 0.5)}{\Gamma(H-0.5)\Gamma(1.5-H)\Gamma(\tau+1.5-H)} \\ &\sim \sigma_\varepsilon^2 \tau^{2H-2} \quad \text{as } \tau \to \infty \end{aligned} \tag{4.48}$$

where $H \in (0, 1)$.

4.7 Critical color categorization of randomness

4.7.1 Blue, white, pink, red, brown and black noise

Following Schroeder (1991, pp. 121–137) we can now present a color categorization of randomness, or irregularity, by collecting the various descriptive exponents and relating them to each other. This comparison of exponents will facilitate the reading of a great variety of interdisciplinary research articles on phenomena of time dependence. In Chapter 8, we'll explain the intimate relationship between our concept of "randomness," as discussed in Chapter 1, and the concept of "irregularity" as defined by the mathematician Lipschitz.

Definition 173 *(1) When the Hurst exponent* $0 < H < 0.5$*, i.e.,* $-0.5 < d < 0$*, the time series of increments is called* antipersistent*. (2) When* $H = 0.5$*, i.e.,* $d = 0$*, the increments are independent or "white," and the time dependence of the series is* neutral *(or* neutrally persistent*). Examples are the increments of Random Walks or Arithmetic Brownian Motions (for speculative prices) and of GBM (for investment returns). The Brownian Motion series is once-integrated "white noise" and is called "brown" noise. Its ACF decays hyperbolically:*

$$\begin{aligned}\gamma(\tau) &= \frac{\sigma_\varepsilon^2 \Gamma(\tau)}{\Gamma(\tau+1)} \\ &= \frac{\sigma_\varepsilon^2 (\tau-1)!}{\tau!} \\ &= \sigma_\varepsilon^2 \tau^{-1} \end{aligned} \tag{4.49}$$

(3) When $0.5 < H < 1$*, i.e.,* $0 < d < 0.5$*, the time series of increments is called* persistent*.*

In the case of extreme antipersistence, $H \downarrow 0$, so that the ACF of the time series decays faster than hyperbolically in a quadratic fashion:

$$\begin{aligned}\gamma(\tau) &= \frac{\sigma_\varepsilon^2 \Gamma(\tau - 0.5)}{\Gamma(\tau + 1.5)} \\ &= \frac{\sigma_\varepsilon^2 (\tau - 1.5)!}{(\tau + 0.5)!} \\ &= \frac{\sigma_\varepsilon^2}{(\tau + 0.5)(\tau - 0.5)} \\ &= \frac{\sigma_\varepsilon^2}{(\tau^2 - 0.25)} \\ &\approx \sigma_\varepsilon^2 \tau^{-2} \quad \text{as } \tau \to \infty \end{aligned} \tag{4.50}$$

At the other extreme of Hurst's limited *randomness spectrum* $H \uparrow 1$, so that the ACF of the time series remains a flat constant and it never vanishes:

$$\begin{aligned}\gamma(\tau) &= \frac{\sigma_\varepsilon^2 \Gamma(\tau + 0.5)}{\Gamma(\tau + 0.5)} \\ &= \sigma_\varepsilon^2 \text{ a constant, as } \quad \tau \to \infty \end{aligned} \tag{4.51}$$

4.7.2 Irregularity exponents

We can make a connection with the stable distributions discussed earlier in Chapter 3, once we realize that, for globally (long-term) dependent time series, for which the autocovariance function has the form

$$\gamma(\tau) = \begin{cases} \tau^\lambda H(\tau) & \text{for } \lambda \in [-1, 0), \text{ or} \\ -\tau^\lambda H(\tau) & \text{for } \lambda \in (-2, -1] \end{cases} \tag{4.52}$$

as the time-interval lengthens, $\tau \to \infty$, and $H(\tau)$ is any slowly varying function at infinity, the *dependence exponent* λ equals

$$\begin{aligned}\lambda &= 2d - 1 \\ &= \upsilon - 1 \\ &= 2H - 2 \\ &= \frac{2}{\alpha_Z} - 2 \\ &= 2\alpha_L - 2 \end{aligned} \tag{4.53}$$

where d is the *difference (order) exponent*, υ is the *spectral exponent* (to be discussed in detail in Chapter 5), H is the aforementioned *Hurst exponent*, α_Z is the *stability exponent* of the Zolotarev parametrization of the stable distributions

Table 4.3 Equivalence of various critical irregularity exponents

Exponents: *Color:*	*Dependence* λ	*Difference* d	*Spectral* υ	*Hurst* H	*Stability* α_Z
Blue noise	$\lambda \downarrow -2$	$d = -0.5$	$\upsilon = -1$	$H \downarrow 0$	*NA*
Antipersistence	$-2 < \lambda < -1$	$-0.5 < d < 0$	$-1 < \upsilon < 0$	$0 < H < 0.5$	*NA*
White noise	$\lambda = -1$	$d = 0$	$\upsilon = 0$	$H = 0.5$	$\alpha_Z = 2$
Persistence (pink)	$0 > \lambda > -1$	$0 < d < 0.5$	$0 < \upsilon < 1$	$0.5 < H < 1$	$1 < \alpha_Z < 2$
Red noise	$\lambda \uparrow 0$	$d = 0.5$	$\upsilon = 1$	$H \uparrow 1$	$\alpha_Z = 1$
Brown noise	*NA*	$d = 1$	$\upsilon = 2$	*NA*	$\alpha_Z = 2/3$
Black noise	*NA*	$1 \leq d \leq 2$	$2 < \upsilon \leq 4$	*NA*	$2/5 \leq \alpha_Z < 2/3$

Note
NA = not applicable.

of Chapter 3, and α_L is the Lipschitz *regularity exponent* (to be discussed in Chapter 8).[10] Thus, the randomness, or irregularity, categorization can be expressed in terms of each of these *critical exponents*. For completeness of definition: $\lambda/2$ is the so-called *time-scaling exponent*.

The complete spectrum of randomness, or irregularity, in terms of the five critical exponents equivalent to the Lipschitz regularity exponent is given in Table 4.3, which provides the essential relationships between the exponents of the first difference of FBM (cf. also Keshner, 1982; Flandrin, 1989).

For example, for the Brownian Motion increments $\varepsilon(t)$, which are white noise:

$$\lambda = -1, \quad d = 0, \quad \upsilon = 0, \quad H = 0.5, \quad \alpha_Z = 2 \tag{4.54}$$

Thus, the time series of Brownian Motion increments is modelled by white noise:

$$\begin{aligned} x(t) &= (1 - L)^0 \varepsilon(t) \\ &= \varepsilon(t) \end{aligned} \tag{4.55}$$

Fractional integration of such white noise, when $d = 0.5$ and $H \uparrow 1$, results in a red noise series (Gilman *et al.*, 1963):

$$x(t) = (1 - L)^{-0.5} \varepsilon(t) \tag{4.56}$$

One complete integer integration of the white noise, when $d = 1$, results in a brown noise series (= Brownian Motion)

$$x(t) = (1 - L)^{-1} \varepsilon(t) \tag{4.57}$$

Visual samples of time series of such white, red and brown noise are given by Figure 4.5.

In the case of $0.5 < H < 1$, the vital property of the FBM is that the persistence of its increments extends forever: *it never dies out* and gives rise to the empirically observed *catastrophes*. The strength of such persistence is measured by the critical H-exponent.

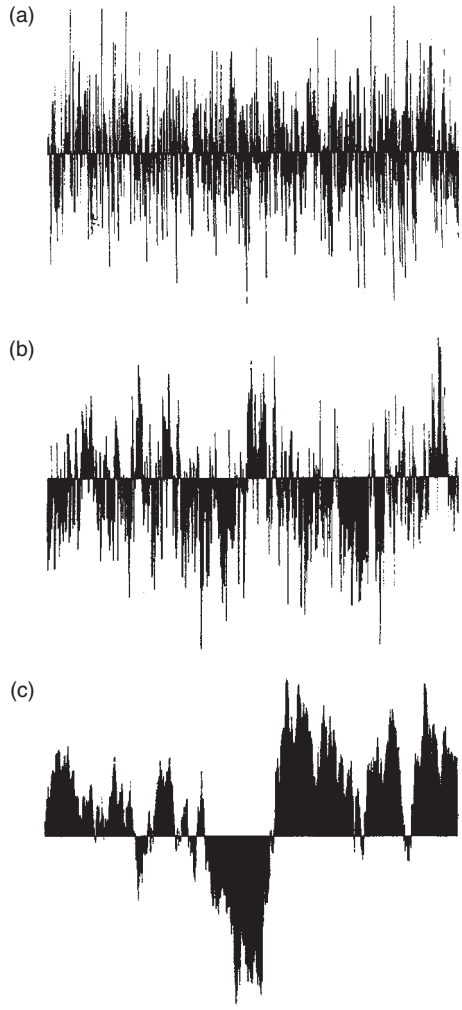


Figure 4.5 Sample of (a) white noise with $P(\omega) = \omega^{-0}$ power spectrum; (b) pink noise with $P(\omega) = \omega^{-1}$ power spectrum; and (c) brown noise with $P(\omega) = \omega^{-2}$ power spectrum.

Example 174 *The rates of return $x(t)$ of the S&P500 stock market index show mild persistence with $H = 0.67$. Indeed, their graph is less irregular than that of ordinary GBM increments. Its fractional dimension D is thus between the dimension of a line, $D = 1$, and the dimension of a plane, $D = 2$:*

$$1 < D = 2 - H = 1.33 < 2 \tag{4.58}$$

In Chapter 8 we'll discuss the fractional dimensions of financial time series in more detail. Curiously, the Dow Jones Industrials stock index shows neutral persistence, according to Li (1991), with $H = 0.5$.

Example 175 *The fractional dimension of GBM increments, with* $H = 0.5$*, is*

$$D = 2 - H = 1.5 = \frac{3}{2} \tag{4.59}$$

The case where $0.5 < d < 1.5$, or, equivalently, $1 < \upsilon < 3$, which cannot be measured directly by the *H*-exponent, but only after one differentiation, has been called the *infrared catastrophe* (Wornell and Oppenheim, 1992). It can be measured by the wavelet multiresolution analysis discussed in Chapter 8. More fractional integration, for example $d = 2$, results in heavily persistent, or pure *black noise*

$$x(t) = (1 - L)^{-2}\varepsilon(t) \tag{4.60}$$

As Schroeder (1991, p. 122) comments:

> Black-noise phenomena govern natural and unnatural catastrophes, like floods, droughts, bear markets, and various outrageous outages, such as those of electrical energy. Because of their black spectra, such disasters often come in clusters.

In contrast, the FBM increments with $0 < H < 0.5$ are antipersistent noise, hence they diffuse more quickly than the Brownian increments. Such FBM increments continuously return to the point they came from.

Remark 176 *Notably this means that the theoretical Random Walk innovations* $\varepsilon(t)$ *are rather exceptional. They exhibit the same stability,* $\alpha_Z = 2$*, and (in-)dependence,* $H = 0.5$*, as Gaussian random variables, but do not necessarily have to be Gaussian! Furthermore, their ACF drops off geometrically with* $\lambda = -1$*. By measuring the financial-economic, e.g., stock price innovations to be close to Gaussian, Granger and Morgenstern (1963) and Granger (1966) inferred that such innovations had a typical spectral shape. However, we'll learn in Chapters 6 and 7 that their inference was erroneous, and that there was nothing typical about that inferred shape, because it was biased by thinking exclusively in term of Gaussian innovations* $\varepsilon(t) \sim N(0, \sigma_\varepsilon^2)$*. For example, the covariance function of modern foreign exchange rates, like the Japanese Yen or the German Deutschemark, shows antipersistence, i.e., a slower drop-off of the ACF than the "typical" spectral shape based on this assumption of Gaussian i.i.d. innovations.*

4.7.3 Stability spectra

It is very important to understand that the Hurst exponent H is a rather limited measure of randomness and distributional stability with a very limited

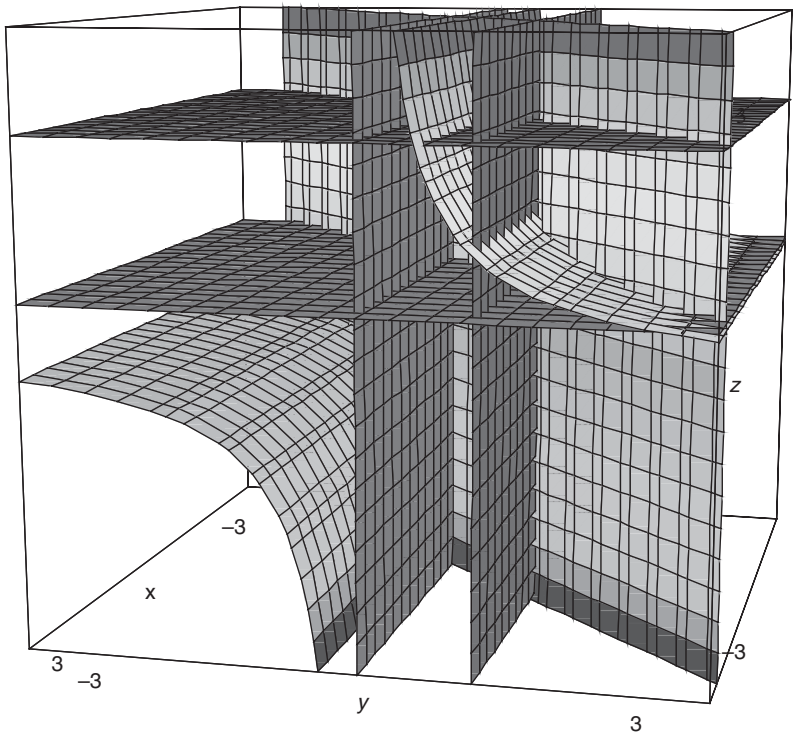


Figure 4.6 Relations between and constraints on d, H and α_Z. The axes measure $x = d$, $y = H$, $z = \alpha_Z$.

measurement domain, and that the α_Z-stability exponent, and the υ-spectral exponent have much more extensive measurement domains. This becomes clear, when we geometrically visualize the mathematical relationships, the constraints, and the respective domains of the various critical irregularity exponents in Figure 4.6.

The implied equality $\alpha_Z = 1/H$ does not hold for all values of α_Z, since the Hurst exponent, per definition, $0 < H < 1$, implies that $1 < \alpha_Z < \infty$, while the parametrized stable distributions of Chapter 3 are defined for the limited domain $0 < \alpha_Z \leq 2$. Apparently there exist empirical *ultra-stable* distributions (not yet parametrized!) in the domain $2 \leq \alpha_Z < \infty$, since we find *in extremo* $\alpha_Z \uparrow \infty$ when $H \downarrow 0$ (and $d \uparrow 0.5$), which is *complete stability*. These distributions are the distributions of *singularities*, or *singularity spectra*, which can be characterized and measured by the stability exponent α_Z. Considering that we have already empirically measured antipersistence in the FX markets, we will discuss such theoretical singularity spectra in Chapter 8.

As we recall from Chapter 3, and as is clearly visible in Figure 4.6, when the Hurst exponent vanishes, $H \downarrow 0$, the Zolotarev stability exponent becomes infinite,

$\alpha_Z \uparrow \infty$. In other words, for very small values of the Hurst exponent, $H \downarrow 0$, we acquire very uncertain measurements regarding Zolotarev's stability exponent α_Z.

In addition, there are now theoretically defined, parametrized stable distributions where $0 < \alpha_Z < 1$, which can also not be measured by the Hurst H-exponent directly, but can be measured by α_Z, if we can compute α_Z in some other fashion. These are the *ultra-unstable* distributions. However, empirically, there appears to be a physical turbulence barrier at $\alpha_Z = 2/5$. In other words, there appears not to exist any empirical α_Z such that $0 < \alpha_Z < 2/5$, even though there are theoretical Zolotarev-parametrized distributions defined for such α_Z values. Again, this is an area open for further theoretical and empirical research.

In conclusion, the best domain for using the H-exponent to compute the stability α_Z-exponent is in the Gaussian neighborhood of $H = 0.5$, where $\alpha_Z = 2$. Still, it is important to recognize that there exists a *stability spectrum* of randomness, or irregularity, completely specified by the stability exponent α_Z.

Remark 177 *Of course, one can still use the H-exponent for measuring infrared and black catastrophes, by measuring the H-exponent after proper integer-differentiation. For example, we hypothesize that* $x(t)$ *is pure black noise and has a spectral exponent* $\upsilon = 4$*, then differentiation of two full times* $(d = 2)$ *should theoretically result in white noise series with a flat spectrum,* $\upsilon = 0$*, so that* $H = 0.5$*. However, when we empirically measure, for example,* $H = 0.2 \rightarrow \upsilon = -0.6$*, then the original series must have a spectral coefficient of* $\upsilon = -0.6 + 4 = 3.4$ *and not* 4*.*

4.8 Software

Benoit 1.3: Fractal System Analysis (for Windows), Trusoft International Inc., 204, 37th Ave. N #133, St. Petersburg, FL 33704 Tel: (813) 925-8131; Fax: (813) 925-8141; sales@trusoft-international.com. See http://www.trusoft-international.com for details. This Benoit software enables you to measure the fractal dimension and/or Hölder–Hurst exponent of your data sets using your choice of method(s) for analysis of self-affine traces of speculative prices.

In the following Exercises you should use the Benoit software, Version 1.3. Once you've accessed Benoit, enlarge the working screen by the maximizing $ button in the upper-right corner of Benoit's initial screen, otherwise you will not see the crucial *OK* button. Always read the Benoit Help descriptions of the methods you use and relate them to the text of this chapter. To feed the empirical data as inputs into the Benoit program using EXCEL, read Benoit's Help instructions on Data Files (Data Formats).

4.9 Exercises

Exercise 178 *Compare the ACFs of the four data series of the Exercises of Chapter 1 against the theoretical benchmarks of Table 4.2. Are the series*

antipersistent, white noise, persistent (= pink) noise, red noise, brown noise or black noise, respectively? How can you tell?

Exercise 179 *Following Kasdin (1995), simulate three types of FBMs with Hurst exponents* $H = 0.2$, $H = 0.5$ *and* $H = 0.8$, *respectively, by using Benoit's self-affine trace generator. Compare your results with the one in this chapter. Generate no more than* 250 *points for each case. Use a vertical range of* 0–25. *Generate both the traces and their first differences. The Benoit program provides three methods for generating synthetic self-affine traces (the successive random addition method, the Fourier Transform method and the wavelet method). Try all three methods and describe in your own words how the results differ from each other. In total you should generate* $2 \times 3 \times 3 = 18$ *pictures. Save the traces and their first differences and display them either in an EXCEL spreadsheet or in Microsoft Power Point. These respective simulations will provide you with some "benchmark" pictures for the following Exercises.*

Exercise 180 *Compute the Hurst exponent and the fractal dimension of our S&P500 data of the Exercises of Chapter 1 using Hurst's Rescaled-Range (R/S) analysis for (1) the original share prices, (2) their total rates of return and (3) for first differences of the rates of total return. (Use the double logarithmic plot window.)*

Exercise 181 *Repeat the preceding Exercise, using the Irregularity (or Roughness)–Length method, which relates the standard deviation of windows of various length to the Hurst exponent, in the fashion described in this chapter. It plots the logarithm of the standard deviation (or Root-Mean-Squared (RMS) error) against the logarithm of the length of the window* τ.

Exercise 182 *Repeat the preceding Exercise, using the Variogram method, which is directly related to the autocovariance function.*

Notes

1 The current unorthodox efforts to characterize nonstationary financial-economic time series using more advanced signal processing technology are comparable with these early out-of-the-mainstream technical efforts by Granger and Morgenstern. For example, econometrician J. B. Ramsey of New York University performed the first wavelet multiresolution analysis (MRA) of macroeconomic data series (Ramsey, 1997).
2 Los (1999, 2000) provides some empirical examples of such "periodicity" for Asian FX markets, using non-parametric methods, based on high frequency data for 1997.
3 Cf. Los (1984) for theoretical discussions and Monte Carlo experiments with empirically estimated Kalman filters for econometric time-varying parameter models, including unstable ones.
4 We'll discuss in Chapter 12 possible insurance against such extreme catastrophic events, in the context of some dramatic hydrological and financial developments in mainland China.

5 Mandelbrot has questioned if Hosking's ARFIMA models were an improvement over his simpler fractionally differenced models, since such models with fractional exponents can trivially represent the integer exponent ARIMA models. But Hosking wanted to show the fractional and integer exponents separately within one modified framework, because they represent different phenomena: non-periodic and periodic cyclicity, respectively.
6 Cf. Meerschaert (1999) for the continuous time form of these long memory dynamic processes.
7 Such classical ACFs support the econometric measurements of Vector Auto-Regression (VARs) models. Classical VARs can represent higher order periodicities, but not the long-term time dependent phenomenon of non-periodic cyclicities, because they are expressed in terms of integer Markov processes. Of course, one can also, unconventionally, model fractional VARs to properly represent globally dependent or long memory processes.
8 And not the incorrect value of $H = 0.53$ provided by Mantegna and Stanley (2000, p. 86), who are proven wrong by their own figure 10.7, which we borrowed as our Figure 4.3.
9 Hölder (1859–1937) was a German mathematician, who devised treatment of divergent series of arithmetic summations, which led to a regularity exponent now recognized to be similar to Hurst's. However, Hölder was thinking about microscopic (physics) phenomena, in contrast to Hurst, who thought about macroscopic (hydrological) phenomena. The Hölder–Hurst exponents are also called critical Lipschitz irregularity exponents.
10 Somewhat confusingly presented in the literature, the Zolotarev stability $\alpha_Z = 1/\alpha_L$, where α_L is the Lipschitz regularity exponent. In the literature, one often finds just α and it is not always clear if the author(s) mean(s) the Zolotarev stability exponent α_Z or the Lipschitz α_L. We hope that this comparison of the various critical exponents and the presentation of their relationships will lift the dense fog between the various scientific disciplines, in particular in finance, physics and engineering, which deal with essentially the same signal processing phenomena.

Bibliography

Akgiray, V. (1989) "Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts," *Journal of Business*, **61**(1), 55–80.

Alexander, Carol (1998) "Volatility and Correlation: Measurement, Models and Applications," chapter 4 in Alexander, Carol (Ed.) (1999), *Risk Management and Analysis*, *Volume 1: Measuring and Modelling Financial Risk*, John Wiley and Sons, New York, NY, pp. 125–171.

Anderson, T. W. (1994) *The Statistical Analysis of Time Series*, John Wiley and Sons, New York, NY.

Baillie, R. T., and T. Bollerslev (1992) "Conditional Forecast Densities From Dynamic Models with GARCH Innovations," *Journal of Econometrics*, **52**(112), 91–113.

Bollerslev, Tim (1986) "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, **31**(3), 307–327.

Bollerslev, Tim (1987) "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics*, **9**(3), 542–547.

Bollerslev, Tim, Ray Y. Chou and Kenneth F. Kroner (1992) "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, **52**(112), 5–59.

Bollerslev, Tim, Robert F. Engle and Daniel B. Nelson (1994) "ARCH Models," chapter 49 in Engle, Robert F., and Dan L. McFadden (Eds) (1994), *Handbook of Econometrics*, Vol. 4, Elsevier-North Holland, Amsterdam.

Box, G. E. P., and G. M. Jenkins (1970) *Time Series Analysis: Forecasting and Control*, 2nd edn, Holden-Day, San Francisco, CA.

Elliott, Robert J., and John van den Hoek (2000) "A General Fractional White Noise Theory and Applications in Finance," *Quantitative Methods in Finance and Bernoulli Society 2000 Conference (Program, Abstracts and Papers)*, 5–8 December 2000, University of Technology, Sydney, pp. 327–345.

Engle, Robert F. (1982) "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, **50**(4), 987–1002.

Engle, Robert F. (Ed.) (1995) *ARCH: Selected Readings*, Advanced Texts in Econometrics, Oxford University Press, Oxford, UK.

Flandrin, Patrick (1989) "On the Spectrum of Fractional Brownian Motion," *IEEE Transactions on Information Theory*, **35**(1), 197–199.

Gilman, D. L., F. J. Fuglister and J. M. Mitchell, Jr (1963) "On the Power Spectrum of 'Red Noise'," *Journal of Atmospheric Science*, **20**(2), 182–184.

Granger, C. W. J. (1966) "The Typical Spectral Shape of an Economic Variable," *Econometrica*, **34**(1), 150–161.

Granger, C. W. J., and O. Morgenstern (1963) "Spectral Analysis of New York Stock Exchange Prices," *Kyklos*, **16**, 1–27. Reprinted as chapter 8 in Cootner, Paul H. (Ed.) (1964), *The Random Character of Stock Market Prices*, The MIT Press, Cambridge, MA, pp. 162–188.

Hosking, J. R. M. (1981) "Fractional Differencing," *Biometrika*, **68**(1), 165–176.

Hurst, H. E. (1951) "Long-term Storage Capacity of Reservoirs," *Transactions of the American Society of Civil Engineers*, **116**, 770–799.

Kasdin, N. Jeremy (1995) "Discrete Simulation of Colored Noise and Stochastic Processes and $1/f^{\alpha}$ Power Law Noise Generation," *Proceedings of the IEEE*, **83**(5), 802–827.

Keshner, M. S. (1982) "1/f Noise," *Proceedings of the IEEE*, **70**, 212–218.

Li, W. (1991) "Absence of $1/f$ Spectra in Dow Jones Daily Average," *International Journal of Bifurcations and Chaos*, **1**, 583–597.

Lo, Andrew W., and A. Craig Mackinlay (1988) "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Tests," *The Review of Financial Studies*, **1**(1), 41–66.

Lo, Andrew W., and A. Craig MacKinlay (1999) "Long-Term Memory in Stock Market Prices," chapter 6 in *A Non-Random Walk Down Wall Street*, Princeton University Press, Princeton, NJ, pp. 147–184.

Los, Cornelis A. (1984) *Econometrics of Models with Evolutionary Parameter Structures*, PhD dissertation, Columbia University, University Microfilms International, Ann Arbor, MI, January.

Los, Cornelis A. (1999) "Nonparametric Testing of the High-Frequency Efficiency of the 1997 Asian Foreign Exchange Markets," *Journal of Multinational Financial Management*, **9**, 3 and 4 October, 265–289.

Los, Cornelis A. (2000) "Nonparametric Efficiency Testing of Asian Stock Markets, Using Weekly Data," in Fomby, Thomas B. and R. Carter Hill (Eds), *Advances in Econometrics: Applying Kernel and Nonparametric Estimation to Economic Topics*, Vol. 14, JAI Press, pp. 329–363.

Mandelbrot, Benoit B. (1965) "Une Classe de Processus Stochastiques Homothétique à Soi; Application à la Loi Climatologique de H. E. Hurst" ("A Class of Self-Similar Stochastic

Processes; Application to the Climatological Law of H. E. Hurst"), *Comptes Rendus (Paris)*, **260**, 3274–3277.

Mandelbrot, Benoit B. (1971) "Limitations of Efficiency and of Martingale Models," *The Review of Economics and Statistics*, **53**(3), 225–236.

Mandelbrot, Benoit B. (1972) "Statistical Methodology for Nonperiodic Cycles: From the Covariance to the R/S Analysis," *Annals of Economic and Social Measurement*, **1**(3), 259–290.

Mandelbrot, Benoit B., and John W. Van Ness (1968) "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review*, **10**(4), 422–437.

Mandelbrot, Benoit B., and J. R. Wallis (1969) "Some Long-Run Properties of Geophysical Records," *Water Resources Research*, **5**(2), 321–340.

Mandelbrot, Benoit B., and M. Taqqu (1979) "Robust R/S Analysis of Long Run Serial Correlation," *Bulletin of the International Statistical Institute*, **48**(2), 59–104.

Mantegna, Rosario, N., and H. Eugene Stanley (2000) *An Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press, Cambridge, UK.

Meerchaert, Mark (1999) "Fractional Diffusion," in *Proceedings of Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics*, American University, Washington DC, 3–5 June 1999.

Moore, Geoffrey H. (1980) *Business Cycles, Inflation, and Forecasting*, published for the National Bureau of Economic Research, Ballinger Publishing Co., Cambridge, MA.

Nelson, Daniel B. (1991) "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, **59**(2), 347–370.

Peters, Edgar E. (1992) "R/S Analysis Using Logarithmic Returns: A Technical Note," *Financial Analysts Journal*, **48**(6), 81–82.

Peters, Edgar E. (1994) *Fractal Market Analysis*, John Wiley and Sons, New York, NY.

Ramsey, J. B. (1997) "The Decomposition of Economic Relationships by Time Scale Using Wavelets: Expenditure and Income," *Studies in Nonlinear Dynamics and Econometrics*, **3**(1), 23–42.

Schroeder, Manfred (1991) "Noises: White, Pink, Brown, and Black," in *Fractals, Chaos, Power Laws*, W. H. Freeman and Co., New York, NY, pp. 121–137.

Whitcher, B., S. D. Byers, P. Gultrop and D. B. Percival (2002) "Testing the Homogeneity of Variance in Time Series: Long Memory, Wavelets and the Nile River," *Water Resources Research*, **38**(5), 1029–1039.

Wornell, G. W., and A. V. Oppenheim (1992) "Estimation of Fractal Signals from Noisy Measurements Using Wavelets," *IEEE Transactions of Signal Processing*, **40**(3), 611–623.

Part II

Financial risk measurement

5 Frequency analysis of financial risk

5.1 Introduction

In Chapter 3 we analyzed the marginal distribution and in Chapter 4 the temporal dependence of investment returns identified as Fractional Brownian Motion (FBM). We'll now prepare to look at these two aspects of the research problem to characterize the long-term temporal risks of such returns simultaneously in their frequency and time domains. As we discussed in the preceding four chapters, Geometric Brownian Motion (GBM) increments are, per definition, independent and stationary (i.i.d.). Their stationarity allows for Fourier analysis, i.e., linear analysis in the frequency domain, since their (co-) variances and therefore their (co-) frequencies are constant. These increment series can be expanded in series of scaling frequencies, the so-called *frequency spectra*. In the next chapter, when we discuss windowed Fourier analysis, we'll determine how the FBM scaling frequency spectra depend on, and vary through, time.

In particular, in this chapter we'll discuss first covariance or correlation functions, which measure the degree of linear dependence, time and frequency convolution, and Fourier (= frequency or spectral) analysis. This chapter will prepare us for the next chapter, where we will visualize nonlinear dependence measurements in the time-frequency domain using Gábor's spectrograms based on the Windowed Fourier Transform.

Our original financial-economic inspiration originated with the work by Granger and Morgenstern (1963), Granger (1966) and Priestley (1981).[1] For the following mathematical details of measuring the time-dependence of varying frequencies we are indebted to Bloomfield (1976), Hsu (1984), Champeney (1990), Nikias and Petropulu (1993) and Körner (1990). Additional and more recent examples of applied Fourier analysis can be found in Folland (1992).

5.2 Visualization of long-term financial risks

5.2.1 Plot of absolute ACF against time horizons

Because of time-reversals and reversals-to-the-mean in time series, it is difficult to detect long-term dependence and geometric scaling laws by just plotting

the Autocorrelation Function (ACF) $\gamma(\tau)$, of first differences of the returns on investments

$$\Delta x(t) = \Delta[\ln X(t) - \ln X(t-1)] \tag{5.1}$$

or of first differences of foreign exchange (FX) rates

$$\Delta X(t) = X(t) - X(t-1) \tag{5.2}$$

The reason is that it is difficult to visually distinguish between the decays of the various ACFs corresponding to various fractional difference constants d, as was originally suggested by Box and Jenkins (1970). An only mildly better identifying picture of these slow geometric declines in dependence, which are indicative for long-term dependence, are plots of the absolute values of the ACFs, $|\gamma(\tau)|$, or of their squared values, $|\gamma(\tau)|^2$, against the time horizons τ, as in Figure 4.4 of Chapter 4.

5.2.2 Time-frequency and time-scale visualizations

In addition, there is a more fundamental problem with this particular identification methodology. The ACF, $\gamma(\tau)$, provides only second-order evidence for linear dependence, i.e., evidence for *weak* linear dependence (cf. Chapters 1 and 3). Following up on our conjecture at the end of Chapter 2, that financial risk involves more than just the second-order moments of variance and covariance of financial variables, we prefer to visualize the shape of their complete distributions, in particular, by way of their third- and fourth-order moments – their skewness and kurtosis. This is similar to what we did for the stable distributions discussed in Chapter 3. Moreover, we would like to visualize complete distributional evidence for all relative frequencies of occurrence for a time series for any form of time dependence, and not only for its correlation, i.e., for its linear dependence.

Thus, our preferred methodology should be to simultaneously visualize the marginal distributional evidence *and* the time-localized dependence evidence of these nonstationary time series, so that we can also better distinguish between serial (short term) time-dependence and global (long term) time-dependence. Such visualization and identification methodology exists already for more than half a century and it is very familiar to signal processing engineers.

In short, we need to analyze time-frequency pictures, or *spectrograms*, of financial time series, as discussed in Chapter 6, and time-scale pictures, or *scalograms*, as discussed in Chapter 7 to enable the required, and proper, identification of financial market risk.

5.3 Correlation and time convolution

In this section, we'll define the ACF of classical time series analysis (Jenkins and Watts, 1968; Box and Jenkins, 1970; Anderson, 1994) and establish its close

relationship with the convolution of signal processing. It's important to recognize that all correlation functions, convolutions and Fourier Transforms discussed in this chapter are scalar products, based on simple integrations.

Definition 183 *The* scalar product*, or inner product, on the space of* $L^2[a, b]$ *of square integrable functions is defined by*

$$\langle x, y \rangle = \int_a^b x(t)y(t)\,dt \tag{5.3}$$

A scalar product of two vectors and a simple integral are essentially the same thing. The two operations exactly coincide in the following situation (Burke-Hubbard, 1998, pp. 159–160). The step functions $f(t)$ are defined for $0 \leq t \leq T$ and are constant, except possibly at the integers. For example, imagine a step function $\bar{Y}(t)$ representing an average quantity of a commodity in time period t and another step function $\bar{X}(t)$ representing the average price of that commodity in time period t. Then, the integral

$$\int_a^b \bar{X}(t)\bar{Y}(t)\,dt \tag{5.4}$$

gives the total revenue of the sale of this commodity over the period $(b - a)$. But the same information is given by the scalar product

$$\begin{aligned}\langle \bar{X}, \bar{Y} \rangle &= \left\langle \begin{bmatrix} \bar{X}_a \\ \vdots \\ \bar{X}_b \end{bmatrix}, \begin{bmatrix} \bar{Y}_a \\ \vdots \\ \bar{Y}_b \end{bmatrix} \right\rangle \\ &= \bar{X}_a\bar{Y}_a + \cdots + \bar{X}_b\bar{Y}_b \\ &= \sum_{t=a}^{b} \bar{X}(t)\bar{Y}(t) \end{aligned} \tag{5.5}$$

In a 2-Dimensional (2D) price–quantity diagram, the scalar product $\langle \bar{X}, \bar{Y} \rangle$ would represent the space under the average price curve.

5.3.1 Covariance functions

Let's now use this scalar product to specify the covariance and correlation functions.

Definition 184 *The* cross-covariance function *of two (random) variables* $x_1(t)$ *and* $x_2(t)$ *is*

$$\begin{aligned} R_{12}(\tau) &= E\{x_1(t)x_2(t-\tau)\} \\ &= \int_{-\infty}^{+\infty} x_1(t)x_2(t-\tau)\,dt \\ &= \int_{-\infty}^{+\infty} x_1(t)L^{\tau}x_2(t)\,dt \end{aligned} \tag{5.6}$$

Of course, we also have

$$\begin{aligned} R_{21}(\tau) &= E\{x_2(t)x_1(t-\tau)\} \\ &= \int_{-\infty}^{+\infty} x_2(t)x_1(t-\tau)\,dt \\ &= \int_{-\infty}^{+\infty} x_2(t)L^{\tau}x_1(t)\,dt \end{aligned} \tag{5.7}$$

where L is the familiar linear lag operator.

The cross-covariance function $R_{12}(\tau)$, or $R_{21}(\tau)$, provides a measure of linear similarity, or linear dependence, between the variables $x_1(t)$ and $x_2(t)$ as a function of the parameter τ, the time shift of one variable with respect to the other. If the cross-covariance function is zero for all time shifts τ, then the two variables are said to be uncorrelated.

Definition 185 *If the (random) variables* $x_1(t)$ *and* $x_2(t)$ *are identical, the covariance function*

$$\begin{aligned} R_{11}(\tau) &= \int_{-\infty}^{+\infty} x_1(t)x_1(t-\tau)\,dt \\ &= \int_{-\infty}^{+\infty} x_1(t)L^{\tau}x_1(t)\,dt \end{aligned} \tag{5.8}$$

is called the autocovariance function *of* $x_1(t)$.

Definition 186 *The normalized quantity* $\gamma(t)$ *defined by*

$$\begin{aligned} \gamma(\tau) &= \frac{\int_{-\infty}^{+\infty} x_1(t)x_1(t-\tau)\,dt}{\int_{-\infty}^{+\infty}[x_1(t)]^2\,dt} \\ &= \frac{\int_{-\infty}^{+\infty} x_1(t)L^{\tau}x_1(t)\,dt}{\int_{-\infty}^{+\infty}[x_1(t)]^2\,dt} \end{aligned} \tag{5.9}$$

is called the autocorrelation function (ACF) *of* $x_1(t)$*, and*

$$\gamma(0) = 1 \tag{5.10}$$

The ACF of the FBM, which we interpreted in Chapter 4 as a fractional summation of white noise processes,

$$x(t) = (1 - L)^{-d}\varepsilon(t) \tag{5.11}$$

with $\varepsilon(t) \sim i.i.d.(0, \sigma_\varepsilon^2)$, is given by

$$\begin{aligned}
\gamma(\tau) &= E\{x(t)x(t-\tau)\} \\
&= \int_{-\infty}^{+\infty} x(t)L^\tau x(t)\,dt \\
&= \int_{-\infty}^{+\infty} (1-L)^{-d}\varepsilon(t)L^\tau(1-L)^{-d}\varepsilon(t)\,dt \\
&= \int_{-\infty}^{+\infty} (1-L)^{-2d}L^\tau\varepsilon^2(t)\,dt \\
&= \int_{-\infty}^{+\infty} \left[\sum_{\tau=0}^{\infty}(-1)^\tau\binom{-2d}{\tau}L^\tau\right]L^\tau\varepsilon^2(t)\,dt \\
&= \left[\sum_{\tau=0}^{\infty}(-1)^\tau\binom{-2d}{\tau}\right]L^{2\tau}\int_{-\infty}^{+\infty}\varepsilon^2(t)\,dt \\
&= \sigma_\varepsilon^2\sum_{\tau=0}^{\infty}c(\tau)
\end{aligned} \tag{5.12}$$

using the results of Chapter 4. The coefficients $c(\tau)$ can again be expressed in terms of the gamma function

$$\begin{aligned}
c(\tau) &= (-1)^\tau\binom{-2d}{\tau} \\
&= \frac{(-1)^\tau(-2d)!}{\tau!(-2d-\tau)!} \\
&= \frac{(-1)^\tau(-2d)(-2d-1)\cdots(-2d-\tau+1)}{\tau!} \\
&= \frac{(2d+\tau-1)\cdots(2d+1)(2d)}{\tau!} \\
&= \frac{(2d+\tau-1)!}{(2d-1)!\tau!} \\
&= \frac{\Gamma(\tau+2d)}{\Gamma(2d)\Gamma(\tau+1)!}
\end{aligned} \tag{5.13}$$

As the time horizon increases, $\tau \to \infty$,

$$c(\tau) \sim \frac{\tau^{2d-1}}{(2d-1)!} \tag{5.14}$$

There are three cases: when $d \downarrow -\frac{1}{2}$,

$$c(\tau) \sim \tau^{-2} \tag{5.15}$$

When $d = 0$, the Gaussian case occurs

$$c(\tau) \sim \tau^{-1} \tag{5.16}$$

When $d \uparrow \frac{1}{2}$,

$$c(\tau) \sim 1 \tag{5.17}$$

Thus, the ACF of the FBM is proportional to

$$\begin{aligned} \gamma(\tau) &\sim \sigma_\varepsilon^2 \tau^{2d-1} \\ &= \sigma_\varepsilon^2 \tau^{2H-2} \end{aligned} \tag{5.18}$$

This clearly shows that the ACF of the FBM is time-dependent, since it scales according to the time horizon τ. How fast it decreases in scale depends on the scaling exponent $2d - 1 = \lambda = 2H - 2$ or $2d + 1 = \lambda + 2 = 2H$, respectively (Table 4.2 in Chapter 4).

5.3.2 Symmetry properties of covariance functions

It is easy to show that, because of symmetry,

$$R_{12}(\tau) = R_{21}(-\tau) \tag{5.19}$$

and

$$R_{11}(\tau) = R_{11}(-\tau) \tag{5.20}$$

5.3.3 Time convolution

Signal engineers prefer to use the concept of time convolution, in contrast to time series statisticians, who prefer to use the concept of a covariance function. Shortly, we'll show that time convolution and the covariance function are equivalent.

Definition 187 *The* convolution *of two variables $x_1(t)$ and $x_2(t)$ is*

$$f(t) = \int_{-\infty}^{+\infty} x_1(u) x_1(t-u)\, du \tag{5.21}$$

which is often symbolically expressed by a "star" symbol $\star$ as

$$f(t) = x_1(t) \star x_2(t) \tag{5.22}$$

5.3.4 Properties of time convolution

Time convolution has three important algebraic properties, which are often used in theoretical Fourier analysis (Hsu, 1984):

(1) Convolution is *commutative*

$$x_1(t) \star x_2(t) = x_2(t) \star x_1(t) \tag{5.23}$$

(2) Convolution is *associative*

$$[x_1(t) \star x_2(t)] \star x_3(t) = x_1(t) \star [x_2(t) \star x_3(t)] \tag{5.24}$$

(3) Convolution is *distributive*

$$x_1(t) \star [x_2(t) + x_3(t)] = x_1(t) \star x_2(t) + x_1(t) \star x_3(t) \tag{5.25}$$

5.3.5 Covariance as time convolution

The cross-covariances of $x_1(t)$ and $x_2(t)$ are related to the convolutions of $x_1(t)$ and $x_2(-t)$, as follows. Let, by the definition of time convolution,

$$\begin{aligned} G_{12}(t) &= x_1(t) \star x_2(-t) \\ &= \int_{-\infty}^{+\infty} x_1(u) x_2[-(t-u)]\, du \\ &= \int_{-\infty}^{+\infty} x_1(u) x_2(u-t)\, du \end{aligned} \tag{5.26}$$

Changing the variable t to τ and the dummy variable u to t, we have

$$\begin{aligned} G_{12}(\tau) &= \int_{-\infty}^{+\infty} x_1(t) x_2(t-\tau)\, d\tau \\ &= R_{12}(\tau) \end{aligned} \tag{5.27}$$

Hence, a cross-covariance equals the following time convolution

$$R_{12}(\tau) = G_{12}(\tau) = x_1(t) \star x_2(-t)|_{t=\tau} \tag{5.28}$$

5.4 Fourier analysis of stationary price innovations

We will now first discuss the Fourier analysis of stationary periodic variables, e.g., of Random Walk price innovations, and, next, the Fourier analysis of stationary aperiodic variables.[2] Fourier analysis is a mathematical technique for transforming the view of a time series from a time-based one to a frequency-based one (Körner, 1990). It analyzes the "frequency content" of a time series.

This particular property is also the drawback of Fourier analysis, since in this transformation from the time domain to the frequency domain, the *timing* of information is lost, because a Fourier Transform (FT) is a global representation of a time series over the whole time domain $(-\infty, +\infty)$. But if a time series does not change much over time – i.e., if it is stationary – this drawback isn't very important. In fact, it's properties make it a very suitable tool for studying linear time-invariant operators (cf. Chapter 3), such as differentiation or integration with integer orders. Such classical research was the basis for the very first forays into empirical periodic analysis of financial price formation (Osborne, 1962).

However, if a given frequency is present in a time series $x(t)$ over only a limited time interval, the FT is unable to accurately detect this frequency and to give any information about its *lifetime* or *coherence* and about the moments of its appearance and of its disappearance.

The following analysis will culminate in the definition of the (constant) risk or power spectral density (PSD) for stationary aperiodic variables. This PSD can be visualized by a *spectrogram*. A spectrogram is a powerful visualization for empirical analysis of stationary random variables. The important Wiener–Khinchin Theorem will show that the risk spectrum is the FT of the auto-covariance function. Thus, we can analyze stationary aperiodic variables $x(t)$ by computing their covariance functions, or equivalently, their PSDs. This Fourier analysis of stationary series will be followed by the even more useful windowed Fourier analysis of slowly changing nonstationary variables in Chapter 6, since we have already observed that most financial market time series are nonstationary.

Definition 188 *A* periodic *variable is any variable for which*

$$x(t) = x(t - \tau) \quad \textit{for all } t \tag{5.29}$$

The smallest constant τ *that satisfies this equation is called the* period τ *of this variable.*

Remark 189 *By iteration, we have for periodic variables the following relationship*

$$x(t) = x(t - n\tau) \quad \textit{for } n = 0, \pm 1, \pm 2, \ldots \tag{5.30}$$

5.4.1 Fourier series for periodic variables

Definition 190 *A periodic variable can be represented equivalently by two trigonometric forms and one complex exponential form of the* Fourier series,

as follows

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{+\infty}(a_n \cos n\omega_0 t + b_n \sin n\omega_0 t) \tag{5.31}$$

$$= C_0 + \sum_{n=1}^{+\infty} C_n \cos(n\omega_0 t - \theta_n) \tag{5.32}$$

$$= \sum_{n=-\infty}^{+\infty} c_n e^{jn\omega_0 t} \tag{5.33}$$

where $\omega_0 = 2\pi/T$ *and the imaginary number* $j = \sqrt{-1}$.

The *Cartesian coefficients* a_n and b_n, respectively the *polar coefficients* C_n, as well as the *exponential coefficients* c_n are collectively called: the *Fourier resonance coefficients*. The second equivalent trigonometric Fourier series is called the *harmonics form*.

Notice that the Fourier series expansion of a periodic time variable describes a periodic variable as a sum of sinusoidal components having different frequencies. The sinusoidal component of *frequency* $\omega_n = n\omega_0$ is called the nth *harmonic* of the periodic variable and n is called the *wave number*. Here $\omega_0 = 2\pi f_0 = 2\pi/T$ is the *fundamental angular frequency* and $f_0 = 1/T$ is the *fundamental frequency* and the first harmonic $C_1 \cos(\omega_0 t - \theta_1)$ is called the *fundamental component* (because it has the same period as the variable $x(t)$). An increase in frequency decreases the wavelength.

The coefficients C_n and the angles θ_n are the *harmonic amplitudes*, i.e., they *scale* the amplitude of the sinusoidal waves, and *phase angles, respectively*, i.e., they *shift* the position of the sinusoidal waves, respectively. Thus, Fourier analysis scales and shifts the sinusoidal *bases* $e^{jn\omega_0 t}$ to achieve a complete analysis of the time series.

Example 191 *Fourier series can be used to approximate target time series* $x(t)$*, in this example, a step function, or square wave, as in Figure 5.1. We start with a mean* $f_0 = C_0 = 1$ *and, successively add the large wave* $f_1 = \cos\omega_0$*, subtract three times smaller and three times more frequent wave* $f_2 = \frac{1}{3}\cos 3\omega_0$ *to create the general "hat" shape. Next we add a five times smaller and faster wave* $f_3 = \frac{1}{5}\cos 5\omega_0$*. In the left column of Figure 5.1 are the target and terms* f_1 *through* f_3*. In the right column are* f_0 *and the succeeding sums, as each term is added to* f_0*. Notice that the approximation improves (i.e. each successive sum approximates the square wave more precisely) as the number of Fourier terms in the series increases. In the last graph, terms* f_5 *and* f_6 *are added (but not shown separately) to show further improvement in the approximation. Notice the* Gibbs phenomenon*, consisting of spurious* $sinc(t)$ *oscillations over the whole time domain. In older sound systems, which use Fourier approximation expansions for communication and transfer of information, this approximation error phenomenon causes a slight*

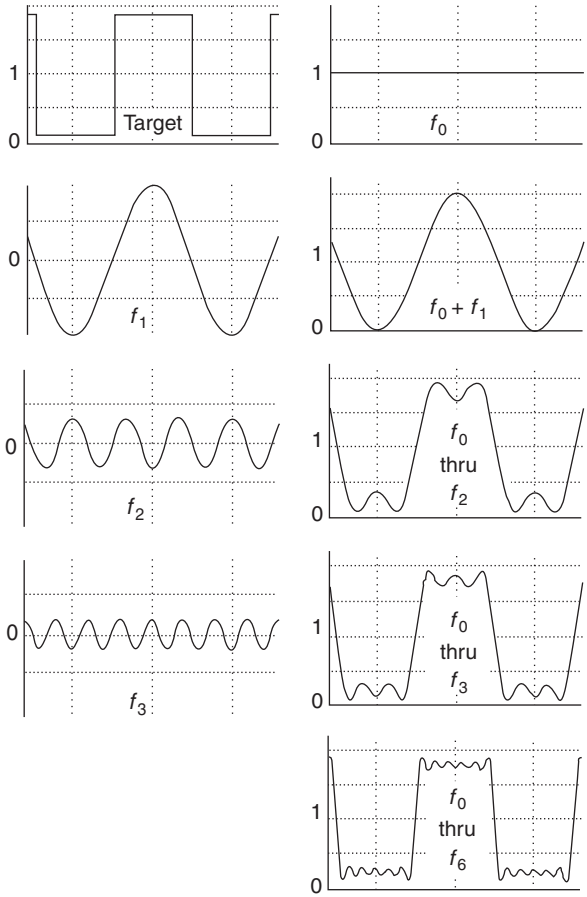


Figure 5.1 Fourier series approximation of a square wave.

"hiss." The Gibbs phenomenon will be discussed in greater detail in Chapter 11, when we are looking at the various finite element numerical solution methods for nonlinear diffusion equations. We'll find that interpolating wavelet expansions more precisely locate the Gibbs phenomenon than the Fourier trigonometric expansions thereby eliminating most of the Gibbs phenomenon.

Remark 192 *It is easy to proof (by checking) the following conversion formulas of the Fourier resonance coefficients.*

For $n \neq 0$

$$C_n = \sqrt{a_n^2 + b_n^2} = 2|c_n| \quad \textit{and} \quad \theta_n = \tan^{-1}\left(\frac{b_n}{a_n}\right) \tag{5.34}$$

and

$$c_n = |c_n| e^{j\phi_n} \quad \textit{where } |c_n| = \tfrac{1}{2}\sqrt{a_n^2 + b_n^2} \quad \textit{and} \quad \phi_n = \tan^{-1}\left(-\frac{b_n}{a_n}\right) = -\theta_n \tag{5.35}$$

or

$$c_n = \tfrac{1}{2}(a_n - jb_n) \quad \textit{and} \quad c_{-n} = \tfrac{1}{2}(a_n + jb_n) = c_n^* \tag{5.36}$$

where the asterisk $*$ *indicates the* complex conjugate*, so that*

$$a_n = 2\,\mathrm{Re}[c_n] \quad \textit{and} \quad b_n = -2\,\mathrm{Im}[c_n] \tag{5.37}$$

For $n = 0$

$$\frac{a_0}{2} = C_0 = c_0 \tag{5.38}$$

Example 193 *The sophisticated heat analysis, conducted by J. B. J. Fourier himself during Napoleon's campaign in Egypt, allows for time-varying Fourier resonance coefficients, which are better analyzed in Chapter 6, when we analyze nonstationary time series. It illustrates how Fourier analysis can be used to solve problems that are difficult to analyze in the time domain, but easier to solve in the frequency domain, as can be seen in Figure 5.2. (adapted from Burke-Hubbard, 1998, p. 13). To determine the temperature at time* t *of a metal bar (in the case of J. B. J. Fourier, the barrel of a cannon in Napoleon's Grand Army) that is cooling, one starts by measuring the bar's initial temperature (at* $t = 0$*), representing it as a temperature function* $x(s, t)$ *that depends on space* s *(= distance along the bar) and time* t*. Next one moves from physical space to the frequency domain, computing its time-dependent FT* $\mathcal{F}(\omega, t)$*, which tells us the coefficient* c_n *for each frequency* $\omega_n = n\omega_0$*, or heat wave number* n*, making up the heat function* $x(s, t)$ *at time* $t = 0$*. The Fourier resonance coefficients at time* $t = 0$*, are given by the formula*

$$c_n(0) = n^{-0.5} \tag{5.39}$$

The Fourier resonance coefficients at time t *are computed with the formula*

$$\begin{aligned} c_n(t) &= c_n(0) e^{(-n^2 t/100)} \\ &= n^{-0.5} e^{(-n^2 t/100)} \end{aligned} \tag{5.40}$$

It's clear that these time-dependent Fourier resonance coefficients decay over time: the heat waves vanish over time. Consider here just the coefficients for times $t = 1, 5, 10$ *and* 50*. For each such time, the coefficients are the same for the entire bar (Fourier resonance coefficients are global coefficients) and the information on space* x *seems to have disappeared. But this space information reappears when*

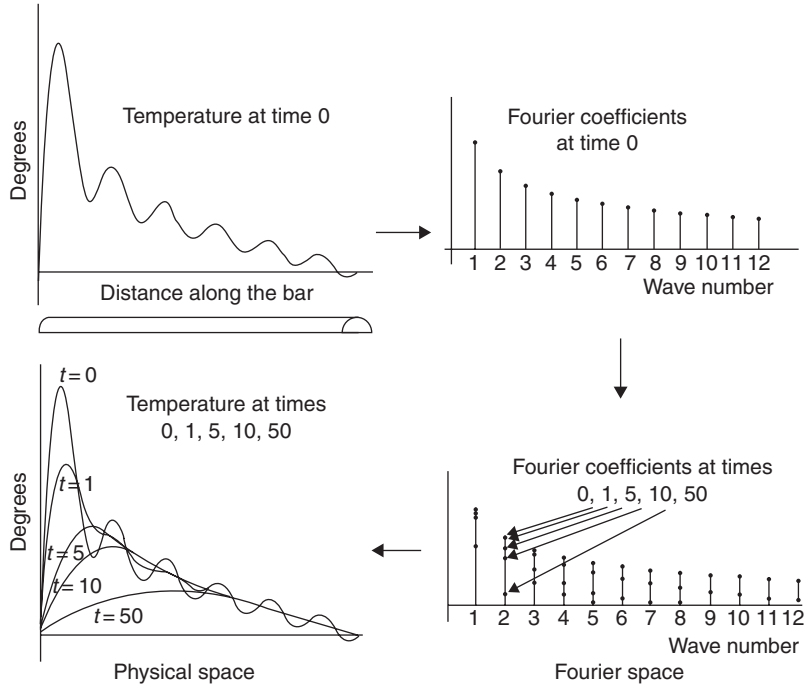


Figure 5.2 Heat diffusion analysis: by following the arrows we find a FT from the measured temperature in physical space to the heat wave frequency domain, and then analyze and return to physical space by the inverse FT.

we return from the frequency domain to the space domain: we invert the resolved FT $\mathcal{F}(\omega, t)$ *to obtain the function* $x(s, t)$*, which provides the exact temperature for each point* s *of the bar at any time* t*.*

5.4.2 Computation of the Fourier resonance coefficients

The constant Fourier resonance coefficients can be computed directly from the data, once we take account of the fundamental orthogonality of the sinusoidal bases.

5.4.2.1 Orthogonality of sinus and cosinus

Definition 194 *A set of functions* $\{\phi_k(t)\}$ *is* orthogonal *on an interval* $a < t < b$*, if, for any two functions* $\phi_m(t)$ *and* $\phi_n(t)$ *in the set* $\{\phi_k(t)\}$*, the following relationship holds*

$$\int_a^b \phi_m(t)\phi_n(t)dt = \begin{cases} 0 & \text{for } m \neq n \\ r_n & \text{for } m = n \end{cases} \tag{5.41}$$

Definition 195 *The set of functions* $\{\phi_k(t)\}$ *is* orthonormal*, when it is orthogonal and* $r_n = 1$ *for* $m = n$.

Using elementary calculus, one can easily show that sines and cosines form an orthogonal set of functions on the interval $-T/2 < t < T/2$, since

$$\int_{-T/2}^{T/2} \cos(m\omega_0 t) = 0 \quad \text{for } m \neq 0 \tag{5.42}$$

$$\int_{-T/2}^{T/2} \sin(m\omega_0 t) = 0 \quad \text{for all } m \tag{5.43}$$

$$\int_{-T/2}^{T/2} \cos(m\omega_0 t)\cos(n\omega_0 t)\,dt = \begin{cases} 0 & \text{for } m \neq n \\ T/2 & \text{for } m = n \neq 0 \end{cases} \tag{5.44}$$

$$\int_{-T/2}^{T/2} \sin(m\omega_0 t)\sin(n\omega_0 t)\,dt = \begin{cases} 0 & \text{for } m \neq n \\ T/2 & \text{for } m = n \neq 0 \end{cases} \tag{5.45}$$

$$\int_{-T/2}^{T/2} \sin(m\omega_0 t)\cos(n\omega_0 t)\,dt = 0 \quad \text{for all } m \text{ and } n \tag{5.46}$$

where $\omega_0 = 2\pi/T$.

Such a well-defined set of orthogonal analytic functions is called a *frame of reference*. When a frame of reference is *complete*, it forms a *basis* for analysis. Such a basis may contain functions that are not necessarily orthogonal (or orthonormal), but its analytic results, i.e., the computed correlation coefficients, are easier to understand when they are. Fortunately, each frame of reference can be orthogonalized. When we discuss wavelets in Chapter 7, we'll discuss more details of these important analytic frames of reference and bases.

5.4.2.2 Valuation of the trigonometric Fourier resonance coefficients

Using these orthogonality relations of the sines and cosines, we can now compute the Cartesian Fourier resonance coefficients a_n and b_n of the Fourier series $x(t)$ and, by using the conversion relations, also the polar Fourier resonance coefficients C_n and θ_n, and the exponential Fourier resonance coefficients of c_n and ϕ_n, as follows:

$$a_n = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\cos(n\omega_0 t)\,dt \quad \text{for } n = 0, 1, 2, \ldots \tag{5.47}$$

$$b_n = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\sin(n\omega_0 t)\,dt \quad \text{for } n = 0, 1, 2, \ldots \tag{5.48}$$

and

$$a_0 = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\,dt \tag{5.49}$$

5.4.2.3 Valuation of complex Fourier resonance coefficients

First, we have the simple mean

$$c_0 = \frac{a_0}{2} = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\,dt \tag{5.50}$$

Then, with the use of the identity

$$e^{-j\theta} = \cos\theta - j\sin\theta \tag{5.51}$$

we find

$$\begin{aligned} c_n &= \tfrac{1}{2}(a_n - jb_n) \\ &= \frac{1}{T}\left[\int_{-T/2}^{T/2} x(t)\cos(n\omega_0 t)\,dt - j\int_{-T/2}^{T/2} x(t)\sin(n\omega_0 t)dt\right] \\ &= \frac{1}{T}\left[\int_{-T/2}^{T/2} x(t)[\cos(n\omega_0 t)\,dt - j\sin(n\omega_0 t)]dt\right] \\ &= \frac{1}{T}\int_{-T/2}^{T/2} x(t)e^{-jn\omega_0 t}\,dt \end{aligned} \tag{5.52}$$

Similarly,

$$\begin{aligned} c_{-n} &= \tfrac{1}{2}(a_n + jb_n) \\ &= \frac{1}{T}\int_{-T/2}^{T/2} x(t)e^{jn\omega_0 t}\,dt \end{aligned} \tag{5.53}$$

These two formulas for c_n and c_{-n}, respectively, can be combined into a single exponential formula

$$c_n = \frac{1}{T}\int_{-T/2}^{T/2} x(t)e^{-jn\omega_0 t}\,dt \quad \text{for } n = 0, \pm 1, \pm 2, \ldots \tag{5.54}$$

These results have led to the powerful analytical identity of Parseval, which provides us with an exact accounting of the total amount of *risk* = *volatility* = *energy* = *power* contained in the financial time series $x(t)$, when decomposed into an infinite series of wave functions.

Proposition 196 (Parseval's identity) *If a_0, a_n and b_n are the coefficients in the Fourier expansion of a periodic function $x(t)$ with period T, then*

$$\frac{1}{T}\int_{-T/2}^{T/2}[x(t)]^2\,dt = \frac{a_0^2}{4} + \frac{1}{2}\sum_{n=1}^{+\infty}(a_n^2 + b_n^2) \tag{5.55}$$

$$= c_0^2 + 2\sum_{n=1}^{+\infty}|c_n|^2 \tag{5.56}$$

$$= \sum_{n=-\infty}^{+\infty}|c_n|^2 \tag{5.57}$$

This mean-square value is called the *risk content* of the periodic function $x(t)$. Thus, the finite estimate of the second moment of the periodic $x(t)$ equals this infinite sum of the squared Fourier resonance coefficients!

5.4.2.4 Orthogonality of complex exponential functions

The complex form of the Fourier series is the most useful. It is the initial platform for our discussion of wavelets in Chapter 7. Consider a set of *complex exponential functions* $\{e^{jn\omega_0 t}\}$ where the fundamental frequency is $\omega_0 = 2\pi/T$. Using elementary calculus, one can show that the mean

$$\frac{1}{T}\int_{-T/2}^{T/2} e^{jn\omega_0 t}\,dt = 0 \quad \text{for } n \neq 0 \tag{5.58}$$

and the (complex) variance

$$\frac{1}{T}\int_{-T/2}^{T/2} e^{jm\omega_0 t}(e^{jn\omega_0 t})^* dt = \begin{cases} 0 & \text{for } n \neq m \\ 1 & \text{for } n = m \end{cases} \tag{5.59}$$

The complex exponential functions $\{e^{jn\omega_0 t}\}, n = 0, \pm 1, \pm 2, \ldots$, form a set of orthogonal basis functions over the interval $-T/2 < t < T/2$. They form a complete frame of reference and, thus, a basis for analysis.

5.4.3 Frequency spectra

Definition 197 *A plot of the magnitude $|c_n|$ of the complex Fourier resonance coefficients c_n versus the angular frequency ω is called the* amplitude spectrum *of the periodic variable $x(t)$. A plot of the phase angle ϕ_n of c_n versus ω is called the* phase spectrum *of $x(t)$.*

Since for periodic series the index n assumes only integers, these two spectra are not continuous curves, but appear only at discrete frequencies $n\omega_0$. They are *discrete frequency spectra* or *line spectra.*

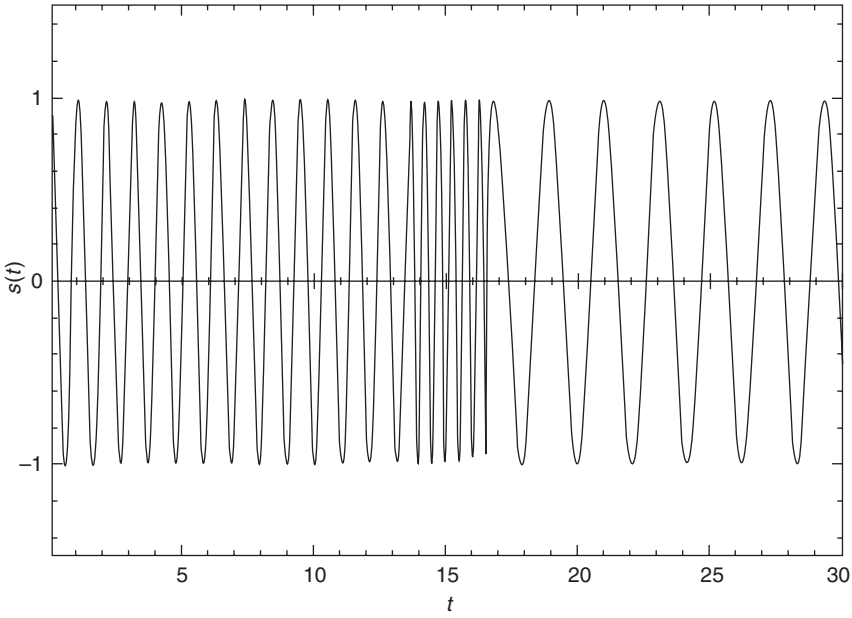


Figure 5.3 A sample signal constructed from sine functions representing three pulsations: $\omega_1 = 6, \omega_2 = 12, \omega_3 = 3$ with the same amplitude. The time coherence of ω_1 and ω_3 is identical. Discontinuities are present at the time of change of frequencies.

Example 198 *Figure 5.3 presents a sampled continuous time signal function $s(t)$ constructed from three successive sine functions with the same coherence or lifetime T_1 for frequency ω_1 and $\omega_3 < \omega_1$, and shorter coherence T_2 for frequency $\omega_2 > \omega_1$ (Bendjoya and Slezak, 1993, pp. 233–234). The signal $s(t)$ shows discontinuities at the time of change of frequencies. Figure 5.4 displays the modulus of the Fourier resonance coefficients $|c_n|$, i.e., its* line spectrum. *Notice the rather good detection of the three frequencies present in $s(t)$ and how the high frequency ω_2 is less precisely detected because of its shorter coherence. The spurious fluctuations all over this spectrum are due to the discontinuities. No information can be obtained from the line spectrum about the sequence of the changes in frequencies, the time at which the different frequencies appear and disappear in $s(t)$, or about their coherence. Only a "timeless" frequency analysis is performed, which can be distorted by discontinuities.*

Example 199 *Figure 5.5 shows examples of the magnitude spectra of several musical variables, which are, clearly, periodic, sinusoidal waves (Kemp, 1991, p. 12). The clarinet, the violin and the highland bagpipe playing the same note (B flat above middle C, 466 Hz). The* interferograms *(=* correlograms*)*

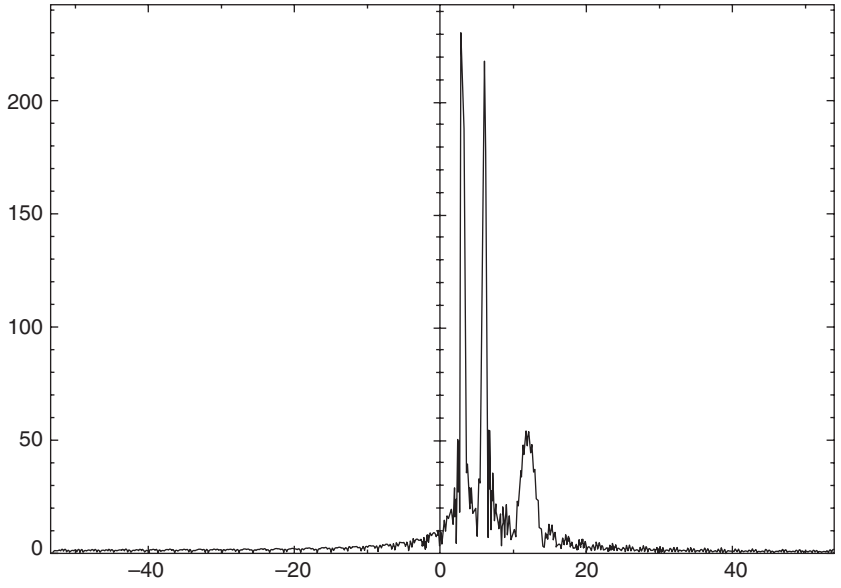


Figure 5.4 The FT of the sampled signal $s(t)$. The three frequencies $\omega_3 < \omega_1 < \omega_2$ are detected, but the frequency ω_2 with he shortest time coherence has the smallest resonance coefficient. The spurious fluctuations all over this spectrum are due to the discontinuities in the signal.

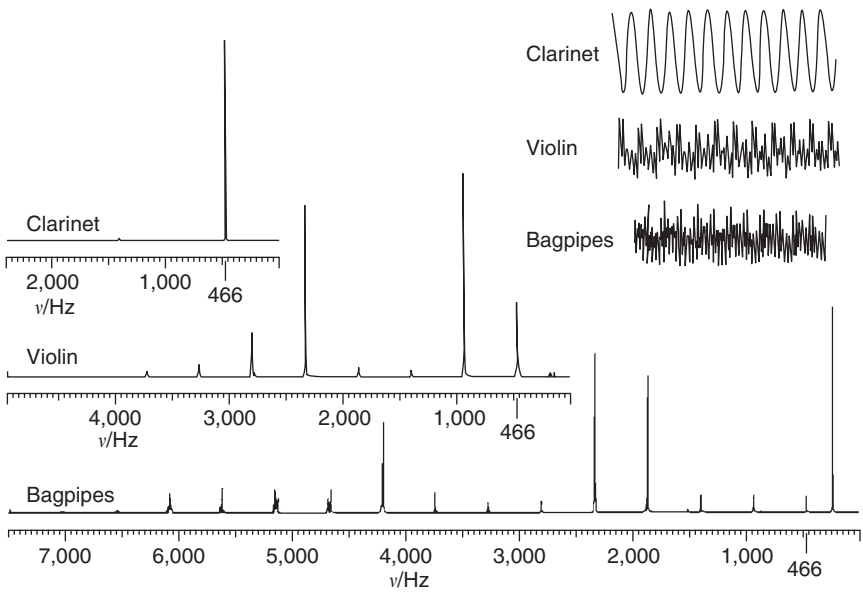


Figure 5.5 Fourier series analysis of pure musical harmonics: dominant frequencies of the clarinet, violin and bagpipe.

at the upper right-hand side were Fourier transformed to show the resonance coefficients of their individual frequencies, or resonances emitted, as shown at the lower left side. Notice that the clarinet is almost pure, with one dominant frequency at 466 Hz. The violin shows several dominating higher frequency overtones in addition to the 466 Hz. And the bagpipe shows both a dominant lower harmonic (from the bas drone) and several dominant higher frequency harmonics.

5.5 Software

The computations of the following Exercises can be executed by using the MATLAB® Signal Processing Toolbox and by the MATLAB® Higher-Order Spectral Analysis (HOSA) Toolbox (Swami *et al*., 1998). Both Toolboxes are available from The MathWorks, Inc., 24 Prime Park Way Natick, MA 01760-1500, USA. Tel (508) 647-7000; Fax (508) 647-7001;

http://www.mathworks.com/products/wavelettbx.shtml.

The HOSA Toolbox (2.0.3) is a collection of MATLAB® M-files containing specialized tools for signal processing with higher order spectra. It was created by Jerry M. Mendel, Chrysostomos L. (Max) Nikias and Ananthram Swami of United Signals and Systems, Inc. The toolbox is a collection of MATLAB® routines whose primary features are functions for: higher order spectrum estimation either by conventional or parametric approaches; magnitude and phase retrieval; adaptive linear prediction; harmonic retrieval and quadratic phase coupling; time-delay estimation and array signal processing.

5.6 Exercises

Run MATLAB® Help, Examples and Demos, Toolboxes, Signal Processing:

Exercise 200 *Filtering a sinusoidal signal: look carefully at the creation of a sinusoidal signal with different frequencies by superposition, the creation and implementation of the IIR (= Infinite Impulse Response filter), the way the various series are plotted, the magnitude versus frequency diagram of the Fast Fourier Transform (FFT) (cf. Brigham, 1988).*

Exercise 201 *Spectral Analysis of the DTMF signal (with sound): study the spectrum diagram*

Exercise 202 *Discrete Fourier Transform (DFT) (try different windows): study the effects on the DFT by changing frequency and amplitude of the signal. Notice that there doesn't exist a DFT for an infinite signal. (It's somewhat confusing that the command for the DFT in MATLAB® is fft. For the inverse DFT it is ifft.)*

Exercise 203 *Continuous FT: study the effects of different modulation frequencies on the FT of the modulated Gaussian pulse. Notice the symmetry of the FT.*

Notes

1 Chapter 6 contains a visual example of the original spectral analysis of the Standard and Poor stock market series by Granger and Morgenstern (1963).
2 Jean Baptiste Joseph Fourier (1768–1830) was a French mathematician, who became famous for his mathematical treatise on the *Theorie Analytique de la Chaleur (Analytical Theory of Heat)*, 1822. He established the partial differential equation governing the heat diffusion in the barrels of the canons of Napoleon's Grand Army to determine how quickly these cannons could be safely reloaded. He solved it by using an infinite series of trigonometric functions. His diffusion (= partial differentiation) equation was used in 1973 to derive the Black–Scholes European option pricing model. This heat diffusion equation is also used in the theory of turbulence, as we will see in Chapter 11.

Bibliography

Anderson, T. W. (1994) *The Statistical Analysis of Time Series*, John Wiley and Sons, New York, NY.

Bendjoya, Ph., and E. Slézak (1993) "Wavelet Analysis and Applications to Some Dynamical Systems," *Celestial Mechanics and Dynamical Astronomy*, **56**, 231–262.

Bloomfield, P. (1976) *Fourier Analysis of Time Series: An Introduction*, John Wiley and Sons, New York, NY.

Box, G. E. P., and G. M. Jenkins (1970) *Time Series Analysis: Forecasting and Control*, 2nd edn, Holden-Day, San Francisco, CA.

Brigham, E. O. (1988) *The Fast Fourier Transform and Its Applications*, Prentice-Hall: Englewood Cliffs, NJ.

Burke-Hubbard, Barbara (1998) *The World According to Wavelets: The Story of a Mathematical Technique in the Making*, 2nd edn, A K Peters, Wellesley, MA.

Champeney, D. C. (1990) *A Handbook of Fourier Theorems*, Cambridge University Press, Cambridge, UK.

Folland, G. (1992) *Fourier Analysis and its Applications*, Wadsworth and Brooks/Cole, Belmont, CA.

Granger, C. W. J. (1966) "The Typical Spectral Shape of an Economic Variable," *Econometrica*, **34**(1), 150–161.

Granger, C. W. J., and O. Morgenstern (1963) "Spectral Analysis of New York Stock Exchange Prices," *Kyklos*, **16**, 1–27. Reprinted as chapter 8 in Cootner, Paul H. (Ed.) (1964) *The Random Character of Stock Market Prices*, The MIT Press, Cambridge, MA, pp. 162–188.

Hsu, Hwei P. (1984) *Applied Fourier Analysis*, Harcourt Brace College Publishers, San Diego, 1984, chapter 6.

Jenkins, G. M., and D. G. Watts (1968) *Spectral Analysis and its Applications*, Holden-Day, San Francisco, CA.

Körner, T. W. (1990) *Fourier Analysis*, Cambridge University Press, Cambridge, UK.

Kemp, William (1991) *Organic Spectroscopy*, 3rd edn, W. H. Freeman and Co., New York.

Osborne, M. F. M. (1962) "Periodic Structure in the Brownian Motion of Stock Prices," *Operational Research*, **10**, May–June, 345–379. Reprinted as chapter 13 in Cootner, Paul H. (Ed.) (1964), *The Random Character of Stock Market Prices*, The MIT Press, Cambridge, MA, pp. 262–296.

Nikias, Chrysostomos L., and Athina P. Petropulu (1993) *Higher-Order Spectral Analysis: A Nonlinear Signal Processing Framework*, PTR Prentice Hall, Englewood Cliffs, NJ.

Priestley, M. B. (1981) *Spectral Analysis and Time Series*, Academic Press, New York, NY.

Swami, Ananthram, Jerry M. Mendel and Chrysostomos L. (Max) Nikias (1998) *Higher-Order Spectral Analysis Toolbox User's Guide*, Version 2, January, The MathWorks, Inc., Natick, MA.

6 Fourier time–frequency analysis of risk

6.1 Introduction

As we discussed in Chapter 5, Fourier series are powerful tools for analyzing periodic variables, such as musical tones of instruments, or for determining the spectral lines of inorganic and organic chemical components. But very few practical problems of financial-economic analysis do involve such rigidly periodic variables. We need a form of Fourier analysis that can deal with aperiodic, but still "cyclical" variables and that can identify Mandelbrot's aperiodic cyclicity prevalent in the long memory financial return series. Therefore, we'll discuss a frequency representation of aperiodic variables by means of the Fourier Transform (FT), which analyzes the frequency contents of *any* time series, periodic or aperiodic, as the case may be.

The many properties of the FT make it particularly suitable for representation in terms of linear time-invariant system operators, such as integer differentiation or integration, discussed in Chapter 4. Its immediate shortcoming is the same as that of the Fourier series: the FT is a global, and not a local, representation of a time series, since it takes an integral, or average, of the available set of time series observations.

In the second half of this chapter we'll discuss the Windowed FT, which is suitable to analyze *transient* phenomena localized in time, although, perhaps, only suboptimally, since the support of the Fourier wave bases remains infinite. This Windowed FT was discovered by Gábor (1946) and produces the colorful spectrograms of instantaneous frequency distributions familiar from speech analysis and other, rather entertaining internet media, such as RealPlayer^{TM}, which lists real-time spectrograms among its audio statistics. Considering that financial time series of investment returns or foreign valuta prices are as nonstationary as speech, Windowed FT forms a powerful, and still heavily under-utilized, research tool for the time–frequency analysis of financial risk (Priestley, 1988; Cohen, 1989).

A very recent, optimal and complete way of analyzing such localized phenomena, which does not suffer from the "infinite support syndrome" of the FT, but which relies instead on finite support, will be discussed in the next chapter, when we focus on time series analysis by finite wavelet bases.

6.2 FT for aperiodic variables

The continuous FT for aperiodic time series is analogously defined to the discrete Fourier series for periodic time series of Chapter 5, as follows (Hsu, 1984; Champeney, 1990).

Definition 204 *The* FT *of time series $x(t)$ (symbolized by $\mathcal{F}$) in the square-integrable L^2 space is defined by the inner product (or correlation):*

$$\begin{aligned} F(\omega) &= \mathcal{F}[x(t)] \\ &= \int_{-\infty}^{+\infty} x(t) e^{-j\omega t} dt \end{aligned} \tag{6.1}$$

Definition 205 *The* inverse FT *of $F(\omega)$ (symbolized by $\mathcal{F}^{-1}$) represents the time series $x(t)$ as an infinite integral ("sum") of sinusoids:*

$$\begin{aligned} x(t) &= \mathcal{F}^{-1}[F(\omega)] \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} d\omega \end{aligned} \tag{6.2}$$

Remark 206 *These two equations are often called the* FT pair*, symbolically denoted by*

$$F(\omega) \leftrightarrow x(t) \tag{6.3}$$

The condition for the existence of the FT $F(\omega)$ is given by

$$\int_{-\infty}^{+\infty} |x(t)|\, dt < \infty \tag{6.4}$$

In other words, the variable $x(t)$ must be *absolutely integrable*. This is the same restrictive condition as exists for martingales (cf. Chapter 2) and, again, it excludes discontinuities and jumps, but also periodic functions, because, for example, for any periodic function:

$$\int_{-\infty}^{+\infty} |x(t)|\, dt = \infty \tag{6.5}$$

Thus, strictly defined, the FT (which is NOT the Fourier series) cannot properly deal with discontinuities or singularities (catastrophes) and periodic functions. However, the Windowed FT can detect other transient phenomena. So let's see how we can understand this form of windowed analysis.

Since any periodic variable $x(t)$ is a function of slow growth, its FT exists in the sense of a generalized function, as follows.

Definition 207 A function $x(t)$ of slow growth *is defined if there exist real numbers C and R such that*

$$x(t) \leq Ct^n \quad \textit{where } |t| > R > 0 \tag{6.6}$$

Immediately related to this concept of slow growth is the concept of a restricting device, such as a *taper*, which "tapers" the influence of the individual observations.

Definition 208 *A continuous function $\phi(t)$ is a* testing function of rapid decay*, or* taper*, if*

$$\lim_{t \to \pm\infty} |t^n \phi^{(r)}(t)| = 0 \quad \textit{for some } n, r \geq 0 \tag{6.7}$$

where the rth derivative

$$\phi^{(r)}(t) = \frac{d^r \phi(t)}{dt^r} \tag{6.8}$$

In the Exercises, you'll find an application of such a taper for FT analysis. Such a taper can be used to define even more generalized functions, as follows.

Definition 209 *A* generalized function of slow growth $g(t)$ *is defined as a symbolic function, such that to each testing function of rapid decay $\phi(t)$ there is assigned a finite number to the inner product:*

$$\langle g, \phi \rangle = \int_{-\infty}^{+\infty} g(t)\phi(t)dt < \infty \tag{6.9}$$

with the linear properties of additivity and homogeneity. Thus

$$\langle g, a_1\phi_1 + a_2\phi_2 \rangle = a_1\langle g, \phi_1 \rangle + a_2\langle g, \phi_2 \rangle \tag{6.10}$$

Remark 210 *If $\phi(t)$ is a taper, then we can use advanced calculus to show that $\phi(t)$ is absolutely integrable:*

$$\int_{-\infty}^{+\infty} |\phi(t)|\, dt < \infty \tag{6.11}$$

Hence, the FT of this taper, $\Phi(\omega)$, exists.

In the next chapter we'll discuss the relaxation of this particular restraining condition of either $\int_{-\infty}^{\infty} |x(t)|\, dt < \infty$ or $\int_{-\infty}^{\infty} |\phi(t)|\, dt < \infty$ for the Wavelet Transform. The Wavelet Transform $\psi(t)$ of $x(t)$ exists, even when $x(t)$ includes discontinuities, jumps, periodicities and cyclicities – in other words, when $x(t)$ includes all the phenomena we observe in empirical financial time series! Because the Wavelet Transform has finite and not infinite support.

But let's now first define the generalized FT.

Definition 211 *The* generalized FT $F(\omega)$ *of a function of slow growth* $x(t)$ *is defined by the commuting integrals*

$$\int_{-\infty}^{+\infty} F(u)\phi(u)du = \int_{-\infty}^{+\infty} x(u)\Phi(u)du \tag{6.12}$$

Remark 212 *It can be shown that all the properties of ordinary FTs also hold for the generalized FTs of functions of slow growth (Hsu, 1984; Körner, 1990).*

6.2.1 Algebraic properties of FTs

FTs have some very useful properties, which makes it easy to *exactly* compute an enormous variety of FTs, which are discussed in a historically very interesting and enjoyable presentation by Körner (1990). Let $F(\omega) \leftrightarrow x(t)$ denote the FT pair. Then it's easy to prove the following nine properties (cf. Hsu, 1984, for the particulars of these proofs). In particular, translation = time shifting and scaling = frequency shifting, will be very useful, when we study empirical financial series that exhibit long time-dependence, such as the Fractional Brownian Motion (FBM) defined in Chapter 4. Here follow the nine fundamental properties of FT pairs often used in theoretical signal processing analysis, and now also in theoretical dynamic asset valuation:

(1) *Convolution* in the time domain = multiplication in the frequency domain:

$$x_1(t) \star x_2(t) = F_1(\omega)F_2(\omega) \tag{6.13}$$

(2) *Multiplication* in the time domain = convolution in the frequency domain:

$$x_1(t)x_2(t) = \frac{1}{2\pi} F_1(\omega) \star F_2(\omega) \tag{6.14}$$

(3) *Linearity* in the time domain = linearity in the frequency domain:

$$a_1x_1(t) + a_2x_2(t) \leftrightarrow a_1F_1(\omega) + a_2F_2(\omega) \tag{6.15}$$

(4) *Translation* (= time shifting) = complex exponential decay in the (imaginary) frequency domain:

$$x(t - t_0) \leftrightarrow F(\omega)e^{-j\omega t_0} \tag{6.16}$$

(5) *Modulation* (= frequency shifting) = complex exponential increase in the (imaginary) time-domain:

$$x(t)e^{j\omega_0 t} \leftrightarrow F(\omega - \omega_0) \tag{6.17}$$

(6) *Scaling up* in the time domain = scaling down in the frequency domain, and *vice versa*:

$$x(ct) \leftrightarrow \frac{1}{|c|} F\left(\frac{\omega}{c}\right) \tag{6.18}$$

(7) *Time-reversal* in the time domain = frequency-reversal in the frequency domain:

$$x(-t) \leftrightarrow F(-\omega) \tag{6.19}$$

(8) *Symmetry* of functions in time and frequency domain:

$$F(t) \leftrightarrow 2\pi x(-\omega) \tag{6.20}$$

(9) *Differentiation* in the time domain = frequency exponential in the frequency domain and *vice versa*:

$$x^{(p)}(t) \leftrightarrow (j\omega)^p F(\omega) \quad \text{and} \tag{6.21}$$

$$(-jt)^p x(t) \leftrightarrow F^{(p)}(\omega) \tag{6.22}$$

Using the fundamental linearity and time shifting properties, we can now find the FT of the Geometric Brownian Motion (GBM) as follows:

$$\begin{aligned} F(\omega) &= \mathcal{F}[x(t)] \\ &= \mathcal{F}[x(t-1) + \varepsilon(t)] \\ &= F(\omega)e^{-j\omega} + \mathcal{F}[\varepsilon(t)] \end{aligned} \tag{6.23}$$

Notice the frequency translation of the time shift! This implies that we can concisely represent the GBM in the frequency domain as follows:

$$F_{\text{GBM}}(\omega) = (1 - e^{-j\omega})^{-1}\mathcal{F}[\varepsilon(t)] \tag{6.24}$$

Similarly, the FT of the FBM is modeled in the frequency domain as follows:

$$F_{\text{FBM}}(\omega) = (1 - e^{-j\omega})^{-d}\mathcal{F}[\varepsilon(t)] \tag{6.25}$$

We'll need these spectral representations later in this chapter, when we focus on the spectral density of the stationary increments of the FBM.

6.2.2 Some exact FTs

Here are some additional important FTs of exact time functions, which can be easily checked (Hsu, 1984):

(1) The FT of a constant is an *impulse function* $\delta(\cdot)$, which is the first derivative of the unit step function $u(\cdot)$:

$$1 \leftrightarrow 2\pi\delta(\omega) = 2\pi\frac{du(\omega)}{d\omega} \tag{6.26}$$

(2) The FT of a complex exponential function results in an impulse with a shifted frequency:

$$e^{j\omega_0 t} \leftrightarrow 2\pi\delta(\omega - \omega_0) \tag{6.27}$$

(3) The FT of a cosine function consists of the sum of two frequency shifted impulse functions:

$$\cos \omega_0 t \leftrightarrow \pi \delta(\omega - \omega_0) + \pi \delta(\omega + \omega_0) \tag{6.28}$$

(4) The FT of a periodic function with period T, which, as we have seen, can always be expressed as a Fourier series of exponential functions, consists of a sequence of equidistant impulses located precisely at the harmonic frequencies of the function (cf. Chapter 5):

$$x(t) = \sum_{n=-\infty}^{+\infty} c_n e^{jn\omega_0 t}, \quad \text{with } \omega_0 = \frac{2\pi}{T}, \tag{6.29}$$

$$\leftrightarrow F(\omega) = 2\pi \sum_{n=-\infty}^{+\infty} c_n \delta(\omega - n\omega_0) \tag{6.30}$$

Remark 213 *To speed up the calculations by reducing the number of computing operations, often we implement the so-called* Fast Fourier Transform (FFT), which *separates its odd and even harmonics. The vector-matrix implementation of this FFT, discovered by Cooley and Tukey (1965), provides an interesting advanced topic in numerical analysis.*

6.2.3 Fourier spectra

The operational *raison d'être* for FTs is to enable the computation of frequency spectra of any continuous or discrete time series.

Definition 214 *The Fourier spectrum $F(\omega) = \mathcal{F}[x(t)]$ is, in general, complex, and thus represented by the sum of real and imaginary parts:*

$$\begin{aligned} F(\omega) &= R(\omega) + jX(\omega) \\ &= |F(\omega)| e^{j\phi(\omega)} \end{aligned} \tag{6.31}$$

where $|F(\omega)|$ is called the magnitude (amplitude) spectrum *of $x(t)$ and $\phi(\omega)$ its* phase spectrum.

When $x(t)$ is a real-valued time series (and, empirically, it always is!), then, using the familiar goniometric identity for a complex exponential,

$$e^{-j\omega t} = \cos \omega t - j \sin \omega t \tag{6.32}$$

its FT can be rewritten as

$$\begin{aligned} F(\omega) &= \int_{-\infty}^{+\infty} x(t)e^{-j\omega t}dt \\ &= \int_{-\infty}^{+\infty} x(t)\cos(\omega t)dt - j\int_{-\infty}^{\infty} x(t)\sin(\omega t)dt \\ &= R(\omega) + jX(\omega) \end{aligned} \tag{6.33}$$

so that, equating the real and imaginary parts of the complex $F(\omega)$, we have the real terms

$$R(\omega) = \int_{-\infty}^{+\infty} x(t)\cos(\omega t)dt \tag{6.34}$$

$$X(\omega) = -\int_{-\infty}^{+\infty} x(t)\sin(\omega t)dt \tag{6.35}$$

which are both easy to compute.

6.2.4 Convolution Theorems

The (general) FT can now be related to the earlier sections of Chapter 5, where we discussed correlation and convolution, by way of two powerful Theorems. These Theorems allow convolutions to be replaced by simple products (cf. Hsu, 1984; Champeney, 1990; Körner, 1990, for the respective proofs). The first Theorem shows that convolution in the time domain can be replaced by a product in the frequency domain. The second Theorem shows that convolution in the frequency domain can be replaced by a product in the time domain.

Theorem 215 (Time Convolution) *If* $\mathcal{F}[x_1(t)] = F_1(\omega)$ *and* $\mathcal{F}[x_2(t)] = F_2(\omega)$ *then*

$$\mathcal{F}[x_1(t) * x_2(t)] = F_1(\omega)F_2(\omega) \tag{6.36}$$

Theorem 216 (Frequency Convolution) *If* $\mathcal{F}^{-1}[F_1(\omega)] = x_1(t)$ *and* $\mathcal{F}^{-1}[F_2(\omega)] = x_2(t)$ *then*

$$\mathcal{F}^{-1}[F_1(\omega) \star F_2(\omega)] = 2\pi x_1(t)x_2(t) \tag{6.37}$$

or, equivalently,

$$\begin{aligned} \mathcal{F}[x_1(t)x_2(t)] &= \int_{-\infty}^{+\infty} [x_1(t)x_2(t)]\, e^{-j\omega t}dt \\ &= \frac{1}{2\pi}F_1(\omega) \star F_2(\omega) \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} F_1(y)F_2(\omega - y)dy \end{aligned} \tag{6.38}$$

Using the second Theorem and setting the frequency $\omega = 0$, we obtain

$$\int_{-\infty}^{+\infty} [x_1(t)x_2(t)]dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F_1(y)F_2(-y)dy$$
$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} F_1(\omega)F_2(-\omega)d\omega \tag{6.39}$$

by changing the dummy variable of integration. If $x(t)$ is real valued, then

$$F(-\omega) = F^*(\omega) \tag{6.40}$$

where $F^*(\omega)$ is the complex conjugate of $F(\omega)$. By substitution in the preceding equation we derive the following financial version of the well-known ergodic Theorem of Parseval, which provides a crucial link between the risk content of a financial time series, as measured in the time domain and in the frequency domain, respectively.

Theorem 217 (Parseval) *If the FT* $\mathcal{F}[x(t)] = F(\omega)$*, then the* risk content *(= second moment) of the aperiodic stationary financial time series* $x(t)$ *is*

$$E[x(t)^2] = \int_{-\infty}^{+\infty} [x(t)]^2 dt$$
$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} |F(\omega)|^2 d\omega$$
$$= \int_{-\infty}^{+\infty} |F(2\pi\nu)|^2 d\omega \tag{6.41}$$

where the angular frequency is $\omega = 2\pi\nu$ *and the frequency* ν *is expressed in Hertz.*

The quantity $|F(\omega)|^2$ is called the *risk* or *power spectrum*, or *power spectral density* (PSD) of $x(t)$. It is the frequency domain equivalent of risk in the time domain.

6.2.5 Wiener–Khintchin Theorem

Using the foregoing results, we can now present the FTs of the various covariance functions.

Corollary 218 *If* $\mathcal{F}[x_1(t)] = F_1(\omega)$ *and* $\mathcal{F}[x_2(t)] = F_2(\omega)$*, then*

$$S_{12}(\omega) = \mathcal{F}[R_{12}(\tau)] = F_1(\omega)F_2(-\omega) \tag{6.42}$$
$$S_{21}(\omega) = \mathcal{F}[R_{21}(\tau)] = F_1(-\omega)F_2(\omega) \tag{6.43}$$
$$S_{11}(\omega) = \mathcal{F}[R_{11}(\tau)] = F_1(\omega)F_1(-\omega) \tag{6.44}$$

The measures $S_{12}(\omega)$ *and* $S_{21}(\omega)$ *are referred to as* cross-risk *or* cross-spectral densities *(CSD), and* $S_{11}(\omega)$ *is, as we saw, the risk spectrum or PSD of* $x_1(t)$.

If $x_1(t)$ is real valued, then

$$\begin{aligned} S_{11}(\omega) &= \mathcal{F}[R_{11}(\tau)] \\ &= F_1(\omega)F_1(-\omega) \\ &= F_1(\omega)F_1^*(\omega) \\ &= |F(\omega)|^2 \end{aligned} \tag{6.45}$$

Thus, we have arrived at the famous Wiener–Khinchin Theorem relating the autocovariance function of a financial time series to its risk spectrum, and *vice versa*.[1] Thereby, the time dependence of a time series is translated into its frequency dependence, and *vice versa*.[2]

Theorem 219 (Wiener–Khinchin) *The autocovariance function* $R_{11}(\tau)$ *and the risk spectral density* $|F(\omega)|^2$ *constitute a FT pair:*

$$\begin{aligned} |F(\omega)|^2 &= \mathcal{F}[R_{11}(\tau)] \\ &= \int_{-\infty}^{+\infty} R_{11}(\tau)e^{-j\omega\tau}d\tau \end{aligned} \tag{6.46}$$

and

$$\begin{aligned} R_{11}(\tau) &= \mathcal{F}^{-1}\left[|F(\omega)|^2\right] \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} |F(\omega)|^2 e^{j\omega\tau}d\omega \end{aligned} \tag{6.47}$$

Since for *periodic* or *random* variables that exist over the entire time interval $(-\infty, \infty)$, the risk contents are infinite,

$$E[x(t)^2] = \int_{-\infty}^{+\infty} [x(t)]^2 dt \to \infty \tag{6.48}$$

the covariance functions, as defined earlier, do not exist as finite numbers, nor do their FTs. Therefore, pragmatically we must work with truncated, approximating average covariance functions, based on the assumed ergodicity of the time series, i.e., the assumed equivalence of the represented time series volatility in the time and frequency domains.

Definition 220 *The* average autocovariance function *of* $x_1(t)$ *is the limit*

$$\bar{R}_{11}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x_1(t)x_1(t-\tau)dt \tag{6.49}$$

Definition 221 *Similarly, the* average cross-covariance function *of* $x_1(t)$ *and* $x_2(t)$ *is the limit*

$$\bar{R}_{12}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x_1(t)x_2(t-\tau)dt \tag{6.50}$$

These definitions assist us to precisely define what we mean by *uncorrelatedness* of two time series, which is defined as their *linear independence*.

Definition 222 *Two variables $x_1(t)$ and $x_2(t)$ are* uncorrelated*, if we can decompose their cross-correlation into a product of two independent time averages:*

$$\begin{aligned}\bar{R}_{12}(\tau) &= \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} x_1(t)x_2(t-\tau)dt \\ &= \left[\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} x_1(t)dt\right]\left[\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} x_2(t-\tau)dt\right] \end{aligned} \tag{6.51}$$

Then, if one time series, say $x_1(t)$, has also a zero average value (e.g. because it is measured as deviations from the mean)

$$\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} x_1(t)dt = 0 \tag{6.52}$$

then their cross-correlation equals zero:

$$\bar{R}_{12}(\tau) = 0 \quad \text{for all } \tau \tag{6.53}$$

Thus, uncorrelatedness of two financial time series is empirically easy to verify: compute the deviations from their means and cross-correlate to see if the result equals zero. However, this procedure only measures uncorrelatedness = linear independence, and does not demonstrate anything about nonlinear independence or global independence, i.e., the kind of independence that financial risk analysts are currently most concerned about. Global dependence has major consequences for the way we conduct financial risk measurement, analysis and management if it does not exist, as we will see in Chapter 12.

Definition 223 *For time series with infinite risk content, the* average risk *of time series $x(t)$ is defined as the approximation*

$$\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}[x(t)]^2dt \tag{6.54}$$

Definition 224 *The* risk spectrum *or PSD of the financial time series $x_1(t)$ is the FT of the average autocovariance function of $x_1(t)$, which does exist, since the*

average autocovariance function is finite. Thus

$$P(\omega) = \mathcal{F}[\bar{R}_{11}(\tau)] = \int_{-\infty}^{+\infty} \bar{R}_{11}(\tau) e^{-j\omega\tau} d\tau \tag{6.55}$$

Then, of course, the average autocovariance function *is the inverse FT of the risk spectrum or PSD:*

$$\bar{R}_{11}(\tau) = \mathcal{F}^{-1}[P(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} P(\omega) e^{j\omega\tau} d\omega \tag{6.56}$$

These are pragmatic, practical definitions for situations which are likely to occur. However, these pragmatic definitions can lead to distortions, in particular when the risk is infinite, as is the case when time series contain discontinuities, or other singularities. These formulas define only approximations to the information content of a financial time series and do not provide a complete analysis.

Corollary 225 *The* average risk *(or mean-square value) of a financial time series $x_1(t)$ is given by the integration of the PSD $P(\omega)$ over the entire frequency range, since*

$$\bar{R}_{11}(0) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} [x_1(t)]^2 dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} P(\omega) d\omega \tag{6.57}$$

In the case of a stationary financial time series, it does not matter if we investigate its risk content in the time domain or in the frequency domain, since they are equivalent representations. But it does very much matter in which domain we investigate the risk content when the time series is nonstationary, since then the risk contents in the time and the frequency domains, respectively, are not equivalent. They cannot be transformed into each other and have to be looked at simultaneously to achieve a complete analysis.

Example 226 *Figure 6.1 shows the truncated financial risk spectrum of the Standard and Poor series computed by Granger and Morgenstern in 1963,* after *an important trend in the mean is removed, by using moving averages of lengths* 80 *and* 36*. In other words, the original time series was nonstationary! The spectrum was computed at* 240 *frequency bands, but only the first* 100 *are shown. A small resonance peak at* 40 *months can be observed, but is not statistically significant. Even after the trend removal, this peak only accounts for slightly less than* 10 *percent of the total remaining variance. Thus, the component corresponding with*

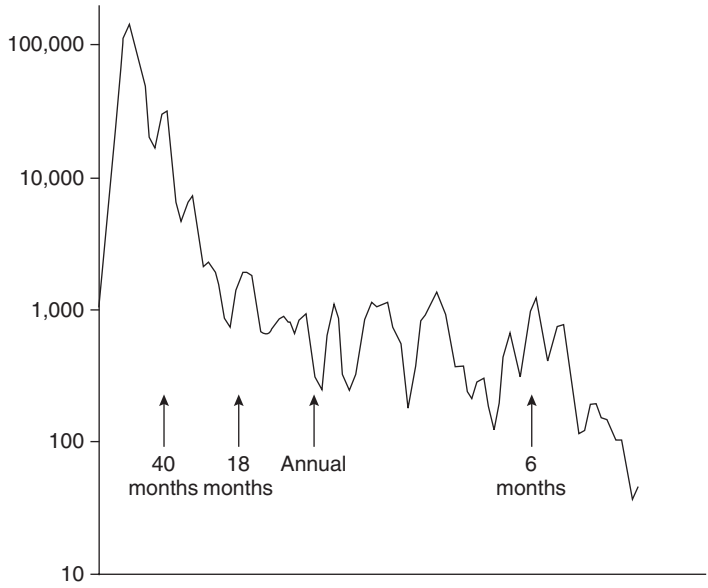


Figure 6.1 Granger and Morgenstern's global risk spectrum of Standard and Poor series, based on annual data, 1875–1952.

the American business cycle of approximately 40 *months, although noticeable, is not of particular importance and is much less pronounced than the components with resonance periods of five years or more (Granger and Morgenstern, 1963).*

Finally, we can now also properly define white noise in spectral terms.

Definition 227 White noise *is defined as any random variable whose risk spectrum or PSD is a constant flat line (= independent of frequency):*

$$P(\omega) = \sigma_{\varepsilon}^2 \tag{6.58}$$

when measured in the 2-dimensional (2D) spectrum-frequency $\{P, \omega\}$*, space.*

Thus, white noise is a very specific kind of noise with a particular characteristic: the *flatness* of its risk spectrum. Therefore, it cannot be considered "general noise," as is often, but erroneously, suggested by statisticians, econometricians, financial analysts, etc. Later on we will find how "general noise" is visualized and measured in both the time and frequency domains.

6.2.6 Average financial risk spectrum of FBM

From Chapter 4, we recall that the FBM provides a useful model for long-term dependent financial time series, whose empirical spectra obey self-similar power

laws of a fractional order $d \in (-0.5, 0.5)$. However, due to the non-stationarity of these financial pricing processes, it is not clear how to measure their time-varying spectra, since the classical measurement by Fourier spectra requires the use of time-averaged measurements based on stationarity. An apparent contradiction exists between the stationarity assumption upon which the usual Fourier spectra are based and the fact that time-varying spectra cannot be associated with stationary processes.

The usual way to approach this difficult measurement issue is as follows. Although the FBM itself is nonstationary, its increments (and hence its derivatives) are stationary. This allows one to associate well-defined average spectral representations with the increments of FBMs.

The *financial risk spectrum or PSD, of the FBM* at frequency zero is the FT of its ACF, which, according to the Wiener–Khinchin Theorem, is the product of the two conjugate FTs of the FBM process, as follows. Recall that the FT of the FBM is modeled by:

$$F_{\mathrm{FBM}}(\omega) = (1 - e^{-j\omega})^{-d} \mathcal{F}[\varepsilon(t)] \tag{6.59}$$

Next, apply the two exponential series expansions for $e^{j\omega}$ and $e^{-j\omega}$, with $j = \sqrt{-1}$, the imaginary number and ω is the angular frequency, and take the limit for $\omega \to 0$. Then we obtain the FT of the ACF of the FBM as follows:

$$\begin{aligned}
P(\omega) &= \mathcal{F}[\gamma(\tau)] \\
&= F(\omega)F(-\omega) \\
&= (1 - e^{-j\omega})^{-d}(1 - e^{j\omega})^{-d}\mathcal{F}[\varepsilon^2(t)] \\
&= \sigma_\varepsilon^2 (1 - e^{j\omega})^{-d}(1 - e^{-j\omega})^{-d} \\
&= \sigma_\varepsilon^2 \left[1 - \left(1 + j\omega + \frac{(j\omega)^2}{2!} + \cdots\right)\right]^{-d} \\
&\quad \times \left[1 - \left(1 - j\omega + \frac{(-j\omega)^2}{2!} - \cdots\right)\right]^{-d} \\
&= \sigma_\varepsilon^2 \left[-j\omega + \frac{\omega^2}{2!} + \cdots\right]^{-d} \left[j\omega + \frac{\omega^2}{2!} + \cdots\right]^{-d} \\
&\sim \sigma_\varepsilon^2 [-j\omega]^{-d}[j\omega]^{-d} \quad \text{as } \omega \to 0 \\
&= \sigma_\varepsilon^2 \omega^{-2d}
\end{aligned} \tag{6.60}$$

Again, for $d = 0$

$$P(\omega) = \sigma_\varepsilon^2 \tag{6.61}$$

the spectral density of white noise. Thus, the spectral density of the anti-persistent FBM increments, or fractionally differenced white noise time series with $d < 0$

Table 6.1 Risk spectrum of FBM increments

Noise characteristic	*Difference*	*Spectrum* $P(\omega)$
Fractionally differenced white noise	$d < 0$	$\sigma_\varepsilon^2 \omega^{-2d} \to 0, \omega \to 0$
White noise	$d = 0$	$\sigma_\varepsilon^2 \omega^{-2d} = \sigma_\varepsilon^2$
Fractionally integrated white noise	$d > 0$	$\sigma_\varepsilon^2 \omega^{-2d} \to \infty, \omega \to 0$

will be zero, when the frequency approaches zero $\omega \to 0$. *Vice versa*, the spectral density of persistent FBM increments, or fractionally integrated white noise with $d > 0$ will be infinite when the frequency approaches zero $\omega \to 0$, as summarized in Table 6.1.

Therefore, neither extreme can be observed in a risk spectrum $P(\omega)$. But, as we've observed in Chapter 4, there are different degrees of fractional integration of white noise in between these two extremes. For fractional differentiation of white noise by $d = 0.5$, the result is blue noise. For fractional integration of the white noise by $d = -0.5$, the result is red noise; for integrating white noise once, $d = 1$, the result is brown noise (= Brownian Motion), and for integrating it once more, $d = 2$, the result is black noise.

Since the financial risk spectrum of the fractional increments of the FBM is scaling and proportional to

$$\omega^{-2d} = \omega^{1-2H} \quad \text{as } \omega \to 0 \tag{6.62}$$

this suggests that the FBM self, which consists of once integrated FBM increments, has a scaling spectral density proportional to

$$\begin{aligned} \omega^{-2(d+1)} &= \omega^{-2(H+0.5)} \\ &= \omega^{-(2H+1)} \quad \text{as } \omega \to 0 \end{aligned} \tag{6.63}$$

since the Hurst exponent $H = d + 0.5$ (cf. Chapter 4).

Remark 228 *Notice that these average financial risk spectra do not depend on time* t*. In other words, the average FBM spectra are* not *time-varying. This fundamental fact results from the stationarity of the FBM increments combined with the linearity of the integration!*

It is also clear that the financial risk spectrum of a scaled FBM, $x(ct)$, is *frequency-scaling* (= characterized by a power law of a fractional order d), which is in accordance to the fact that its second-order moments, represented by the ACF

$\gamma(\tau)$ are *time-scaling*:

$$\gamma(\tau) \sim \sigma_\varepsilon^2 \tau^{2d-1} \tag{6.64}$$

Recall that, according to the scaling property of FTs, there is the transform pair

$$x(ct) \leftrightarrow \frac{1}{|c|} F\left(\frac{\omega}{c}\right) \tag{6.65}$$

Thus, the financial risk spectrum of the scaled FBM is

$$\begin{aligned}
\mathcal{F}[\gamma(\tau)] &= \frac{1}{|c|} F\left(\frac{\omega}{c}\right) \frac{1}{|c|} F\left(\frac{-\omega}{c}\right) \\
&= \frac{1}{|c|^2} P\left(\frac{\omega}{c}\right) \\
&= \begin{cases} \dfrac{\sigma_\varepsilon^2}{|c|^2}\left(\dfrac{\omega}{c}\right)^{-2d} = c^{2d-2}\sigma_\varepsilon^2 \omega^{-2d} = c^{2H-3}\sigma_\varepsilon^2 \omega^{-(2H-1)} \\ \qquad\qquad \text{for the FBM increments, and} \\ \dfrac{\sigma_\varepsilon^2}{|c|^2}\left(\dfrac{\omega}{c}\right)^{-2(d+1)} = c^{2d}\sigma_\varepsilon^2 \omega^{-2(d+1)} = c^{2H-1}\sigma_\varepsilon^2 \omega^{-(2H+1)} \\ \qquad\qquad \text{for the FBM self} \end{cases}
\end{aligned} \tag{6.66}$$

McCulloch *et al.* (2001) provide an alternative approach to spectral measures of the stable distributions of Chapter 3.

6.3 Hurst exponent identification from risk spectrum

It is this frequency scaling property of the FBM, which allows us to compute its financial risk spectrum to determine the Hurst exponent. How? Plot the logarithm of the financial risk spectrum $P(\omega)$ of the FBM against the logarithm of frequency ω:

$$\begin{aligned}
\ln \mathcal{F}[\gamma(\tau)] &= \ln\left[\frac{1}{|c|^2} P\left(\frac{\omega}{c}\right)\right] \\
&= -(2H+1)\ln\omega + [(2H-1)\ln[c] + \ln\sigma_\varepsilon^2] \\
&= -b\ln\omega + C
\end{aligned} \tag{6.67}$$

The slope coefficient of the resulting negative line is

$$b = (2H+1) \tag{6.68}$$

so that

$$H = \frac{b-1}{2} \tag{6.69}$$

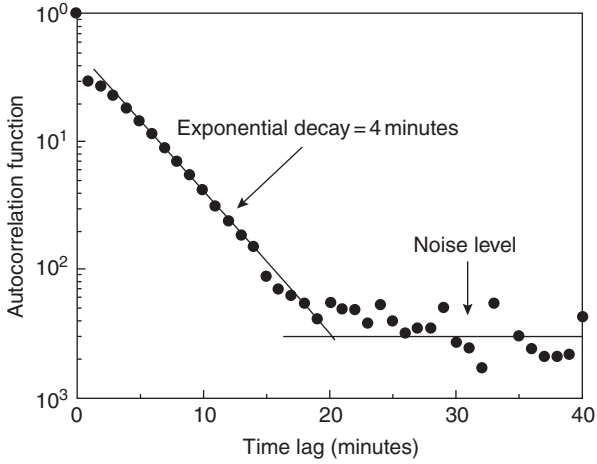


Figure 6.2 Semi-log plot of the autocorrelation function $\gamma(\tau)$ for the S&P500 index, sampled at a 1-minute time scale.

and the intercept is

$$C = (2H - 1)\ln[c] + \ln \sigma_\varepsilon^2 \tag{6.70}$$

With H so computed and the intercept C and the scaling constant c known, we can even find the value of the noise variance σ_ε^2. This could not be done with Mandelbrot's logarithmic plot in Chapter 4. Let's now look at some empirical examples.

Example 229 *Figure 6.2, displayed also in Mantegna and Stanley (2000, p. 55; courtesy of Gopikrishnan* et al.*, 1998), provides the semi-logarithmic plot of the autocorrelation function for the S&P500 index, sampled at the* $\Delta t = 1$ *minute time scale. The straight line corresponds to exponential decay with a characteristic decay time of* $\tau = 4$ *minutes. It is apparent that after about* 20 *minutes the correlations are at the level of pure noise.*

Example 230 *Figure 6.3, displayed in Mantegna and Stanley (2000, p. 56; adapted from Mantegna and Stanley, 1996), shows the spectral density of the S&P500 index, of which high-frequency minute-by-minute data were recorded during the four-year period from January 1984 to December 1987. The empirical behavior of the index is clearly described by the linear slope coefficient* $\lambda/2 = (H-1) \approx -0.5$ *in the time window from approximately* 30 *trading minutes to* 100 *trading days, corresponding with the independent increment case,* $H = 0.5$*. Such a linear slope coefficient is characteristic for the particular financial market investigated. Mantegna and Stanley (2000, p. 56) mention comparable studies*

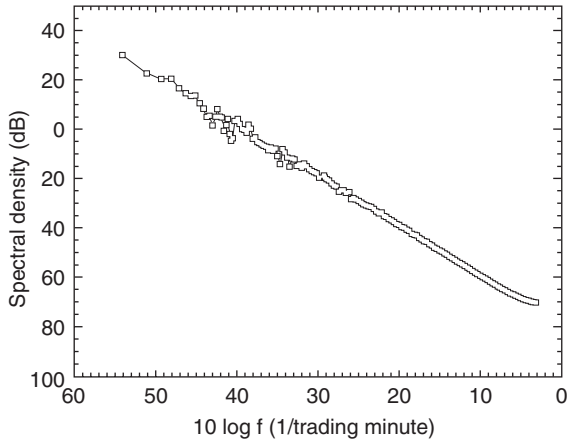


Figure 6.3 Spectral density of high-frequency data from the S&P500 index.

analyzing daily data on stock indices of New York (the New York Composite index), Frankfurt (the DAX index), and Milan (the MIB index) exchanges, with the empirical results of $\lambda = -1.04, -1.06$ *and* -1.14*, or, equivalently,* $H = 0.48,\ 0.47$ *and* 0.43*, respectively (Mantegna, 1991). These values show the presence of weak long-term dependence, in particular, of anti-persistence, since the empirical values of* H *are always slightly lower than* 0.5*. Based on our own empirical research, we concur with their conclusion that the strength of such long-term dependence is dependent on the particular financial market and that it "seems to be larger for less efficient markets." Using high-frequency data for the S&P500 index, one finds that the market's volatility, which is a measure of the market's risk, has two regimes: for very short trading horizons (*$\tau < 30$ *trading minutes), superdiffusive (*$-\lambda/2 < -0.5$*), or anti-persistent (*$H < 0.5$*) behavior is observed. In the longer term (*$30 < \tau < 10^4$ *minutes* $= 167$ *hours* $= 7$ *days) the behavior is closer to diffusive or neutrally persistent (*$\lambda/2 = -0.5$ *or* $H = 0.5$*). In the very short-term* $\lambda/2 = -0.8$ *or* $H = 0.2$ *with strong superdiffusive or ultra-anti-persistent behavior, which borders on chaos. This anti-persistent behavior is most likely due to the fact that in the very short-term the time series has a memory of only a few minutes, as shown in the preceding example. For another example of spectral analysis of heavy-tailed data,* cf. *Mikosch (1998).*

6.4 Heisenberg Uncertainty Principle

The fundamental Uncertainty Principle of Heisenberg states that there exists no time series with finite risk which is *compactly supported* both in the time and frequency domains.[3] Since this principle has important consequences for our financial

market risk analysis (Los, 2000), we will provide a very simple proof using the concepts of *equivalent time duration* and *spectral bandwidth*.

Definition 231 *The* equivalent time duration T_{D} *of* $x(t)$ *is defined by*

$$T_{\mathrm{D}} = \frac{1}{x(0)} \int_{-\infty}^{+\infty} |x(t)|\, dt \tag{6.71}$$

where $x(0) \neq 0$.

Definition 232 *The* equivalent spectral bandwidth W_{B} *of* $x(t)$ *is defined by*

$$W_{\mathrm{B}} = \frac{1}{F(0)} \int_{-\infty}^{+\infty} |F(\omega)|\, d\omega \tag{6.72}$$

where $F(0) \neq 0$.

Proposition 233 (Uncertainty Principle of Heisenberg): *The product of the equivalent spectral bandwidth and time duration of a time series* $x(t)$ *cannot be less than a certain minimum value.*

$$W_{\mathrm{B}} T_{\mathrm{D}} \geq 2\pi \tag{6.73}$$

Proof By definition of T_{D}, we have

$$\begin{aligned} x(0) T_{\mathrm{D}} &= \int_{-\infty}^{+\infty} |x(t)|\, dt \\ &\geq \int_{-\infty}^{+\infty} x(t) dt \\ &= \left[\int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt \right]_{\omega=0} \\ &= F(0) \end{aligned} \tag{6.74}$$

Similarly, by definition of W_B, we have

$$\begin{aligned} F(0) W_{\mathrm{B}} &= \int_{-\infty}^{+\infty} |F(\omega)|\, d\omega \\ &\geq \int_{-\infty}^{+\infty} F(\omega) d\omega \\ &= \left[\int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} d\omega \right]_{t=0} \\ &= 2\pi x(0) \end{aligned} \tag{6.75}$$

Thus, we obtain

$$x(0)T_{\mathrm{D}} \geq F(0) \geq \frac{2\pi x(0)}{W_{\mathrm{B}}} \tag{6.76}$$

from which we conclude that

$$W_{\mathrm{B}} T_{\mathrm{D}} \geq 2\pi \tag{6.77}$$

■

6.5 Windowed FT for transient price innovations

Classical time series analysis devotes most of its efforts to the design of time-invariant and frequency-invariant operators, that modify essentially stationary time series properties. The FT dominates linear time-invariant time series analysis, because the sinusoidal waves $e^{j\omega t}$ are the constant eigenvectors of linear time-invariant difference operators (cf. Chapter 3 for the relevant properties of linear time-invariant operators). As we discussed earlier, this makes it possible to compute the Hurst exponent and thus to establish some indication about the irregularity and the time and frequency scaling of financial time series.

Let's be a bit more precise about this particular aspect of the use of the FT, by defining the characteristic function of the linear time-invariant operator in terms of its eigenvalues, i.e., the solutions of its characteristic function (Mallat, 1999).

Definition 234 *A function satisfying the equation*

$$\Lambda\{x(t)\} = \lambda x(t) \tag{6.78}$$

is called an eigenfunction (*or* characteristic function) *of the operator* Λ*, and the corresponding value of* λ *is called an* eigenvalue *(or* characteristic value*) of* Λ.

Definition 235 *A* linear time-invariant (*convolution*) operator L *is entirely specified by the eigenvalue* $H(\omega)$*, which is the FT of the linear function* h *at the frequency* ω*:*

$$\begin{aligned} Le^{j\omega t} &= \int_{-\infty}^{+\infty} h(u) e^{j\omega(t-u)} du \\ &= e^{j\omega t} \int_{-\infty}^{+\infty} h(u) e^{-j\omega u} du \\ &= e^{j\omega t} H(\omega) \end{aligned} \tag{6.79}$$

The exponential basis $e^{j\omega t}$, which represents a sinusoidal wave, is the eigenvector *of the linear convolution operator L.*

Consequently, the eigenfunction of a linear time-invariant system is an exponential function, because we have the FT pair

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} d\omega \tag{6.80}$$

and

$$F(\omega) = \int_{-\infty}^{+\infty} x(t) e^{-j\omega t} dt \tag{6.81}$$

i.e., the Fourier coefficient obtained by correlating $x(t)$ and $e^{-j\omega t}$, and thus

$$Lx(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) H(\omega) e^{j\omega t} d\omega \tag{6.82}$$

This construction explains the global character of the FT. The linear operator L amplifies, or attenuates, the sinusoidal component $e^{j\omega t}$ of $x(t)$ by the *transfer function* $H(\omega)$, which is the *frequency filter* of $x(t)$. Since the support of the sinusoidal wave $e^{j\omega t}$ covers the whole real line $(-\infty, +\infty)$, the Fourier coefficient $F(\omega)$ depends on the values $x(t)$ for all times $t \in \mathbb{R}$.

It is precisely this global "mix" of information spread over all the times considered by the set of observations why the FT $F(\omega)$ can excellently analyze the *global* frequency contents of the irregularity, or true risk, of $x(t)$, but not its *local* frequency contents of its irregularity. Thus, we can determine the overall risk level of the financial investment return series $x(t)$, but Fourier analysis cannot assist us with the determination of its local risk content, i.e., the market risk level of this financial series at a particular time t. But this is precisely the kind of risk information that fund managers require for proper market risk management by continuous *dynamic hedging*!

On the other hand, when each of the stacked frequencies can be precisely identified, we can produce a frequency or risk analysis that is localized in frequency as well as in time. Such an approach will require the understanding of the *time–frequency localization* of the systematic part of a time series, as already in 1946 had been achieved by Gábor's or Windowed Fourier Analysis (Cohen, 1989; Delprat *et al.*, 1992). In fact, there exist now two time–frequency localization transforms:

(1) the Gábor Transform.
(2) the Wavelet Transform (to be discussed in Chapter 7).

Gábor's Transform or Windowed Fourier Transform (WFT) replaces the FT's infinitely supported sinusoidal wave by the product of a sinusoid and a compact *taper*, which is localized in time (Allen, 1977).

6.5.1 Gábor's WFT

As we noted in the preceding section, the Uncertainty Principle states that there is no finite risk time series $x(t)$ which is compactly supported in both the time and frequency domains. In other words, the risk spread of a variable and its FT cannot be simultaneously arbitrarily small. Motivated by quantum mechanics, the Hungarian physicist Gábor (1946) defined elementary time–frequency "atoms" or "kernels" as wave forms that have minimal spread in the time–frequency plane. He demonstrated the importance of localized time–frequency time series analysis, when he implemented his invention in the form of the first *holograph*.[4] It took until the 1980s before his vision came to complete fruition in Wavelet Transforms and (multidimensional) wavelet multiresolution analysis, discussed in Chapter 7 and as now used in 3-dimensional (3D) image compacting and transmission over the internet.

Definition 236 Gábor atoms (*or* kernels) *are constructed by* time translation *(by period* τ*) and* frequency modulation *(by frequency* ξ*) of the original* time window $g(t)$*:*

$$g_{\tau,\xi}(t) = g(t-\tau)e^{j\xi t} \tag{6.83}$$

such that $\int_{-\infty}^{+\infty} g^2(t-\tau)dtd\tau = 2\pi$.

Notice that a Gábor atom is the product of a sinusoidal wave $e^{j\xi t}$ with a finite risk symmetric window g. Thus, the risk of $g_{\tau,\xi}(t)$ is symmetrically concentrated in the time neighborhood of τ over an interval of size σ_t, measured by the standard deviation of $|g|^2$, and it has a frequency center ξ. The atom $g_{\tau,\xi}(t)$ can be viewed as changing analyzing filters which adapt according to the frequency change in the time series $x(t)$ associated with the frequency ξ in the neighborhood of time horizon τ. The time and frequency spreads of these atoms are constant. The whole family of Gábor atoms is generated by time and frequency translations of one specific atom $g(t)$.

Example 237 *Figure 6.4 shows the Gábor atom* $g_{0,\xi}$ *for three frequencies: a: high frequency* ξ_1*, b: middle frequency* ξ_2 *and c: low frequency* ξ_3*, (Bendjoya and Slézak, 1993, p. 235).*

Definition 238 *The* FT of the Gábor atom *is a frequency translation by* ξ*:*

$$\begin{aligned} G_{\tau,\xi}(\omega) &= \mathcal{F}[g_{\tau,\xi}(t)] \\ &= \int_{-\infty}^{+\infty} g_{\tau,\xi}(t)e^{-j\omega t}dt \\ &= \int_{-\infty}^{+\infty} g(t-\tau)e^{-j(\omega-\xi)t}dt \\ &= G(\omega-\xi)e^{-j(\omega-\xi)\tau} \end{aligned} \tag{6.84}$$

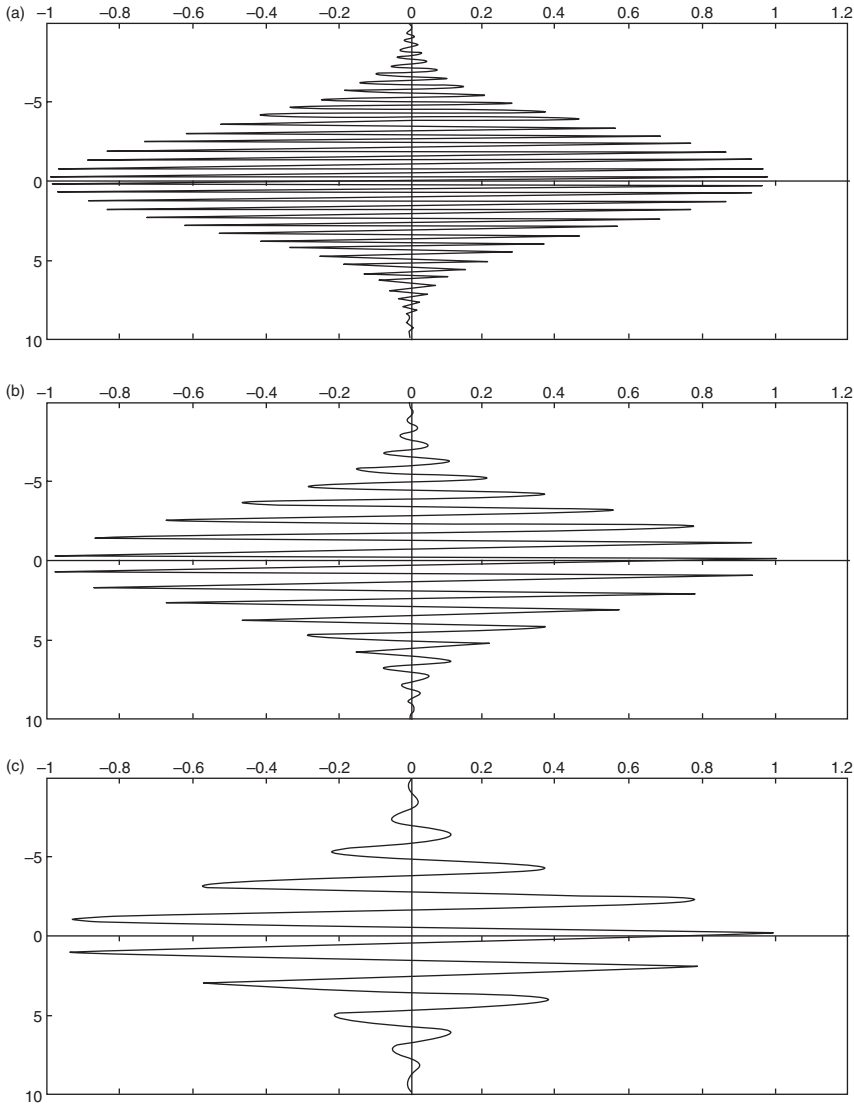


Figure 6.4 Gábor's atom g_{0,ξ_i} as a function of time for three frequencies: (a) high ξ_1, (b) middle ξ_2 and (c) low ξ_3.

Thus, the risk of $G_{\tau,\xi}(\omega)$ is localized near the frequency ξ, over an interval of size σ_ω, which measures the domain where the Gábor resonance coefficient $G(\omega)$ is non-negligible.

The original FT represents a time series as the sum of sinusoidal waves in which the resonance coefficients are correlation coefficients. As we discussed

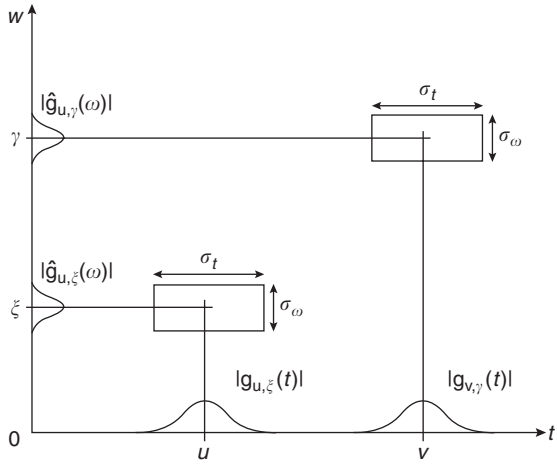


Figure 6.5 Heisenberg boxes of two windowed Fourier atoms $g_{u,\xi}$ and $g_{v,\gamma}$.

earlier, these *sinusoids* are very well localized in frequency, but not in time, since their support has infinite length $(-\infty, +\infty)$. This is a consequence of their exact periodicity. To represent the frequency behavior of a financial time series locally in time, the time series should be analyzed by functions that are localized both in time and frequency, i.e., that are compactly supported in both the time *and* frequency domains, like Gábor's atoms.

In a *time-frequency plane* (t, ω), the risk spread of the atom $g_{u,\xi}$ is measured in the mean squares sense and can symbolically be represented by the *Heisenberg box* illustrated in Figure 6.5. Gábor's Heisenberg box is centered at (u, ξ) and has a time dispersion σ_t and a frequency dispersion σ_ω. Although the shape of this time–frequency box may vary depending on the time width of the window g, the Uncertainty Principle proves that its area satisfies the following inequality

$$\sigma_t \sigma_\omega \geq \frac{1}{2} \tag{6.85}$$

where σ_t is the standard time deviation and σ_ω the standard frequency deviation of a time series $x(t) \in L^2$.

Remark 239 *The area of the Heisenberg box is minimal and this inequality is an equality if and only if the window, kernel or density function, g is Gaussian, in which case the atoms $g_{\tau,\xi}$ are called* Gábor chirps. *In other words, g is a Gábor chirp, if there are constants $(\tau, \xi, c, b) \in \mathbb{R}^2 \times \mathbb{C}^2$ such that*

$$g_{\tau,\xi}(t) = ce^{-b(t-\tau)^2} e^{j\xi t} \tag{6.86}$$

If the time series $x(t)$ is non-zero with a compact support, then its FT in the frequency domain cannot be zero on a whole frequency interval. Similarly, if its

FT is compactly supported, then the time series $x(t)$ cannot be zero on a whole time interval. Hence, even if the Heisenberg constraint is satisfied, it is impossible to have a function in L^2 space, which is compactly supported both in time and frequency domain. This means that there cannot exist an instantaneous frequency analysis for finite risk time series. Thus, time–frequency localization is achievable only in the mean squares sense as visualized by the Heisenberg box.

We can also represent Gábor's Transform as a scalar function, with two arguments, time horizon τ and frequency ξ, as follows.

Definition 240 Gábor's Transform *or* WFT *correlates (= computes the inner product of) the time series $x(t)$ with each Gábor atom $g_{\tau,\xi}(t)$ to produce the following* resonance coefficients:

$$\begin{aligned} G(\tau,\xi) &= \langle x(t), g_{\tau,\xi}(t)\rangle \\ &= \int_{-\infty}^{+\infty} x(t) g^*_{\tau,\xi}(t) dt \\ &= \int_{-\infty}^{+\infty} x(t) g(t-\tau) e^{-j\xi t} dt \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega) G^*_{\tau,\xi}(\omega) d\omega \end{aligned} \tag{6.87}$$

The last equation follows from Parseval's (ergodic) Formula, since we have (without proof):

Theorem 241 (Parseval's Formula)

$$\int_{-\infty}^{+\infty} x(t) h^*(t) dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} F(\omega) H^*(\omega) d\omega \tag{6.88}$$

Remark 242 *For $h = x$ it follows from Parseval's Formula that*

$$\int_{-\infty}^{+\infty} |x(t)|^2 dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |F(\omega)|^2 d\omega \tag{6.89}$$

i.e., Parseval's Theorem.

The original time series $x(t)$ can be reconstructed from Gábor's resonance coefficients by the following double integral:

$$\begin{aligned} x(t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} G(\tau,\xi) g(t-\tau) e^{j\xi t} d\xi d\tau \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x(t) g(t-\tau) e^{j\xi t} g(t-\tau) e^{-j\xi t} dt d\xi d\tau \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x(t) g^2(t-\tau) dt d\tau \end{aligned} \tag{6.90}$$

The time series $x(t)$ can thus be viewed as a sum of localized waves weighted by the profile of the chosen taper or window g. Gábor's Transform has a *constant time–frequency resolution*. This resolution can be changed by rescaling the window g. It is a complete, stable and redundant representation of the systematic part of the time series. Hence, it is invertible and easy to model.

6.5.2 Spectrograms: varying spectral densities

Gábor's WFT can be written both as a time integral and as a frequency integral. Measuring time-varying harmonics is the most important application of WFTs. For example, when listening to the human voice, we perceive sounds with frequencies that vary in time. They are clearly nonstationary: they are modulated over time.

A spectral line of $x(t)$ creates high amplitude windowed Fourier resonance coefficients $S(\tau, \xi)$ at frequencies $\xi(\tau)$ that depend on the time τ. The time evolution of such spectral components is therefore analyzed by following the location of such large amplitude coefficients. These color or grey level coded visualizations of the amplitude resonance coefficients are called spectrograms.

Definition 243 *A* spectrogram *is the squared modulus of the WFT, i.e., the time varying spectral density*

$$P_S(\tau, \xi) = |S(\tau, \xi)|^2 = \left| \int_{-\infty}^{+\infty} x(t) g(t - \tau) e^{-j\xi t} dt \right|^2 \tag{6.91}$$

A spectrogram measures the risk of financial time series $x(t)$ in the time–frequency neighborhood of (τ, ξ) specified by the Heisenberg box of $g_{\tau,\xi}(t)$. This means that we can now measure and visualize the *localized risk* of a financial time series, instead of the average risk.

6.5.3 Examples of spectrograms and sonograms

We will now show a few nonfinancial and financial examples, using these new technologies of signal processing, like the creation of histograms and spectrograms of the increments of financial rate of return series or the increments of foreign exchange rate series.[5] We will conclude that such financial increment series consist of series of singularities or jumps, with random arrival times and with modulated amplitudes, i.e., non-stationarities, like the sonogram of human laughter.

Example 244 *Figure 6.6 presents Gábor's time–frequency analysis using a grey level coding, of the time series presented in Chapter 5. (Adapted from Bendjoya and Slezak, 1993, p. 236.) The highest value of the Gábor resonance coefficients are coded in black and the lowest value in white. The Gábor Transform can detect the frequencies present in the time series $x(t)$ and also their temporal location.*

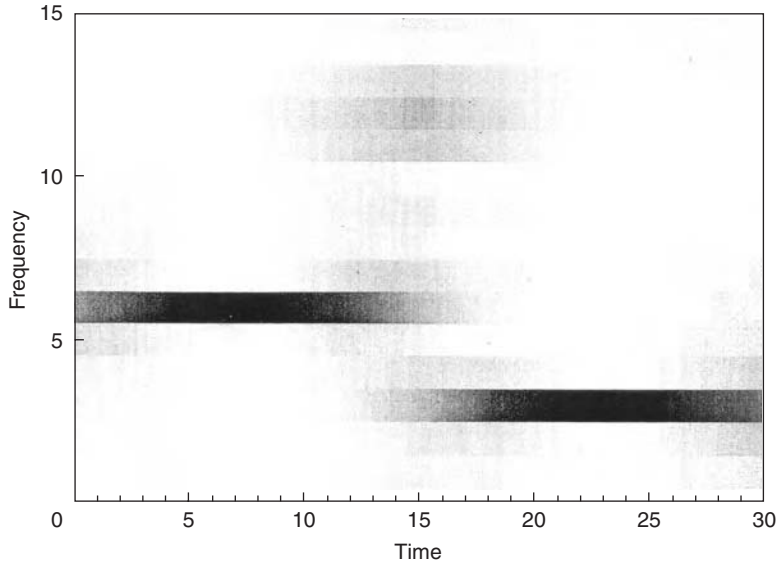


Figure 6.6 Time–frequency analysis by the Gábor Transform with σ adapted to the time coherence of frequencies ω_1 and ω_3. The Gábor resonance coefficients are coded in grey level, with the highest values in black and the lowest values in white. The abscissa represents time and the ordinate the frequency. The highest frequency with the shortest time coherence is worse detected than the other, lower frequencies for which the window size is well suited.

However, it's also clear from this spectrogram that the Gábor Transform has difficulty with detecting frequencies with a coherence shorter than the size of the window. In this spectrogram, the lower coefficients are associated with the highest frequency ω_2 present in the time series with a very short coherence T_2. Such high frequency could be interpreted as "noise" instead of a bona fide systematic signal. The contribution of the short coherence frequency is very easily underestimated. In short, Gábor's WFT is well suited only for financial time series with a coherence at least equal to and preferably larger than the temporal size of the window.

Example 245 *Daubechies (1992, p. 5) shows three spectrograms $P_S(\tau, \xi)$ of the same periodic time series $f(t)$, with only two discontinuities or unit steps $u(t)$ where the arrows are, as in Figure 6.7. The spectrograms show that the basic time series consists of the sum of two sinusoidal time series each with a different frequencies of* 500 *Hz and* 1,000 *Hz, respectively. The two discontinuities are clearly detected and show up as impulses $\delta(t)$ in the spectrograms, cutting through all frequencies. Notice the demonstration of Heisenberg's Uncertainty Principle occasioned by three widths of windows g: in the first spectrogram of panel (b): the emphasis is on precise determination of the basic frequencies of*

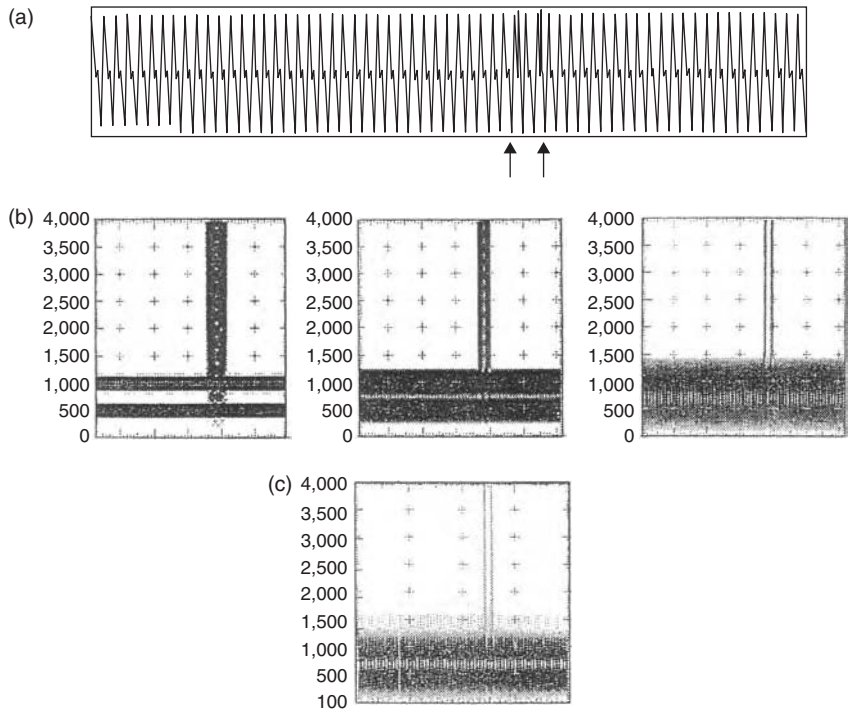


Figure 6.7 Signal, spectrograms and scalogram (a) the signal $f(t)$. (b) Three spectrograms or WFTs $P_S(\tau, \xi)$ of $f(t)$ with three different window widths. Actually $\ln|S(\tau, \xi)|$ is plotted, using grey levels, high values = black, zero = white, intermediate grey levels are assigned proportional to $\ln|S(\tau, \xi)|$ in the time t (abscissa), frequency ω (ordinate) plane. (c) Wavelet Transform $P_W(\tau, a)$ of $f(t)$. To make the comparison with (b) the $|W(\tau, a)|$ is plotted with the same grey level coloration and a linear frequency axis (i.e., the ordinate corresponds to a^{-1}).

the sinusoids, but this blurs the precise localization of the impulses. Going to the right, the precision of the localization of the impulses is increased, but this blurs the precise determination of the basic frequencies. In panel (c): we've an example of a wavelet-based scalogram, which will be discussed in Chapter 7. A scalogram is an excellent devise for time localization of transient events, but is less useful for the precise determination of basic frequencies. Thus spectrograms and scalograms should be used in tandem for a complete time–frequency analysis of financial time series.

Example 246 *Figure 6.8 demonstrates that the spectrogram $P_S(\tau, \xi)$ is a great device to visualize non-stationarity, in particular of time-varying frequencies, or*

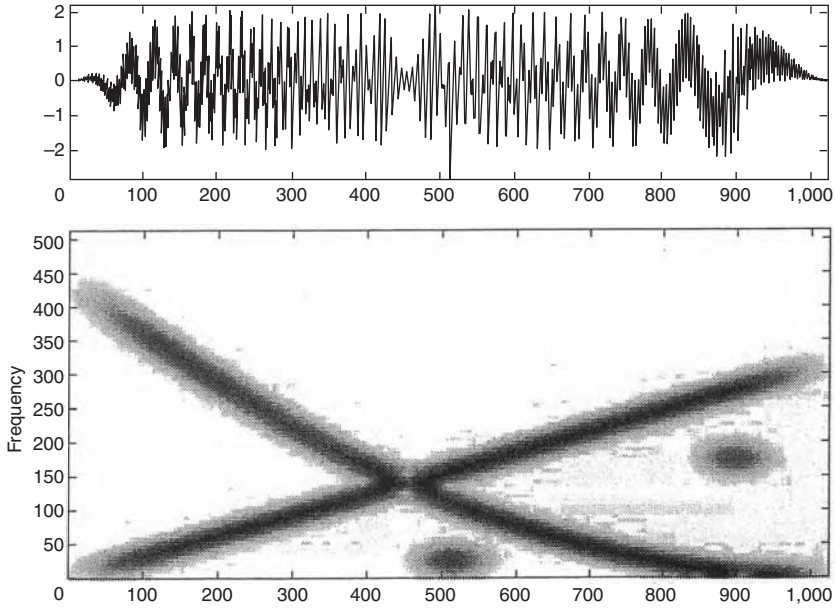


Figure 6.8 Spectrogram $P_S(\tau, \xi)$ of time series with two superimposed time-varying frequencies. Dark areas indicate large amplitude resonance coefficients $|S(\tau, \xi)|^2$.

so-called "chirps." The time series of $T = 1{,}000$ *observations at the top is analyzed in the spectrogram below it. The time series includes a linear chirp, whose frequency increases linearly over time, a quadratic chirp, whose frequency decreases quadratically over time, and two modulated Gaussian noise functions located at* $t = 512$ *and* $t = 896$ *(Mallat, 1999, p. 72).*

Example 247 *Human speech time series are very high frequency series, highly nonstationary and they are known to exhibit frequency- and phase-coupling phenomena, i.e., cross-correlation phenomena over time. The following analysis was performed on a data set of* $T = 1{,}400$ *observations by some of the Higher-Order Spectral Analysis (HOSA) Toolbox MATLAB®-files to illustrate such non-stationarity features (Swami* et al.*, 1998, pp. 1–122/125). Time is measured in milliseconds. Figure 6.9 shows the sonogram of the laughter data in the first panel. The corresponding binned histogram in the second panel shows that its univariate frequency distribution is asymmetrical. The mean, standard deviation, skewness and kurtosis (*cf. *Chapter 1) are computed as* 0.5621, 536.69, 0.1681 *and* 1.3277, *respectively, indicting that these data are non-Gaussian, and that the univariate probability density function (p.d.f.) is not symmetrically distributed. Figure 6.10 shows the spectrogram* $P_S(\tau, \xi)$ *of the laughter chirp. An FFT length of* 512 *observation is used, so that the (Hanning) taper* g *has length* $512/2 = 256$, *with an*

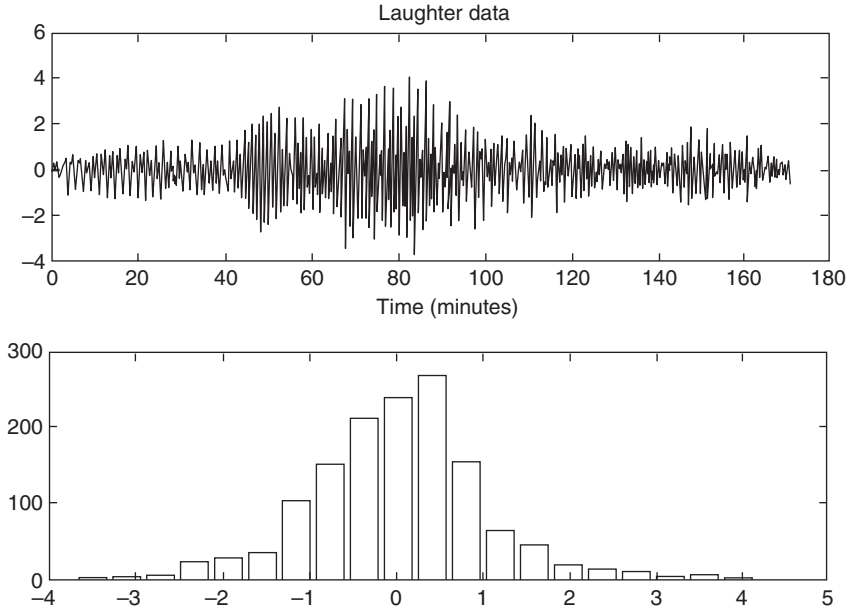


Figure 6.9 Laughter data and their global histogram.

overlap of 240 *observations. The spectrogram shows three dominant frequency tracks, approximately around* 550, 1,100 *and* 1,550 *Hz. The last frequency track or formant begins around* 30 *milliseconds. Additional fragments are visible around* 1,800 *Hz and* 2,100 *Hz. The spectrogram indicates that human laughter is essentially harmonic and that its frequencies appear to be approximately harmonically related. However, not all its fundamental frequencies occur at the same time!*

Example 248 *Figure 6.11 shows that data for the changes in the three-month Treasury Yield contains numerous one-day spikes, which strongly suggests that such series are not continuous, but that they are highly discontinuous (Chapman and Pearson, 2001, p. 86). They can be characterized as series of singularities with modulated amplitudes, i.e., a nonstationary jump process, with random arrival times of these jumps. These financial time series of changes in short-term cash rates of Figure 6.11 look very much like the sonogram of the high frequency nonstationary laughter data in Figure 6.9. However, the series of yield increments are unevenly distributed series of modulated singularities, while the laughter data is a sonogram with a continuous sound wave, consisting of superimposed sound waves of a few fundamental frequencies. This small, but crucial difference is not observable from the binned frequency distributions, which look very much alike. Under close inspection it is somewhat visible in the time domain. But it would be most*

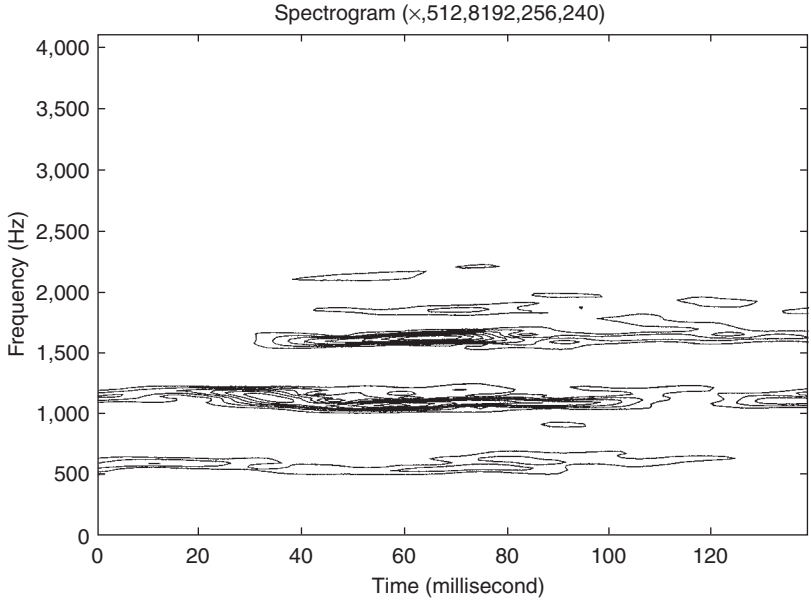


Figure 6.10 Spectrogram of laughter data with three dominant harmonics, approximately around 550 Hz, 1,100 Hz and 1,550 Hz. The last coherent harmonics begins around 30 milliseconds. Additional harmonic fragments are visible around 1,800 Hz and 2,100 Hz. Notice how much more (local) information a spectrogram displays than the statistician's global histogram in Figure 6.9.

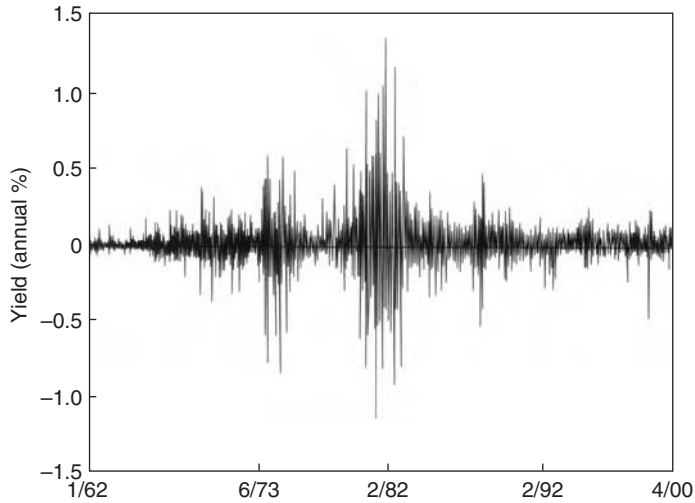


Figure 6.11 Changes in the daily level of the three-month Treasury yield, daily data, January 1962 to April 2000.

clearly visible in the time–frequency spectrogram. We've not been able to obtain this particular yield data set and to produce a spectrogram of it. But in the following example of foreign exchange rate (FX) increments we were able to analyze a very similar data set. It will show that FX increments are not like laughter, since there are no natural harmonics identifiable in the data. The FX increments consist of series of singularities (for which we can compute and visualize a singularity spectrum). It is a clear example of the situation where the frequency distributions of the FX increments and of laughter are similar, but where their respective time distributions are completely different.

Example 249 *Figure 6.12 shows the comparison between the modulated spectrogram* $P_S(\tau, \xi) = |S(\tau, \xi)|^2$ *of minute-by-minute FX data compared with the flat spectrogram of white noise. We compare* 10,800 *increments of the empirical German Deutschemark DEM/US Dollar (USD) rate in the first week of June 1997 in the left panel with the same number of observations on simulated white noise in the right panel with the same constant variance as the FX increments. The frequencies are standardized between* 0 *and* 1 *and are measured by the vertical axis, while the time intervals are measured by the horizontal axis. The lowest frequencies are at the bottom and the highest frequencies at the top. Notice, first, that the German FX increments have low financial risk (light grey and white) in the low frequencies and high financial risk (dark grey and black) in the high frequencies: the risk spectrum is modulated over time. The series is nonstationary. In contrast, the white noise has a constant financial risk over all frequencies and is evenly distributed over time: it is clearly stationary and has a flat spectrum. The financial risk of the German FX increments is intermittently distributed over time. The financial risk*

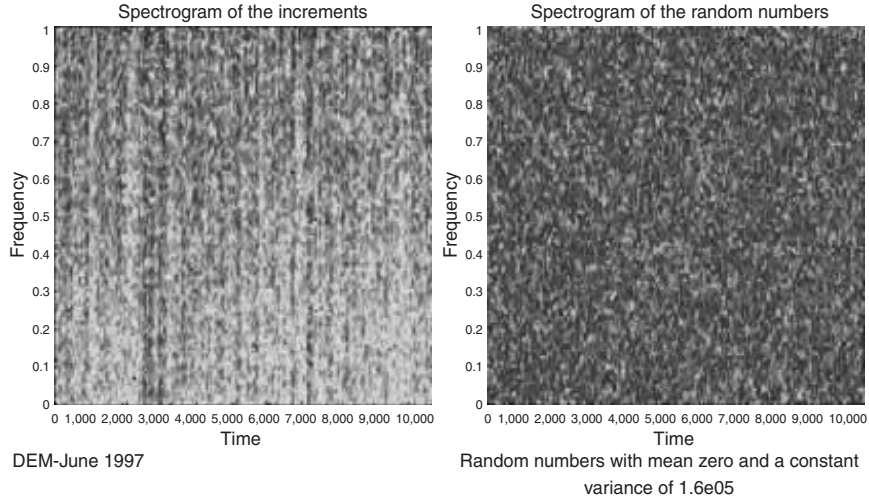


Figure 6.12 Comparison of the modulated spectrogram of empirical DEM/USD increments with the flat spectrogram of white noise.

of the white noise is evenly distributed over time. Thus, to emphasize, this time–frequency visualization shows that the German FX increments in Figure 6.12 are nonstationary, or modulated: their frequency distribution varies over time. Moreover, the German FX increments are unevenly distributed over time. In contrast: white noise is stationary, or unmodulated, and evenly distributed over time. This empirical analysis clearly calls into question the use of the GBM for the modeling of FX prices. In Chapter 8 we will see that scalograms improve the time location of the spectrograms and that such scalograms contain information about these singular FX increments, which can be extracted from the scalogram/spectrogram by modeling the FX process as an FBM.

6.6 Software

The computations of the following Exercises can be executed by using Benoit 1.3: Fractal System Analysis for Windows, Trusoft International Inc., 204 37th Ave. N #133, St Petersburg, FL 33704 Tel (813) 925-8131; Fax (813) 925-8141; sales@trusoft-international.com. See http://www.trusoft-international.com for details.

They can also be executed by using the MATLAB® Signal Processing Toolbox and the MATLAB® HOSA Toolbox (Swami *et al*., 1998). Both Toolboxes are available from The MathWorks, Inc., 24 Prime Park Way Natick, MA 01760-1500, USA. Tel (508) 647-7000; Fax (508) 647-7001; http://www.mathworks.com/products/wavelettbx.shtml.

The HOSA Toolbox (2.0.3) of Nikias and Petropulu (1993) is a collection of MATLAB® M-files containing specialized tools for signal processing with higher order spectra. It was created by Jerry M. Mendel, Chrysostomos L. (Max) Nikias and Ananthram Swami of by United Signals and Systems, Inc. The toolbox is a collection of MATLAB® routines whose primary features are functions for: higher-order spectrum estimation either by conventional or parametric approaches; magnitude and phase retrieval; adaptive linear prediction; harmonic retrieval and quadratic phase coupling; time-delay estimation and array signal processing.

6.7 Exercises

Exercise 250 *Using the Benoit 1.3 software, compute the Financial Risk Spectrum for (1) the original share prices of Chapter 1, (2) their total rates of return, and (3) the first differences of the rates of total return. Use initially simple averaging of the spectrum. Then try tapering with a rapid-decay function to avoid Gibbsian edge effects and to minimize spectral leaking. Spectral leaking is the phenomenon that financial risk is added to the estimate at some wave number from neighboring wave numbers. In that case, there is no sharp discrimination between the risk levels at different frequencies. Spectral leaking is a particular concern if the spectrum is very red, i.e., if the long-wavelength components have greater risk than the short-wavelength components. This is always the case for self-affine traces, like the globally dependent financial time series, like we discuss in this book.*

Exercise 251 *Run MATLAB® Help, Help Desk (HTML), Online Manuals (in PDF). In the Signal Processing Toolbox User's Guide, study and implement the examples in MATLAB® for the DFT (pp. 1/43–1/45), Spectral Analysis (pp. 3/5–3/11), and FFT-based Time –Frequency Analysis (pp. 4 and 26). Find out what is a risk spectrum or PSD (= power spectral density)? What is a CSD (= cross-spectral density)? What is a periodogram? What is the Nyquist rate? What is a spectrogram?*

Run the following MATLAB® Help, Examples and Demos, Toolboxes, HOSA, Case Studies:

Exercise 252 *The classical sunspot data series (annual,* 1700–1987*): notice the two representations of the data, as time series and as histogram. Why is differencing helpful? Interpret the summary statistics of the histogram: mean, variance, skewness and kurtosis. How and why do we compute a singular value plot? Interpret the various representations of the risk spectrum and the harmonic analysis. Are the sunspot data periodic or cyclic? If periodic, what is their period? If cyclic, why? What's the difference between periodicity and cyclicity? Is the business cycle of GDP returns periodic or cyclic?*

Exercise 253 *Canadian lynx data (annual, 1821–1934): do the same as for the sunspot data.*

Exercise 254 *Speech data (laughter): this is an example of high-frequency (HF) data. What is a spectrogram? How do you interpret it? How can you determine if a series is stationary and why is that important? How do you determine if data is harmonic? How do you determine if data is Gaussian? What is linearity and how do you determine it (*cf. *Chapter 3)? What is a bi-spectrum and why it is useful? What is a cumulative spectrum? What is quadratic frequency coupling? What is self-coupling? What other terms don't you know and do you need to define?*

Exercise 255 *Run MATLAB® Fourier analysis on a data set of one day of High Frequency Foreign Exchange (HFFX tick-by-tick and minute-by-minute) data for some Asian currencies, available from the author, or from an international commercial, currency-trading bank like ABN-AMRO. Compute histograms and their summary statistics, risk (power) spectra and cumulative spectra. Conduct harmonic analysis on all series and determine the Fourier "signature,"or "fingerprint,"of each of the nine series, using MATLAB®'s programming facilities (You must also figure out how the EXCEL Link works to feed the raw input data in). Compute a spectrogram. Determine if the series are stationary, harmonic and Gaussian or not. Compute a bi-spectrum and determine if the HFFX series are linear.*

Notes

1 The American mathematician Norbert Wiener (1894–1964), is best known for his development of an interdisciplinary approach to the study of communication and control

processes in living organisms and machines, for which he coined the word *cybernetics*. He was doing work on automated control systems for anti-aircraft guns during the Second World War, when he wrote his famous "yellow peril" report on the optimal tracking solution in the frequency domain. Wiener also contributed to the theory of stochastic processes and the theory of Brownian motion, by constructing a rigorous mathematical description of physical processes that are subject to random change. He helped build the mathematics department at the Massachusetts Institute of Technology (MIT) into an outstanding research facility and taught there from 1919 to 1960.

2 Before Aleksandr Yakovlevich Khinchin graduated in 1916 he had already written his first paper on a generalization of the Denjoy integral. This paper began a series of publications by Khinchin on the properties of functions retained after deleting a set of density zero (probability zero) at a given point. He summarized these results in *Fundamental Mathematica* in 1927. In that same year he became a professor at Moscow University and published his *Basic Laws of Probability Theory*. Between 1932 and 1934 he laid the foundations for the theory of stationary random processes, culminating in a major paper in *Mathematische Annalen* in 1934. At Moscow University, Khinchin build the influential school of probability theory, together with Kolmogorov and Gnedenko. From the 1940s on he was interested in the theory of statistical mechanics and he helped to develop Shannon's ideas on information theory. Khinchin published his famous *Mathematical Principles of Statistical Mechanics* in 1943 and in 1951 he extended it into his *Mathematical Foundations of Quantum Statistics*. It included his fundamental treatment of *local* limit theorems for sums of identically distributed random variables.

3 German theoretical physicist Werner Karl Heisenberg (1901–1976) was one of the leading scientists of the twentieth century. The physical principles underlying the mathematics of quantum mechanics remained mysterious until 1927, when Heisenberg – following conversations with Bohr and Albert Einstein – discovered the uncertainty principle. An important book of Heisenberg published in 1928, *The Physical Principles of Quantum Theory*, described his ideas. The previous year he had become a professor at the University of Leipzig, and in 1932 he was awarded the Nobel Prize for physics. He remained in Germany during the Nazi period and became director of the Kaiser Wilhelm Institute, also heading the unsuccessful German nuclear weapons project. In 1958, Heisenberg became director of the Max Planck Institute for Physics and Astrophysics. He spent his later years working toward a general theory of subatomic particles.

4 The British scientist and inventor Dennis Gábor was born in Budapest, Hungary (1900–1979) and won the Nobel prize for physics (1971) for his invention (in 1947) and later development of holography, a means of numerically producing 3D photographic images without using a lens. Gábor began his career as an industrial research engineer in Germany but went to England with the rise of the Nazis in 1933. He began teaching in 1949 at the Imperial College of Science and Technology in London and became professor of applied electronic physics in 1958. In 1968, he was appointed staff scientist at CBS Laboratories in Stamford, Connecticut and stayed in the United States.

5 A recent interesting example of the application of the WFT in options markets is Benhamou (2002) of Goldman Sachs International, Fixed Income Strategy, Swaps, Division in London, who is impressed by the non-lognormal densities of discrete Asian options and examines the effects of fat-tailed distributions on price as well as on the delta. Using this technology he finds that fat tails lead to larger jumps in the (hedging) delta.

Bibliography

Allen, J. B. (1977) "Short-Time Spectral Analysis, Synthesis and Modification by Discrete Fourier Transform," *IEEE Transactions ASSP*, **25**(3), 235–238.

Bendjoya, Ph., and E. Slézak (1993) "Wavelet Analysis and Applications to Some Dynamical Systems," *Celestial Mechanics and Dynamical Astronomy*, **56**, 231–262.

Benhamou, Eric (2002) "Fast Fourier Transform for Discrete Asian Options," *The Journal of Computational Finance*, **6**(1), 49–68.

Champeney, D. C. (1990) *A Handbook of Fourier Theorems*, Cambridge University Press, Cambridge, UK.

Chapman, David A., and Neil D. Pearson (2001) "Recent Advances in Estimating Term-Structure Models," *Financial Analysts Journal*, **57**(4), 77–95.

Cohen, L. (1989) "Time – Frequency Distributions: A Review," *Proceedings of IEEE*, July, 941–981.

Cooley, J. W., and J. W. Tukey (1965) "An Algorithm for the Machine Calculation of Complex Fourier Series," *Mathematics of Computation*, **19**(90), 297–301.

Daubechies, Ingrid (1992) *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA. (Notes from the 1990 CBMS-NSF Conference on Wavelets and Applications, at Lowell, MA.)

Delprat, N., B. Escudić, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian and B. Torrésani (1992) "Asymptotic Wavelet and Gábor Analysis: Extraction of Instantaneous Frequencies," *IEEE Transactions in Information Theory*, **38**, 644–664.

Duhamel, P., and M. Vetterli (1990) "Fast Fourier Transform: A Tutorial Review and a State of the Art," *Signal Processing*, **19**(4), 259–299.

Gábor, D. (1946) "Theory of Communications," *Journal of the Institute of Electrical Engineering, London III*, **93**, 429–457.

Granger, C. W. J., and O. Morgenstern (1963) "Spectral Analysis of New York Stock Exchange Prices," *Kyklos*, **16**, 1–27. Reprinted as chapter 8 in Cootner, Paul H. (Ed.) (1964) *The Random Character of Stock Market Prices*, The MIT Press, Cambridge, MA, pp. 162–188.

Hsu, Hwei P. (1984) *Applied Fourier Analysis*, Harcourt Brace College Publishers, San Diego, CA.

Gopikrishnan, P., M. Meyer, L. A. N. Amaral and H. E. Stanley (1998) "Inverse Cubic Law for the Distribution of Stock Price Variations," *European Physics Journal*, B, **3**, 139–140.

Körner, T. W. (1990) *Fourier Analysis*, Cambridge University Press, Cambridge, UK.

Los, Cornelis A. (2000) "Frequency and Time Dependence of Financial Risk," *The Journal of Performance Measurement*, **5**(1), 72–73.

Mallat, Stéphane (1999) *A Wavelet Tour of Signal Processing*, 2nd edn, Academic Press, Boston, MA.

Mantegna, R. N. (1991) "Lévy Walks and Enhanced Diffusion in Milan Stock Exchange," *Physica*, A, **179**, 232–242.

Mantegna, Rosario, N., and H. Eugene Stanley (1996) "Turbulence and Financial Markets," *Nature*, **383**, 587–588.

Mantegna, Rosario, N., and H. Eugene Stanley (2000) *An Introduction to Econophysics: Correlation and Complexity in Finance*, Cambridge University Press, Cambridge, UK.

McCulloch, J. Huston, John P. Nolan and Anna K. Panorska (2001) "Estimation of Stable Spectral Measures," *Mathematical and Computer Modelling*, **34**, 1113–1122.

Mikosch, T. (1998) "Periodogram Estimates from Heavy-Tailed Data," in Adler, Robert J., Raisa E. Feldman and Murad S. Taqqu (Eds) (1998) *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Birkhäuser, Boston, MA, pp. 241–257.

Nikias, Chrysostomos L., and Athina P. Petropulu (1993) *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework*, PTR Prentice Hall, Englewood Cliffs, NJ.

Priestley, M. (1988) *Non-Linear and Non-Stationary Time Series Analysis*, Academic Press, San Diego, CA.

Swami, Anathram, Jerry M. Mendel and Chrysostomos L. (Max) Nikias (1998) *Higher-Order Spectral Analysis Toolbox User's Guide*, Version 2, January, The MathWorks, Inc., Natick, MA.

7 Wavelet time–scale analysis of risk

7.1 Introduction

In this chapter we continue to simultaneously analyze the marginal distributions and the temporal dependence of investment returns, as in Chapter 6, but we do it in a time–scale frame of reference, instead of in a time–frequency frame of reference. *Scale* is proportional to the inverse of frequency: $a \sim 1/\omega$.

Our basic model of analysis for the investment returns and foreign exchange rates is again the Fractional Brownian Motion (FBM) presented earlier in Chapter 4. In that chapter we discussed the analysis of stationary and of slowly varying nonstationary financial time series. In this chapter we discuss the analysis of financial time series that contain numerous transient, nonstationary characteristics, such as drifts, trends, discontinuities in higher derivatives of the series, the beginnings and ends of particular events, as well as the self-similarity and scaling exhibited by the FBM.

As we discussed in Chapter 5, classical Fourier analysis is *periodic* wave analysis. It expands signals or functions (of time) in terms of sinusoidal basis functions, or, equivalently, in terms of complex exponentials. Therefore, it is especially suited for the harmonic analysis of periodic, time-invariant, or stationary phenomena. Next, in Chapter 6, we approached the analysis of nonstationary phenomena in a time–frequency frame of reference by breaking the data set up into a sequence of finite horizon "windows," and implemented Fourier analysis to each consecutively overlapping window, in a moving average fashion. The problem with the windowed approach is that the Gábor–Fourier Transforms (FTs) are still not strictly localized. This non-localization leads to approximation, "time-smearing," and thus some time ambiguity of the analytic results.

In contrast, the wavelets discussed in this chapter comprise a complete set of finite basis functions, precisely localized in both time and frequency (or scale), which, in linear resonance combinations, can provide an extremely flexible, efficient and complete representation of a time series.

These wavelet basis functions have their risk concentrated in time. When correlated with a time series, the magnitudes of the resulting wavelet resonance or correlation coefficients provide a tool for the analysis of nonstationary, transient, rapidly or sharply developing dynamic processes (Wang, 1995; Ogden and Parzen, 1996). Such nonlinear dynamic phenomena often incorporate scaling behavior.

Wavelets are very good tools for detecting, quantifying and modeling scaling behavior at various resolutions. Thus, wavelet Multiresolution Analysis (MRA) is an improvement over Gábor's Windowed Fourier Analysis. By using non-overlapping, scaling and shifting windows wavelet MRA localizes the significant resonance correlations accurately, both in time and in frequency.

This chapter is heavily indebted to signal processing engineers and mathematicians like Mallat (1989a,b,c; 1999), Bruce *et al*. (1996), Burrus *et al*. (1998), Burke-Hubbard (1994, 1998), Flandrin (1989, 1992), Jawerth and Sweldens (1994), Kaplan and Kuo (1993), Rioul and Vetterli (1991), and Strang (1994). Holschneider (1995) of the Centre National de la Recherche Scientifique (CNRS = National Center for Scientific Research) in Marseille, France, and Cohen and Kovacevic (1996) provide excellent and detailed mathematical overviews of wavelet analysis, as do Benedetto and Frazier (1994). The Wavelet Transform (WT) was first introduced by Morlet *et al*. (1982) to obtain time–frequency information from seismic time series. The two (already) classic texts in wavelet theory are written by two mathematical founders of the subject: Yves Meyer (1990) and Ingrid Daubechies (1992). Meyer's book requires a research-level background in mathematics, but Daubechies' text is acessible to a somewhat wider audience. Simpler introductions to wavelets, using only linear algebra, can be found in Chui (1992a) and in Frazier (1999).

In the meantime, wavelet analysis and its applications have become a truly interdisciplinary research methodology. Excellent first introductions to bridge the still existent chasm between signal processing and statistical analysis using wavelets can be found in Chui (1992b) and in Ogden (1997). In finance, this ingenious wavelet analysis has already provided us with the tools to identify and, perhaps even forecast, the pricing and trading processes that characterize our financial markets. The first instances of this new empirical analysis are only now slowly emerging in the financial literature, although much confusion remains regarding what the best ways are to use these powerful analytic signal processing tools in finance (Jensen, 1997). This situation is not unlike when the first instances of econometrics appeared in economics 60 years earlier.

But there is no doubt that wavelet MRA is extremely powerful and will lead to many new discoveries in both finance and economics (cf. Ramsey *et al*., 1995; Ramsey and Zhang, 1996; Ramsey and Zhang, 1997; Aussem *et al*., 1998; Ramsey and Lampart, 1998a,b; Gençay *et al*., 2001), as it already has in medical and biomedical, seismic and oceanographic signal and image processing, in quantum mechanics and asteroid family identification from cluster analysis, meteorology and turbulence research (Ruskai *et al.*, 1992; Meyers *et al*., 1993; Lau and Weng, 1995).

Wavelets have been used to solve serious electronic communications problems and, combined with fractals, they have been applied to time series that are chaotic, as we will discuss in Chapter 8. Since wavelets are self-similar and scaling, there is a natural affinity between wavelet MRA and fractal models, in particular in the research of the self-similar cascading risk levels of the vortices in turbulence research (Massopust, 1994; Wornell, 1995), as we will see in

Chapter 11. Moreover, the use of wavelets as basis functions for the discretization and numerical solution of nonlinear diffusion equations (e.g. used in the valuation and dynamic hedging of American and exotic options) have already achieved excellent success (Bendjoya and Slezak, 1993; Meyer, 1993).

7.2 Wavelet analysis of transient pricing

Wavelet analysis has a diverse historical background. In reflection seismology in the 1970s and 1980s, Morlet *et al*. (1982) found that modulated pulses sent underground have a time duration that is too long at high frequencies to separate the reflections of fine, closely spaced layers of rock, because of the Heisenberg Uncertainty Principle (cf. Chapter 6). Instead of emitting pulses of equal time and frequency duration, he then thought of sending shorter waveforms at high frequencies. Such waveforms are obtained by scaling a single basis function, called a (Morlet) wavelet.[1]

Alex Grossman of the Marseille Theoretical Physics Center recognized in Morlet's approach some ideas that were close to his own analysis of coherent (= correlating) quantum states (Grossman and Morlet, 1984). Thus, nearly forty years after Gábor *et al*., reactivated a collaboration between theoretical physics and signal processing, this ultimately led to the formalization of the Continuous Wavelet Transform (CWT).

The basic ideas of wavelet time–scale analysis were already familiar to mathematicians and engineers working with the harmonic Fourier analysis discussed in Chapters 5 and 6. Thus, the acceptance of wavelets was rather rapid within the community of signal processing engineers. Wavelet analysis is now invading other applied fields in cognitology, biology and medicine, like computer vision, machine sensors, neurology, e.g., the study of electroencephalographs (EEGs) to find extreme brain waves, and in cardiology, e.g., the study of electrocardiograms (ECGs) to identify cardiac arrhythmias (Aldroubi and Unser, 1996).[2]

Wavelet analysis has now also reached the financial markets, to determine the periodicities, aperiodic cyclicities, intermittencies and arrhythmias of the financial time series produced in great abundance by these markets. It's our expectation that a study of the spectrograms and scalograms of the financial markets can assist us with their financial and economic diagnosis to prevent financial crises and other market inefficiencies (Jensen, 1997).[3]

The specific mathematical methods of wavelet analysis have been developed mainly by the French mathematician Yves Meyer (1985, 1993) and his colleagues. Complete wavelet MRA was discovered by Stéphane Mallat in 1988 (Mallat, 1989a,b). Since then, research on wavelets has become truly international. It is particularly active in the United States, where it is led by the work of mathematicians and scientists such as Ingrid Daubechies at Rutgers University and AT&T Bell laboratories, and Ronald Coifman and Victor Wickerhauser at Yale University (Coifman and Wickerhauser, 1992; Wickerhauser, 1994; Buckheit and Donoho, 1995).

After a lapse of more than thirty years, the thread of analyzing the fractality or self-affinity of financial-economic time series was picked up by Ramsey and Zhang (1996, 1997) at the Courant Institute of New York University by implementing wavelet analysis. It is currently a wide open field of research, ripe for a more complete exploration by students in finance and economics (Gençay *et al.*, 2001).

7.2.1 Wavelet Transform

Let's see how wavelet analysis works. Similar to the windowed Fourier Transform (WFT), the WT decomposes a 1-dimensional (1D) time series into 2-dimensional (2D) time–scale ($\sim$ frequency^{-1}) space. In particular, while Fourier analysis breaks down a time series of investment returns into constituent orthogonal sinusoids of different frequencies (= constant periodicities), wavelet analysis breaks down such a time series into constituent orthogonal wavelets of different scales.

Similar to the Gábor Transform, the WT replaces the basic sinusoidal waves of the FT by a family of basic wavelets generated by translations and dilations of one particular wavelet atom. Figure 7.1 compares an infinite sine wave basis for Fourier analysis with a finite Daubechies(20) wavelet basis.

Definition 256 *A* continuous wavelet atom $\psi_{\tau,a}(t)$ *is a wave function of zero average, centered around amplitude zero, with finite risk:*

$$E\{\psi_{\tau,a}(t)\} = \int_{-\infty}^{+\infty} \psi_{\tau,a}(t)dt = 0 \tag{7.1}$$

which is translated *by a limited time interval* τ *and* scaled*, or dilated, by a scale parameter a as follows:*

$$\psi_{\tau,a}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-\tau}{a}\right) \tag{7.2}$$

This scaled and translated wavelet is time centered around τ, like the Gábor atom. If the frequency center of ψ is η, then the frequency center of the dilated

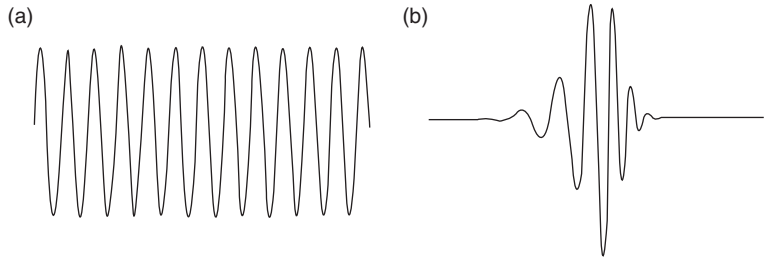


Figure 7.1 (a) A sine wave and (b) a Daubechies' wavelet ψ_{D20}.

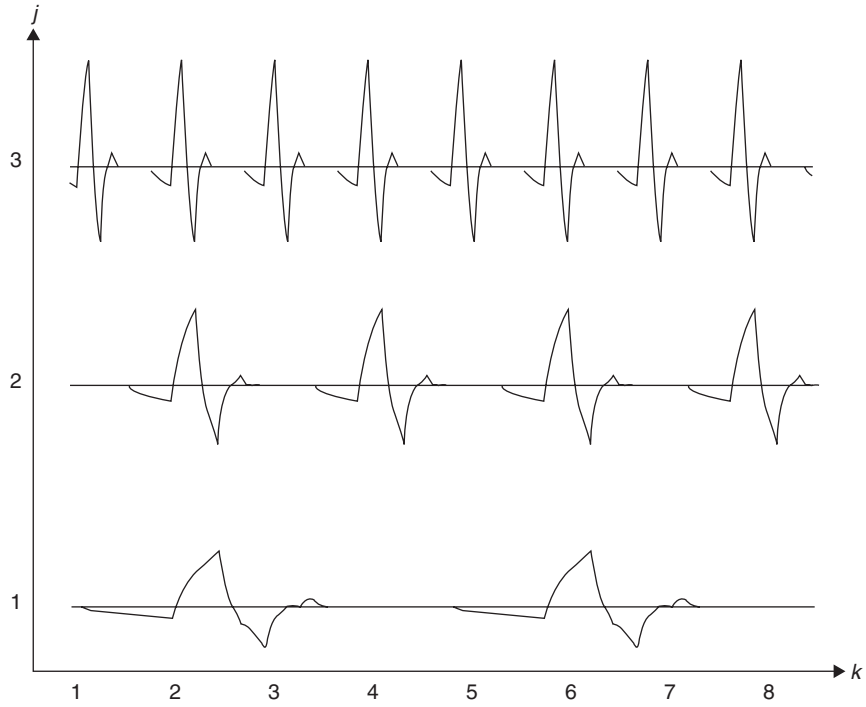


Figure 7.2 Self-similarity of wavelets: translation (every fourth k) and scaling of a wavelet ψ.

wavelet is $\xi = \eta/a$ (thus the scale $a \propto 1/\xi$). Figure 7.2 shows the translation and scaling of a Daubechies(4) wavelet, which we'll define later in this chapter. A continuous wavelet is simply a finite risk function with a zero mean. Besides its scaling and dilating Heisenberg box, the most important feature of a wavelet is the number of its vanishing moments:

$$\int_{-\infty}^{+\infty} t^r \psi(t) dt = 0 \quad \text{for } 0 \leq r < n \tag{7.3}$$

This vanishing moments property of wavelets makes it possible to analyze the local regularity of a time series $x(t)$. A theorem characterizes fast decaying wavelets with r vanishing moments as the rth derivatives of a fast decaying function (cf. Chapter 6 for such testing functions of rapid decay, or *tapers*). We will meet these fast decaying wavelets again in Chapter 8, when we discuss the crucial Lipschitz irregularity analysis.

Usually wavelet analysis is done by orthonormal wavelets, to effectuate the completeness, or exhaustiveness of the analysis.

Definition 257 *An* orthogonal wavelet $\psi_{\tau,a}(t)$ *is a wavelet with the orthogonality property*

$$\int_{-\infty}^{+\infty} \psi_{\tau,a}(t)\psi_{v,b}(t)dt = 0 \quad for\ \tau \neq v \ \ or \ \ a \neq b \tag{7.4}$$

Definition 258 *An* orthonormal wavelet $\psi_{\tau,a}(t)$ *is an orthogonal wavelet with the normalization*

$$\int_{-\infty}^{+\infty} \psi_{\tau,a}(t)\psi^*_{\tau,a}(t)dt = \int_{-\infty}^{+\infty} |\psi_{\tau,a}(t)|^2 dt = 1 \tag{7.5}$$

These few introductory definitions enable us now to define the CWT, which forms the basis for wavelet MRA.

Definition 259 *The* CWT *of* $x(t)$ *at position* τ *and scale* a *is an inner product computed by correlating (or convoluting) the time series* $x(t)$ *with a wavelet atom*

$$\begin{aligned} W(\tau, a) &= \int_{-\infty}^{+\infty} x(t)\psi^*_{\tau,a}(t)dt \\ &= \int_{-\infty}^{+\infty} x(t)\frac{1}{\sqrt{a}}\psi^*\left(\frac{t-\tau}{a}\right)dt \\ &= x(t) \star \psi_{\tau,a}(-t) \end{aligned} \tag{7.6}$$

Thus, the CWT resonance coefficient is the correlation (convolution) between the time series and the appropriate wavelets, as in Figure 7.3. In a (Morlet) Wavelet Transform, a wavelet is correlated with different sections of a financial time series $x(t)$. The inner product of a section and the wavelet is a new function. The volume of the area delimited by that function and computed by the integral is the wavelet resonance (or correlation) coefficient. Sections of the time series $x(t)$ that look like the wavelet give large resonance coefficients, as seen in Figure 7.3(c) and (d). (The (scalar) product of two negative functions is positive.) Slowly changing sections of $x(t)$ produce small resonance coefficients, as seen in (e) and (f). Accordingly, the time series $x(t)$ is analyzed at different scales, using wavelets of different widths.

Thus, the dilating and translating wavelet atoms can be used as the orthonormal basis for a unique, complete observation system, which allows continuously varying levels of resolution, like a microscope. Similar to Gábor's WFT, a WT can measure the time–frequency variation of spectral components, but it has a sharper, more localized time–frequency resolution than the WFT. One of the reasons is that wavelets tend to be irregular, fractal and asymmetric, while sinusoids are smooth, periodic and symmetric (Bruce *et al*., 1996).

The CWT can operate at any scale, from that of the original financial time series up to some maximum scale, which is determined by trading off the need for

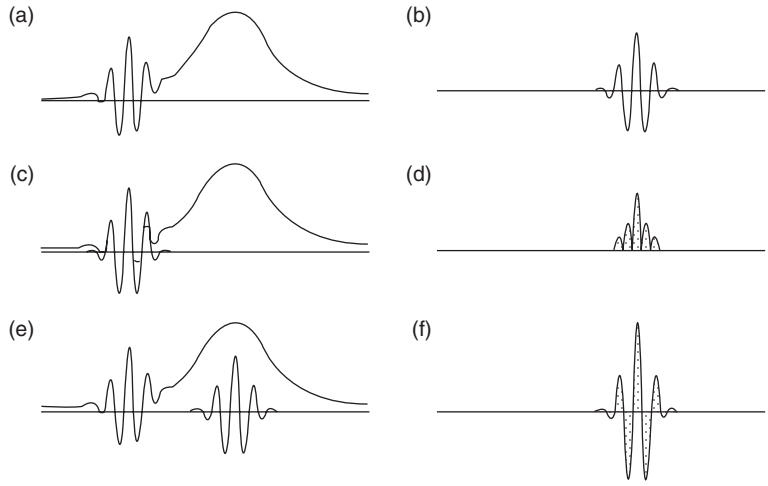


Figure 7.3 Wavelet coefficients are "correlation" or "resonance" coefficients. Here a wavelet is correlated with an irregular signal. Different sections of the signal produce different resonance coefficients.

detailed analysis with available computational power. The CWT is not only continuous in terms of scaling, it is also continuous in terms of shifting (translation): during computation, the analyzing wavelet $\psi_{\tau,a}(t)$ is shifted smoothly over the full domain of the analyzed series $x(t)$. When done, one has computed the wavelet resonance coefficients produced by different sections of the signal, translated by τ, and at different scales a.

The CWT $W(\tau, a)$ has four very useful mathematical properties:

(1) It's linear: $W(\tau, a)\{\gamma_1 x_1(t) + \gamma_2 x_2(t)\} = \gamma_1 W(\tau, a)\{x_1(t)\} + \gamma_2 W(\tau, a)\{x_2(t)\}$;
(2) It's invariant under translation $W(\tau, a) = W(\tau - \tau_0, a)$;
(3) It's invariant under dilation $W(\tau, a) = (1/k)W(k\tau, ka)$, using $k = 1/\sqrt{a}$;
(4) It's localized in time and frequency.

How to make analytic sense of the resulting multitude of wavelet resonance coefficients? How can we interpret them? Usually one makes a *visualization plot* in which the abscissa represents the position $t - \tau$ along the time axis, the ordinate represents the scale a, and the grey scale or color at each (τ, a) point represents the magnitude of the wavelet coefficient $|W(\tau, a)|$ as in Figure 7.4 (This is a grey scale example of a *scalogram*, defined in Section 7.2.3). These plots of wavelet coefficient resemble an irregular surface viewed from above. You can also represent the same coefficients in a 3-dimensional (3D) plot as in Figure 7.5. Notice the slope of the ridges from the small-scale to the large-scale coefficients. The maxima of these ridges are the maxima lines. The speed of their decay from the large to the

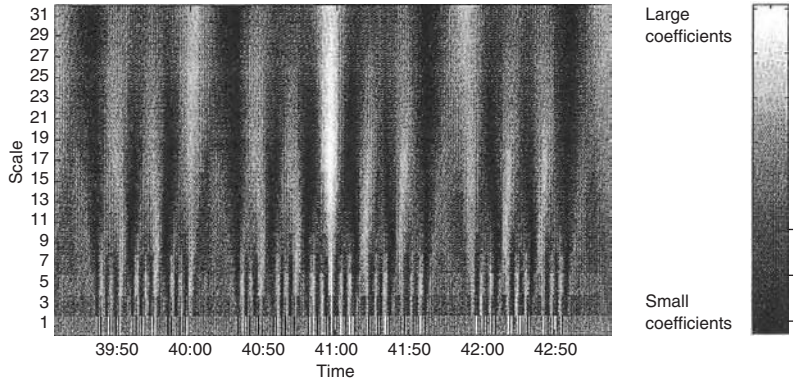


Figure 7.4 A scalogram: a plot of the magnitude of wavelet coefficients.

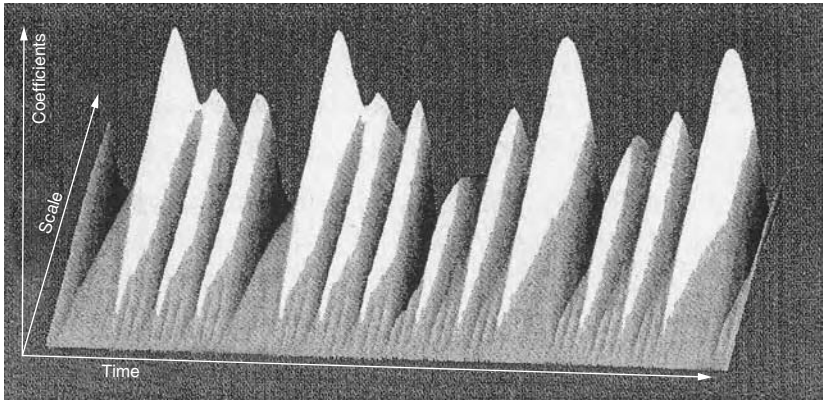


Figure 7.5 A 3D scalogram: a plot of the magnitude of the wavelet coefficients in three dimensions.

small scales can be used for very sophisticated irregularity analysis of singularities and to compute singularity spectra, as we will see in Chapter 8.

By applying Parseval's Formula (cf. Chapter 6), the CWT resonance coefficient, which is a time integral, can also be written as a frequency integral (Walker, 1997):

$$W(\tau, a) = \int_{-\infty}^{\infty} x(t)\psi^*_{\tau,a}(t)dt$$
$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\Psi^*_{\tau,a}(\omega)d\omega \qquad (7.7)$$

where $F(\omega)$ is the FT of $x(t)$ and $\Psi^*_{\tau,a}(\omega)$ is (the complex conjugate of) the FT of the wavelet atom $\psi_{\tau,a}(t)$:

$$\Psi_{\tau,a}(\omega) = \sqrt{a}\Psi(a\omega)e^{j\omega\tau} \tag{7.8}$$

In the time integral, the financial time series (e.g. of investment returns) $x(t)$ is correlated with the wavelet $\psi_{\tau,a}(t)$. Its risk is concentrated in a positive time interval centered at τ. In the equivalent frequency integral, the FT of the time series $F(\omega)$ is correlated with the FT of the wavelet $\Psi_{\tau,a}(\omega)$. This means that $\Psi(\omega) = 0$ for $\omega < 0$. The risk of $\Psi_{\tau,a}(\omega)$ is concentrated over a positive frequency interval centered at τ/a, whose size is scaled by a^{-1}.

7.2.2 Relationship between frequency and scale

In the time–frequency plane, a wavelet atom $\psi_{\tau,a}$ is again symbolically represented by a Heisenberg box centered at $(\tau, \eta/a)$, as in Figure 7.6. The time and frequency spread are proportional to a and a^{-1}, respectively. When the scale a varies, the height and width of the Heisenberg rectangle change, but *its area or volume remains constant*. When a decreases, i.e., when the time resolution decreases, the frequency support of the wavelet is shifted to the higher frequencies, and *vice versa*, in accordance with Heisenberg's Uncertainty Principle, discussed in Chapter 6.

As Figure 7.6 shows, the higher scales a correspond to the most dilated ("stretched") wavelets. The more dilated the wavelet $\psi_{\tau,a}(t)$, the longer the portion of the time series $x(t)$ with which it is being compared, and thus the coarser the

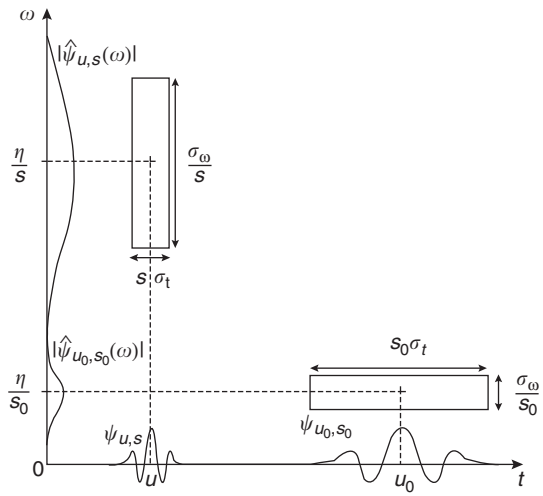


Figure 7.6 Heisenberg boxes of two wavelets. Smaller scales decrease the time dispersion, but increase the frequency support, which is shifted towards higher frequencies.

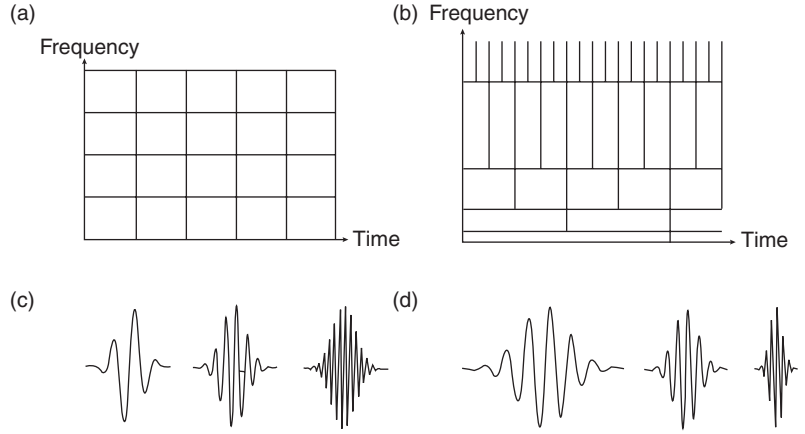


Figure 7.7 Time–frequency resolution and basis functions of the Windowed FT and the Wavelet Transform: (a) tiling of the time–frequency plane for the WFT, (b) for the WT, (c) corresponding basis functions for the WFT and (d) for the WT.

time series features being measured by the wavelet resonance coefficients $W(\tau, a)$. Therefore, there is an inverse correspondence between the scale parameter a and the radian frequency ω:

> Low scale $a \Leftrightarrow$ compressed wavelet $\Leftrightarrow$ rapidly changing time series details $\Leftrightarrow$ high frequency ω.
> High scale $a \Leftrightarrow$ dilated wavelet $\Leftrightarrow$ slowly changing, coarser time series features $\Leftrightarrow$ low frequency ω.

Remark 260 *Gábor's Windowed or Short-Term Fourier Transform (STFT) obtains frequency information with limited precision, and that precision is determined by the size of the window of the particular Gábor atom, which remains the same for all frequencies. In contrast, the WT uses wavelet windows that vary according to their scale (= "inverted frequency"). This can be clearly seen in Figure 7.7, which compares the basis functions and time–frequency resolution of Gábor's WFT and the WT in* tiling diagrams. *The tiles in these diagrams represent the essential concentration in the time–frequency plane of a given basis function (Herley* et al.*, 1993; Strang, 1993). Notice in (d) that the shape of the wavelet basis functions is invariant under the changes in frequency. This is what produces the precise and unambiguous interpretation of a time–frequency analysis by a WT.*

7.2.3 Scalograms: varying scaled and localized densities

Thus, the wavelet coefficient $W(\tau, a)$ depends on the values of $x(t)$ and its FT $F(\omega)$ in the time–frequency region, where the risk of the wavelet atom $\psi_{\tau,a}$ and

its FT $\Psi_{\tau,a}(\omega)$ is concentrated. Measuring time-varying frequencies is again the most important application of WTs. Sharp transitions in the time series $x(t)$ create large amplitude wavelet resonance coefficients $W(\tau, a)$ at scales a localized at time τ. The time evolution of such spectral or scale components is analyzed by following the location of such large amplitude coefficients. As we saw earlier, these visualizations are called scalograms (Figures 7.4 and 7.5). They are already used on an experimental basis in economics (Ariño and Vidakovic, 1995), and in finance (Jensen, 1997).

Definition 261 *A* scalogram*, or* local wavelet spectrum*, is the localized and scaled wavelet density (= modulus squared of the WT)*

$$P_W(\tau, a) = |W(\tau, a)|^2 = \left| \int_{-\infty}^{\infty} x(t) \psi^*_{\tau,a}(t) dt \right|^2 \tag{7.9}$$

Thus, a scalogram measures the *localized risk* of a financial time series $x(t)$ in the time–scale neighborhood of (τ, a) specified by the Heisenberg box of $\psi_{\tau,a}(t)$. If η denotes the frequency center of the base wavelet, then the frequency of a dilated wavelet is $\xi = \eta/a$.[4]

Definition 262 *The* normalized scalogram *is*

$$\frac{\xi}{\eta} P_W(\tau, a) \tag{7.10}$$

Example 263 *Figure 7.8 displays the scalogram of a wavelet analysis of our example of the Gábor Windowed Fourier Analysis in Chapters 5 and 6, using a Morlet wavelet for* 25 *different scale levels:*

$$\psi_{\tau,a}(t) = e^{-t^2/2} e^{j\omega_m t} + e^{-\omega_m^2/2} \tag{7.11}$$

for $\omega_m = 6$*, where the term in* ω_m^2 *ensures admissibility (it is negligible for* $\omega > 5$*) and* $j = \sqrt{-1}$ *(adapted from Bendjoya and Slezak, 1993, p. 240). Scale a and frequency* ω *are related by* $a = \omega_m/\omega$*. Again, the abscissa measures time t and the ordinate the scale a. A grey coding is used with the largest resonance coefficients in black and the smallest in white. Notice the differences compared with the Gábor analysis. First, all three monochromatic frequencies present in the signal* $x(t)$ *with the same amplitude are detected in the same fashion. The three detected coherent frequencies have the same weight. Moreover, the discontinuities are detected by the two cones pointing towards the locations of these singularities at the small scales. The width of these cones contains information about the type of singularity, as will be discussed in Chapter 8.*

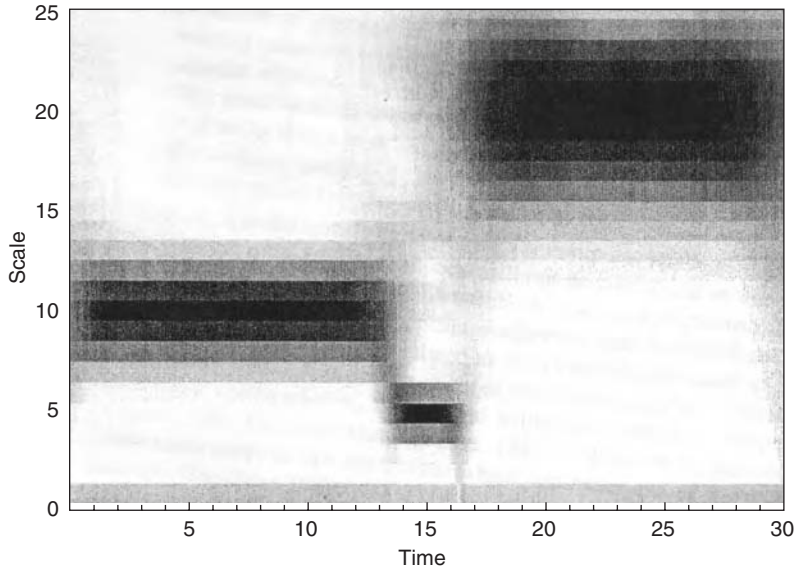


Figure 7.8 A scalogram with modulus $|W(\tau, a)|$ using a Morlet (6) wavelet for 25 different scales along the ordinate and time along the abscissa, with grey level coding. The three detected coherent frequencies $\omega_3 < \omega_1 < \omega_2$ in the signal $x(t)$ have the same weight. Discontinuities are visible at all scales through the cones pointing toward the location of their singularities at the smallest scale. The grey strip at the smallest scale is a finite sampling effect.

Example 264 *Figure 7.9 demonstrates that the normalized scalogram* $(\xi/\eta)P_W(\tau, a)$ *is also a device to visualize nonstationarity, in particular of time-varying frequencies, or "chirps." The time series of the one but last example in Chapter 6 of* $T = 1{,}000$ *observations is analyzed in the following scalogram. As we noted, the time series includes a linear chirp, whose frequency increases linearly over time, a quadratic chirp, whose frequency decreases quadratically over time, and two modulated Gaussian noise functions located at* $t = 512$ *and* $t = 896$*. Compare this scalogram with the spectrogram in Chapter 6. Despite the appearance to the contrary, the scalogram represents the data analysis more truthfully than the spectrogram, since the scalogram visualizes also the relative* epistemic uncertainty *of the computed frequencies, in particular, of the higher frequencies, as required by the Heisenberg Principle and the Heisenberg boxes.*

The reason for the varying relative epistemic (knowledge) uncertainty in scalograms is perfectly clear from the dyadic time–scale tiling diagram in Figure 7.10, which shows how a smooth sinusoidal function and an isolated singularity are represented in a scalogram. In contrast, the spectrograms in Chapter 6 represent

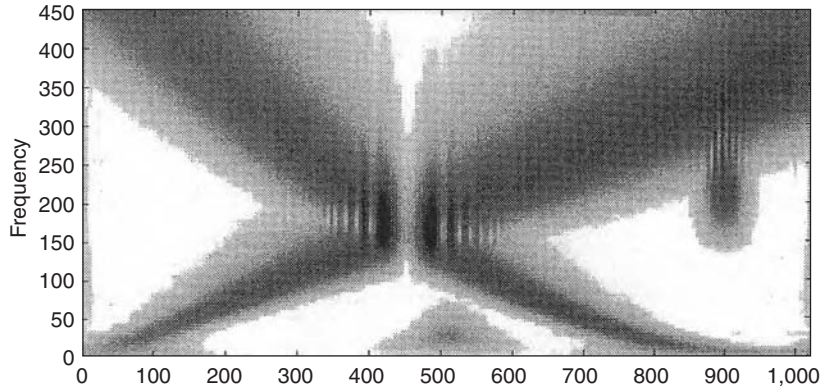


Figure 7.9 Normalized scalogram $(\xi/\eta)\,P_W(\tau, a)$ computed from the "chirped" time series of Figure 6.8. Dark points indicate large amplitude wavelet coefficients.

illusory analytic precision, where there can't be any, because it is tiled by scale-invariant, equally sized Heisenberg boxes.

Example 265 *Figure 7.11 provides an empirical example of a 3D normalized scalogram measuring a sharp discontinuity or break in a financial time series. This scalogram is computed from* 44,640 *minute-by-minute increments of the Thai baht quotations of July 1997. The large discontinuity at the beginning of the time axis (from front towards the right) is caused by the financial crisis of July 2, 1997 and is represented by very large wavelet resonance coefficients (measured along the vertical axis from* 0 *to* 100 *percent) over the various scales a along the scale axis (from the front, where* $a = 1$ *minute, towards the left, where* $a = 60$ *minutes* $= 1$ *hour). Notice that most of the risk of the discontinuity is concentrated on the smallest scale = highest frequency of Foreign Exchange (FX) trading: the discontinuity generated a short-lived vortex.*

Related to the scalogram is the scalegram or wavelet spectrum, which is the wavelet analog of the average risk spectrum.

Definition 266 *A* scalegram*, or* global (average) wavelet spectrum *is the scaled wavelet density (= average modulus squared of the WT):*

$$\begin{aligned} P_W(a) &= \int_{-\infty}^{\infty} |W(\tau, a)|^2 d\tau \\ &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} x(t)\psi^*_{\tau,a}(t)dt \right|^2 d\tau \end{aligned} \tag{7.12}$$

The wavelet spectrum is the scalogram projected (integrated) onto the scale or (inverted) frequency axis. It provides the wavelet equivalent of the classical

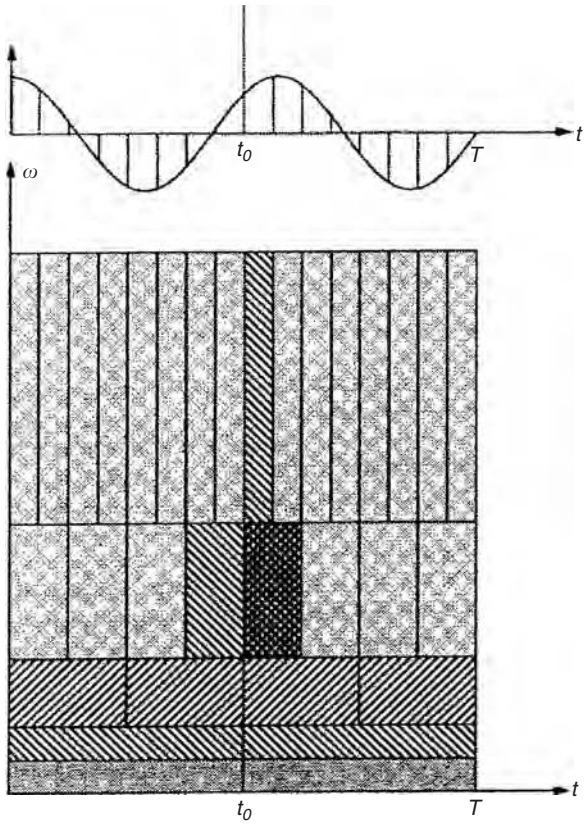


Figure 7.10 Time–scale tiling for a sinusoidal function with an isolated singularity represented by a "cone" in the scaleogram at t_0. The abscissa represents time. The ordinate represents either increasing frequency or decreasing scale $(1/a)$.

marginal frequency distribution (cf. Chapter 2) and Fourier risk spectrum (cf. Chapters 5 and 6).[5] Like the average risk spectrum, it is used to look at the components of a signal as a function of scale (or frequency), with disregard to location. The wavelet spectrum based on the WT contains much the same information as the risk spectrum based on the FT. The wavelet spectrum can be used to identify the homogeneous H-exponent(s) from scaling financial time series, as we will see in the next chapter. This, in turn, is used to identify special models of nonlinear deterministic behavior called *transient chaos* or *intermittency* (Scargle, 1997).

Remark 267 *When the noise model is based on counting photons, as in Chapter 1, this noise simply adds a constant (independent of scale) to the true average wavelet spectrum. That constant is the mean counting rate of the photon*

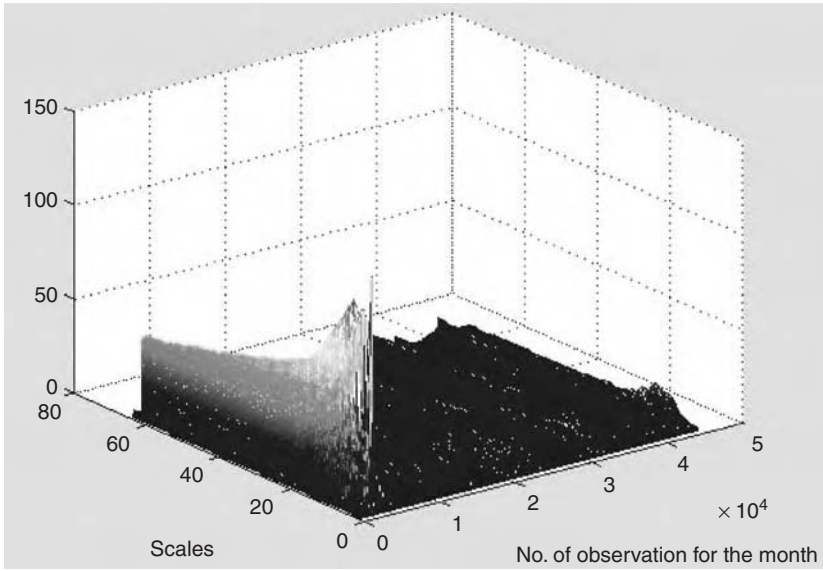


Figure 7.11 Empirical 3D scalogram of Thai baht increments in July 1997.

count. In practice, it is quite easy to identify the true scalegram, although the identification remains uncertain when the signal-to-noise ratio is low.

7.2.4 Frame theory and wavelet bases

Subsampling of Gábor's WFT, or of the CWT, defines a *complete* representation of the financial time series, if any such time series can be reconstructed from linear combinations of discrete families of windowed Fourier atoms $\{g_{\tau_n,\xi_k}\}_{(n,k)\in\mathbb{Z}^2}$, or, respectively, of wavelet atoms $\{\psi_{\tau_n,a_j(n,j)\in\mathbb{Z}^2}\}$. *Frame theory* discusses what conditions the families of wavelets must satisfy if they are to provide *stable and complete representations* of time series. Completely eliminating redundancy is equivalent to building an orthogonal basis of the time series space. The following discussion is meant to generalize for WTs what we already have learned from the FTs.

Definition 268 *A* frame *is a family of vectors $\phi_k(t)$, which can represent any financial time series with finite risk by the sequence of its inner products with the vectors of the family. More precisely, a family $\{\phi_k(t)\}_{k\in\mathbb{Z}}$ of vectors in the real square-integrable, or Hilbert space $L^2(\mathbb{R})$, is a frame of this L^2 space, if there are two constants $A > 0$ and $B > 0$ such that, for any $x(t)$ in the space $L^2(\mathbb{R})$,*

$$A\|x(t)\|^2 \leq \sum_{k\in\mathbb{Z}} |\langle x(t), \phi_k(t)\rangle|^2 \leq B\|x(t)\|^2 \tag{7.13}$$

where

$$\|x(t)\|^2 = \int_{-\infty}^{+\infty} |x(t)|^2 dt \tag{7.14}$$

is the risk content of $x(t)$*, aka the* norm *of financial time series* $x(t)$.

Definition 269 *If* $A = B$*, the frame is said to be* tight.

In general a frame is a stable, redundant and not necessarily linear representation of a financial time series. It is a generalization of the more familiar fundamental concept of the *basis* of a linear space, which we've encountered earlier. For example, an orthogonal basis is a complete, tight frame. The frame vectors $\{\phi_k(t)\}_{k\in\mathbb{Z}}$ are supposed to be of *unit norm*, that is

$$\|\phi_k(t)\|^2 = \int_{-\infty}^{+\infty} |\phi_k(t)|^2 dt = 1 \tag{7.15}$$

When is a frame not tight but redundant (Lawton, 1990)?

Definition 270 *A frame is* redundant*, if* $1 < A$.

There exist some specific bases for the wavelet decomposition space. For example,

Definition 271 *A* Riesz basis *is a frame of linearly independent vectors* $\{\phi_k(t)\}_{k\in\mathbb{Z}}$.

If the frame vectors $\{\phi_k(t)\}_{k\in\mathbb{Z}}$ form a Riesz basis, then $A \leq 1 \leq B$.

Definition 272 *An* orthonormal basis *is a tight (= complete = unique) Riesz basis.*

Thus, a frame is an orthonormal basis if and only if $A = B = 1$.

A financial time series $x(t)$ can always be expanded into a series of terms, as follows.

Definition 273 *An* expansion *is the linear decomposition*

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k \phi_k(t) \tag{7.16}$$

where k *is an integer index for the finite or infinite sum,* a_k *are the real-valued expansion coefficients, and* $\phi_k(t)$ *are a set of real-valued functions of* t*, called the* expansion set.

We can always expand, decompose or analyze a financial time series into a series of terms, but the question is: when is such an analytic decomposition complete and thus unique? That depends on the set of terms $\{\phi_k(t)\}$. If this set is complete and thus unique, the expansion is.

Definition 274 *If an expansion (= linear decomposition) is unique, the set of frame vectors* $\{\phi_k(t)\}$ *is called a* basis *for the class of financial time series* $x(t)$ *that can be so decomposed.*

Remark 275 *For example, the set of exponential eigenvectors* $\{e^{j\omega t}\}$ *form the expansion set for the FT. Since it is a unique expansion, this particular expansion set forms a basis.*

One of the crucial consequences of dealing with an orthogonal basis is that the expansion coefficients a_k can always be computed by the inner product (or correlation):

$$\begin{aligned} a_k &= \langle x(t), \phi_k(t) \rangle \\ &= \int_{-\infty}^{+\infty} x(t)\phi_k(t)dt \end{aligned} \tag{7.17}$$

As we discussed in Chapter 5, this was, indeed, the case with the computation of the FT resonance coefficients, and, we will see, it is also the case with a properly defined wavelet basis! We need also to define the extend or span of the basis set and what is maximally included in such a span.

Definition 276 *The* span *of a basis set, span* $\{\phi_k(t)\}$*, is the set of all financial time series* $x(t)$ *that can be decomposed in terms of this set of bases:*

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k \phi_k(t) \tag{7.18}$$

Definition 277 *The* closure *of the space spanned by the basis set,* $\overline{\text{span}\{\phi_k(t)\}}$*, contains not only all variables that can be expressed by a linear combination of the basis functions* $\phi_k(t)$*, but also the variables which are the limit of these infinite expansions.*

The closure is usually denoted by an over-bar, as we will see in Definition 291, when we discuss wavelet MRA, which is a particular form of unique expansion. To do so, we need the definition of a wavelet expansion.

Definition 278 *A* wavelet expansion *is the two-parameter (or 2D) expansion, such that*

$$x(t) = \sum_{j=0}^{+\infty} \sum_{n=-\infty}^{+\infty} a_{j,n} \psi_{j,n}(t) \tag{7.19}$$

where the integer indices $j, n \in \mathbb{Z}^2$ *and the* $\psi_{j,n}(t)$ *are the wavelet expansion functions that (usually) form an orthogonal basis.*

Similar to the preceding general frame definitions, we have for the wavelet bases the following specific definitions.

Definition 279 *An* orthogonal wavelet basis *is a complete set of orthogonal wavelets:*

$$\{\psi_{j,n}\}_{(j,n)\in\mathbb{Z}^2} \tag{7.20}$$

Definition 280 *An* orthonormal wavelet basis (*or* tight wavelet frame) *is a complete set of orthonormal wavelets.*

The term complete in these articular definitions means that there are no redundant wavelets in these set and that the set is unique.

Definition 281 *The set of expansion coefficients* $a_{j,n}$ *are called the* Discrete Wavelet Transform (DWT) *of* $x(t)$ *and the wavelet expansion is the* inverse DWT.

The following is a most remarkable and powerful theorem. It is clearly the foundation for the success and current popularity of wavelet MRA. For its proof we refer to the aforementioned mathematical wavelet literature, in particular to Mallat (1989a).

Theorem 282 (Wavelet expansion) *Any financial time series* $x(t)$ *with finite risk can be decomposed over a orthogonal wavelet basis*

$$x(t) = \sum_{j=0}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle x(t), \psi_{j,n} \rangle \psi_{j,n} \tag{7.21}$$

This theorem is remarkable and powerful, because it states that *any* financial time series can be so *completely* analyzed. There is no approximation involved!

In summary:

(1) A wavelet basis is a set of *building blocks* to represent a function or time series $x(t)$. It is a 2D expansion set (usually a basis) for some class of 1- (or higher) dimensional functions.
(2) The wavelet expansion gives a *time–scale (frequency) localization* of $x(t)$. Most of the risk of the financial time series $x(t)$ is well represented by a few expansion coefficients $a_{j,n}$.
(3) The computation of these expansion coefficients from $x(t)$ can be done *efficiently* in discrete time.

7.2.5 DWT Systems

We will now present some specific examples of first-generation wavelet sets of DWTs, which are all generated from a single scaling function, or a wavelet, by simple time translation and frequency scaling.

Almost all useful wavelet sets also satisfy the so-called *multiresolution conditions*. This means that if a set of data series can be represented by a weighted sum of laterally shifted wavelets $\psi(t-k)$, then a larger set (including the original) can be represented by the weighted sum of $\psi(2t-k)$. The lower resolution coefficients can be computed from the higher resolution coefficients by a tree-structure algorithm, called a *filter bank*. Mallat (1989a) provided the mathematical basis for such an MRA, as we will discuss in detail in Section 7.3.

We will now first define some families of wavelets, which can be quite diverse.

7.2.5.1 Haar wavelet

The first recorded mention of the term "wavelet" was in 1909, in the PhD thesis of Alfréd Haar (1910).[6] Haar realized that one can construct a very simple piecewise constant function whose dilations and translations generate a complete *dyadic* orthonormal basis in Hilbert space, i.e., in real quadratic linear space $L^2(\mathbb{R})$, as follows.[7]

Definition 283 *A* dyadic *orthonormal wavelet basis in Hilbert space* $L^2(\mathbb{R})$ *is defined by the set of dyadically scaling orthonormal wavelets:*

$$\begin{aligned}\psi_{j,n}(t) &= \frac{1}{\sqrt{2^j}}\psi\left(\frac{t-2^j n}{2^j}\right)\\ &= 2^{-j/2}\psi(2^{-j}t-n)\end{aligned} \tag{7.22}$$

The discrete *dyadic scale* parameter $a_j = 2^j$, while the translation interval is $\tau_n = 2^j n$. The factor $2^{-j/2}$ maintains a constant norm independent of scale j. Thus, we have the definition of the Haar wavelet.

Definition 284 *The discrete* Haar wavelet *is defined by:*

$$\psi^H(t) = \begin{cases} +1 & \text{if } 0 \le t < 0.5 \\ -1 & \text{if } 0.5 \le t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{7.23}$$

The Haar wavelet is the (basic) wavelet that appears most useful for the analysis of financial time series, in particular of the increments or rates of return of pricing series, since they are produced by independently shifting demand and supply (curves), often in the form of discretely recorded tick data from trading transactions, and contain many singularities.

7.2.5.2 Other families of wavelets

Of course, there exists a whole family of considerably more sophisticated wavelets other than the simple Haar wavelet. For example, other rather simple wavelets are the discrete time triangle wavelet and the continuous time Gábor wavelet.

Definition 285 *The discrete* Triangle Wavelet *is*

$$\psi^T(t) = \begin{cases} +t & \text{if } 0 \le t < 0.25 \\ 0.5 - t & \text{if } 0.25 \le t < 0.75 \\ -1.0 + t & \text{if } 0.75 \le t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{7.24}$$

Definition 286 *The continuous* Gábor wavelet *is a particular Gábor chirp (*cf. *Chapter 6)*

$$\psi^G(t) = e^{j\eta t} g(t) \tag{7.25}$$

with a Gaussian window

$$g(t) = \frac{1}{\sigma^2 \pi^{0.25}} e^{-t^2/2\sigma^2} \tag{7.26}$$

for $\sigma^2\eta^2 \gg 1$.

The *Gábor wavelet* wavelet family, which has a Gaussian flavor, is often used in theoretical *continuous time* MRA, where it provides elegant solutions for difficult problems, as we will see in Chapter 8.

The family of wavelets is growing rapidly, since customized sets of wavelets can be carefully created to satisfy selected situations by applying particular zero-moment conditions, as Daubechies (1988) first demonstrated. There now exists already a remarkable *Daubechies(N) wavelet* family, where N = order of zero-moments of the Daubechies wavelets. But there are no explicit closed form Daubechies wavelets, except the Daubechies(1) wavelet, which is the same as the Haar wavelet. The $N > 1$ order Daubechies wavelets are all defined numerically by sets of recursive equations which define the filter coefficients of these wavelets. This is a similar situation as the no-closed form of most stable frequency distributions in Chapter 3.

7.3 Mallat's MRA

An efficient way to implement the DWT in the form of an MRA was invented in 1986 and developed in 1988 by Mallat (1989a–c). The operational Mallat algorithm is in fact a classical scheme known to signal processing engineers as a *two-channel subband coder*, or tree analysis. This very practical filtering algorithm yields a Fast Wavelet Transform, similar to the Fast Fourier Transform (FFT).

For many financial time series, the low-frequency content of a time series $x(t)$ is the most important part, since it gives the series its recognizable identity. Its high-frequency content, on the other hand, imparts its flavor or nuance. Mallat showed that one can completely decompose a time series $x(t)$ in terms of *approximations* (A), provided by so-called scaling functions, and *details* (D), provided by the wavelets. The approximations are the high-scale, low-frequency components of the time series. The details are the low-scale, high-frequency components.

This decomposition process can be iterated, with successive approximations being decomposed in turn, so that one time series $x(t)$ is broken down in many lower-resolution components. This is called the *wavelet decomposition tree*. Since the decomposition process is iterative, in theory it can be continued indefinitely. In reality, the decomposition can proceed only until the individual details consist of a single observation. For example, when one observes minute-by-minute data, the one-minute data point provides the smallest detail of resolution. The choice of the *wavelet filters* determines the shape of the wavelet we use to perform the analysis.

The following discussion of Mallat's MRA is adapted from Burrus *et al*. (1998) and from Hubbard (1998).

7.3.1 Low- and high-pass filters

Mathematically, for the MRA of the financial time series $x(t)$, one needs two closely related basic functions. In addition to the wavelet $\psi(t)$, which provides the details, one needs a second basis function, called the *scaling function*, which provides the low frequency approximation, e.g., like an average or mean. This scaling function and the wavelets are conjugated, as we will see. One cannot exist without the other. Mallat (1989a) proves that, using a combination of these scaling functions and wavelets, a very large class of time series can be represented by the following decomposition equation of scaling functions and wavelets:

$$\begin{aligned} x(t) &= A + D \\ &= \sum_{n=-\infty}^{+\infty} c_n \varphi_n(t) + \sum_{j=0}^{+\infty} \sum_{n=-\infty}^{+\infty} d_{j,n} \psi_{j,n}(t) \end{aligned} \tag{7.27}$$

where the approximation (A) is provided by the 1D linear combination of the scaling functions, which form the so-called *low-pass filters*:

$$\begin{aligned} A &= \sum_{n=-\infty}^{+\infty} c_n \varphi_n(t) \\ &= \sum_{n=-\infty}^{+\infty} c_n \varphi(t-n) \end{aligned} \tag{7.28}$$

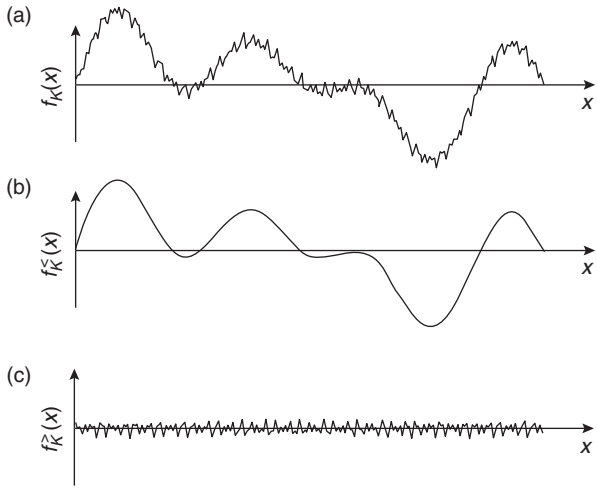


Figure 7.12 Time signal observations on $f(x)$ in panel (a) are subject to low-pass filtering in panel (b) and subject to high-pass filtering in panel (c).

and the details (D) by the 2D linear combination of the dyadic wavelets, which form the so-called *high-pass filters*:

$$D = \sum_{j=0}^{+\infty} \sum_{n=-\infty}^{+\infty} d_{j,n} \psi_{j,n}(t)$$
$$= \sum_{j=0}^{+\infty} \sum_{n=-\infty}^{+\infty} d_{j,n} \psi(2^{-j} t - n) \tag{7.29}$$

Example 287 *Figure 7.12 shows a primitive form of MRA or decomposition: 1D time series $f_K(x)$ (a) subject to low-pass filtering, (b) as indicated by the $<$ sign, and high-pass filtering, (c) as indicated by the $>$ sign. This kind of analysis is discussed in finance by, for example, Fama and French (1988). Obviously,*

$$f_K(x) = f_K^<(x) + f_K^>(x) \tag{7.30}$$

where $1/K$ is the scale of this filtering. This is a time series which possesses structures on only two very different scales: a small scale (of the order of a few millimeters) and a large scale (of the order of a few centimeters). The filter scale $1/K$ is chosen to be intermediate, say, $c.1$ cm. The passage from $f_K^<(x)$ to $f_K^>(x)$ may be generalized to filters of arbitrary shape, such as the WTs.

The coefficients of this MRA expansion, or DWT, are again computed as inner products, or basis correlations, as follows.

Definition 288 *The discrete (approximation)* scaling coefficients *are computed by the inner product*

$$c_n = \langle x(t)\phi_n(t)\rangle = \int_{-\infty}^{+\infty} x(t)\phi_n(t)dt, \quad \text{with } n \in \mathbb{Z} \tag{7.31}$$

Definition 289 *The discrete (detail)* wavelet resonance coefficients *are computed by the inner product*

$$d_{j,n} = \langle x(t)\psi_{j,n}(t)\rangle = \int_{-\infty}^{+\infty} x(t)\psi_{j,n}(t)dt, \quad \text{with } j, n \in \mathbb{Z} \tag{7.32}$$

Let's look now at some more clarifying definitions and see why this decomposition into approximating scaling functions $\phi_n(t)$ and detailing wavelets $\psi_{j,n}(t)$ leads to a *complete* MRA of the time series $x(t)$, due to the conjugation between these two functions.

Definition 290 *A set of (time)* scaling functions *is defined in terms of integer translates of the real square integrable basic scaling function by*

$$\varphi(t) = \varphi(t-n), \quad n \in \mathbb{Z}, \ \ \varphi \in L^2(\mathbb{R}) \tag{7.33}$$

Thus, a scaling function is a strictly *periodic* function, as defined in Chapter 3. It exactly repeats itself with a lag n. Therefore, most of the concepts of the periodic wave theory of FTs can be applied to scaling functions.

Definition 291 *The square-integrable real (Hilbert)* subspace of $L^2(\mathbb{R})$ spanned *by these scaling functions is defined as*

$$\mathbf{V}_0 = \overline{\text{span}\{\varphi_n(t)\}} \tag{7.34}$$

for all integers k from $-\infty$ *to* $+\infty$. *The over-bar denotes* closure. *This means that*

$$x(t) = \sum_{n=-\infty}^{+\infty} a_n\varphi_n(t) \quad \text{for any } x(t) \in \mathbf{V}_0 \tag{7.35}$$

One can increase the size of the subspace spanned by changing the time scale of the time scaling functions. A 2D set of functions is generated from the basic

scaling function by translation and scaling, as follows

$$\varphi_{j,n}(t) = 2^{-j/2}\varphi(2^{-j}t - n) \tag{7.36}$$

whose span over n is

$$\begin{aligned}\mathbf{V}_j &= \overline{\text{span}_n\{\varphi_n(2^{-j}t)\}} \\ &= \overline{\text{span}_n\{\varphi_{j,n}(t)\}} \quad \text{for all integers } n \in \mathbb{Z}\end{aligned} \tag{7.37}$$

This means that any time series $x(t)$ can be linearly expanded strictly in terms of scaling functions as follows:

$$\begin{aligned}x(t) &= \sum_{n=-\infty}^{+\infty} a_n \varphi_{j,n}(t) \\ &= \sum_{n=-\infty}^{+\infty} a_n \varphi_n(2^{-j}t - n) \quad \text{for any } x(t) \in \mathbf{V}_j\end{aligned} \tag{7.38}$$

For $j < 0$, the span can be larger, since the $\varphi_{j,n}(t)$ is narrower and is translated in smaller steps. Therefore, it can represent finer detail. For $j > 0$, $\varphi_{j,n}(t)$ is wider and is translated in wider steps. The wider scaling functions can represent only coarse information, and the space they span is smaller. Thus, the change in scale provides a change in resolution. The scale j indicates the *resolving power* of the analysis, similar to the resolving power of lenses in optics and photography.

7.3.2 MRA equation

Mallat formulated these intuitive ideas of scale and resolution into mathematical requirements for a *complete* MRA, by requiring a nesting of the spanned spaces $\mathbf{V}_j$ as follows

$$\mathbf{V}_{j+1} \subset \mathbf{V}_j \quad \text{for all } j \in \mathbb{Z} \tag{7.39}$$

with

$$\mathbf{V}_\infty = \{0\} \quad \text{and} \quad \mathbf{V}_{-\infty} = L^2(\mathbb{R}) \tag{7.40}$$

Thus, the linear space that contains low resolution will also contain the linear spaces of high resolution. This means that at the zero resolution, the only finite risk time series is 0, while at the infinite resolution all finite risk time series are perfectly reproduced. In other words, because of the definition of the spanned spaces $\mathbf{V}_j$, the spaces must satisfy the natural dyadic scaling condition

$$\varphi(t) \in \mathbf{V}_j \quad \Leftrightarrow \quad \varphi(2t) \in \mathbf{V}_{j+1} \tag{7.41}$$

which ensures that elements in a space are simply scaled versions of the elements in the next space. Thus, $\mathbf{V}_{j+1}$ is obtained from $\mathbf{V}_j$ by factor 2 rescaling.

This *dyadic nesting* of the spans of $\varphi(2^{-j}t - n)$, denoted by $\mathbf{V}_j$, is achieved by requiring that $\varphi(t) \in \mathbf{V}_1$, which means that if $\varphi(t) \in \mathbf{V}_0$, it is also $\varphi(t) \in \mathbf{V}_1$, the space spanned by $\varphi(2t)$. The resolution of V_j is generated by a basis which is obtained by 2^{-j} translations of a $2^{=j}$ rescaled φ. The φ is such a function that integer translations or lateral time shifts of φ creates a Riesz basis of $\mathbf{V}_\infty$. As we discussed earlier, a Riesz basis is a frame of linearly independent vectors.

Thus, we have now arrived at Mallat's formal definition of an MRA.

Definition 292 (Mallat's MRA) *A sequence* $\{\mathbf{V}_j\}_{j\in\mathbb{Z}}$ *of closed subspaces of* $L^2(\mathbb{R})$ *is an* MRA*, if and only if the following six properties are satisfied:*

(1) For all $(j, n) \in \mathbb{Z}^2$, $x(t) \in \mathbf{V}_j \Leftrightarrow x(2^{-j}t - n) \in \mathbf{V}_{j+1}$ *(*$\mathbf{V}_j$ *is* 2^{-j} dyadic translation invariant*)*

(2) For all $j \in \mathbb{Z}$, $\mathbf{V}_{j+1} \subset \mathbf{V}_j$ (nesting of resolutions)

(3) For all $j \in \mathbb{Z}$, $x(t) \in \mathbf{V}_j \Leftrightarrow x(2^{-1}t) \in \mathbf{V}_{j+1}$ (dyadic scaling of resolutions)

(4) $\lim_{j\to\infty} \mathbf{V}_j = \cap_{j=-\infty}^{+\infty} \mathbf{V}_j = \{0\}$ (at zero resolution, finite risk is 0)

(5) $\lim_{j\to-\infty} \mathbf{V}_j = Closure(\cup_{j=-\infty}^{+\infty} \mathbf{V}_j) = L^2(\mathbb{R})$ (at infinite resolution, perfect reproduction of finite risk)

(6) There exist a function φ*, such that* $\{\varphi(t - n)\}_{n\in\mathbb{Z}}$ *is a* Riesz basis *of* $\mathbf{V}_0$.

Remark 293 *When the Riesz basis is an orthogonal basis, the MRA is orthogonal, and its base atom is called a* scaling function*. It is always possible to orthogonalize any MRA. This implies that scaling functions always exist. However, orthogonalities impose constraints, such as that a compactly supported orthogonal scaling function cannot be symmetric and continuous, as Daubechies (1988, 1992) proved.*

The definition of an MRA implies that the scaling function $\varphi(t)$ can be expressed in terms of an expansion, i.e., a weighted sum of shifted $\varphi(2t)$ as follows (Strang, 1989).

Definition 294 *The* MRA (dilation *or* scaling) equation *is*

$$\varphi(t) = \sum_{n=-\infty}^{+\infty} h(n)\sqrt{2}\varphi(2t - n), \quad \textit{for any } n \in \mathbb{Z} \tag{7.42}$$

where the coefficients $h(n)$ *are real or complex numbers called the* scaling (function) coefficients *(= the* scaling filter *or* scaling vector*) and the* scaling factor $1/\sqrt{2}$ *maintains the norm of the scaling function.*

An equivalent way to present the MRA equation is

$$\frac{1}{\sqrt{2}}\varphi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{+\infty} h(n)\varphi(t - n), \quad \text{for any } n \in \mathbb{Z} \tag{7.43}$$

Its FT is

$$\Phi(\omega) = \frac{1}{\sqrt{2}}\mathbf{H}\left(\frac{\omega}{2}\right)\Phi\left(\frac{\omega}{2}\right) \tag{7.44}$$

where $\mathbf{H}(\omega)$ is the transfer function, i.e., the FT of $h(n)$.

This recursive equation is fundamental to MRA and is analogous to a differential equation with coefficients $h(n)$ and a closed form solution $\varphi(t)$, that may or may not exist, or be unique or nonunique. For example, only for the Daubechies(1) wavelet (= Haar wavelet) exists an explicit expression, which is the Haar scaling function. For all higher order Daubechies wavelets only *numerical* solutions exist in the form of computed $h(n)$ coefficients, although the square modulus of the transfer function, $|\mathbf{H}(\omega)|^2$, is explicit and often fairly simple.

The coefficients $h(n)$ form a so-called *conjugate mirror filter*, which entirely determines the scaling function and most of its properties. In particular, the scaling function is compactly supported, if and only if $h(n)$ has a finite number of zero coefficients. In fact, we have the following crucial MRA design Theorem of Mallat and Meyer, which we present again without its proof, since it can be found in Mallat (1989a–c).

Theorem 295 (Mallat and Meyer MRA design) *Let $\varphi \in L^2(\mathbb{R})$ be any integrable scaling function. The Fourier series of the MRA coefficients $h(n)$, which are computed by the inner product*

$$h(n) = \left\langle \frac{1}{\sqrt{2}}\varphi\left(\frac{t}{2}\right), \varphi(t-n)\right\rangle \tag{7.45}$$

satisfies the two dyadic equations

$$\text{for all } \omega \in \mathbb{R}, \quad |\mathbf{H}(\omega)|^2 + |\mathbf{H}(\omega+\pi)|^2 = 2 \tag{7.46}$$

and

$$\mathbf{H}(0) = \sqrt{2} \tag{7.47}$$

Conversely, if $\mathbf{H}(\omega)$ is 2π periodic and continuously differentiable in a neighborhood of $\omega = 0$, if it satisfies the two dyadic equations, and if

$$\min_{\omega\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right]} |\mathbf{H}(\omega)| > 0 \tag{7.48}$$

then

$$\Phi(\omega) = \prod_{p=1}^{+\infty} \frac{\mathbf{H}(2^{-p}\omega)}{\sqrt{2}} \tag{7.49}$$

is the FT of a scaling function $\varphi \in L^2(\mathbb{R})$.

Remark 296 *This important FT of the scaling function is currently used for designs of new MRAs, which show up in the image compression filters, like the JPEG and MPEG filters of digital cameras and in the storage of digital movies on the Internet, and which are used for the digital restoration and coloration of old classic black and white movies in Hollywood (Mulcahy, 1996, 1997).*

7.3.2.1 Examples of MRA scaling filters

The following are examples of the scaling function and the corresponding MRA equation for some of the wavelet families we've introduced earlier in this chapter:

(1) The *Haar (= Daubechies(1)) scaling function* is

$$\varphi(t) = \begin{cases} 1 & \text{if } 0 \leq t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{7.50}$$

and its corresponding *Haar MRA equation* is

$$\varphi(t) = \varphi(2t) + \varphi(2t-1) \tag{7.51}$$

i.e., the MRA equation with two scaling coefficients

$$\begin{aligned} h(0) &= h(1) \\ &= \frac{1}{\sqrt{2}} = 0.70711 \end{aligned} \tag{7.52}$$

(rounded to five digits).

Figure 7.13 (left) shows how the Haar MRA equation corresponds with the graph of its scaling function.

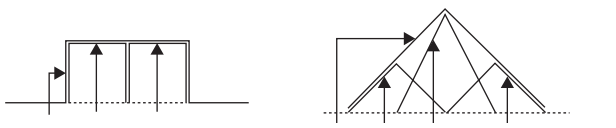


Figure 7.13 Haar (left) and triangle (right) scaling functions and their respective MRA equations.

(2) The *triangle scaling function* is

$$\varphi(t) = \begin{cases} t & \text{if } 0 \leq t < 0.5 \\ 1 - t & \text{if } 0.5 \leq t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{7.53}$$

and the corresponding *triangle MRA equation* is

$$\varphi(t) = \frac{1}{2}\varphi(2t) + \varphi(2t - 1) + \frac{1}{2}\varphi(2t - 2) \tag{7.54}$$

i.e., the MRA equation is with three scaling coefficients

$$h(0) = \frac{1}{2\sqrt{2}} = 0.35355 \tag{7.55}$$

and

$$h(1) = h(2) = \frac{1}{\sqrt{2}} = 0.70711 \tag{7.56}$$

Figure 7.13 (right) shows how the triangle MRA equation corresponds with the graph of its scaling function.

(3) The *Daubechies(4) scaling function* is the MRA equation with four scaling function coefficients

$$h(0) = 0.48296 \tag{7.57}$$

$$h(1) = 0.83652 \tag{7.58}$$

$$h(2) = 0.22414 \tag{7.59}$$

and

$$h(3) = -0.12941. \tag{7.60}$$

These scaling function coefficients are crucial for the "regularity" or "zero-moment" properties of the scaling filter. They summarize the density of information in the time series $x(t)$.

7.3.3 Relationship between wavelets and scaling functions

But what is the general relationship between wavelets and scaling functions? The wavelets are used to build a basis on which to represent the details of a time series that are gained between a particular resolution represented by a scaling function, and the next finer resolution, as is seen from the following set of definitions.

Definition 297 Sets of wavelets $\{\psi_{j,n}(t)\}$ *are sets of functions that span the differences between the spaces spanned by the various scales of a scaling function.*

Remark 298 *Usually it is required that the scaling functions and their corresponding wavelets are orthogonal, because orthogonal functions allow simple computation of expansion coefficients by inner (correlation) products and, consequently, as we will see at the end of this chapter, have a Parseval's Theorem that allows the* complete partitioning *of the financial time series' risk in the WT's time–scale domain.*

Definition 299 *The* orthogonal complement *(or* disjoint difference*) of* $\mathbf{V}_j$ *in* $\mathbf{V}_{j+1}$ *is* $\mathbf{W}_j$*. This means that all elements of* $\mathbf{V}_j$ *are orthogonal to all elements of* $\mathbf{W}_j$*, or the inner products of all scaling functions and wavelets equal zero*

$$\langle \varphi_{j,n}(t)\psi_{j,l}(t)\rangle = \int_{-\infty}^{\infty} \varphi_{j,n}(t)\psi_{j,l}(t)dt = 0 \tag{7.61}$$

for all appropriate $j, l, n \in \mathbb{Z}$*.*

This relationship between the orthogonal spaces is indicated as follows

$$\mathbf{V}_{j+1} = \mathbf{V}_j \oplus \mathbf{W}_j \tag{7.62}$$

where the symbol $\oplus$ indicates that the space $\mathbf{V}_{j+1}$ consist of the subspace $\mathbf{V}_j$ and its orthogonal complement $\mathbf{W}_0$. But then

$$\begin{aligned} \mathbf{V}_2 &= \mathbf{V}_1 \oplus \mathbf{W}_1 \\ &= \mathbf{V}_0 \oplus \mathbf{W}_0 \oplus \mathbf{W}_1 \end{aligned} \tag{7.63}$$

and, in general, the whole real square-integrable (Hilbert) space is completely divided up as follows

$$L^2(\mathbb{R}) = \mathbf{V}_0 \oplus \mathbf{W}_0 \oplus \mathbf{W}_1 \oplus \mathbf{W}_2 \oplus \cdots \tag{7.64}$$

where $\mathbf{V}_0$ is the initial space spanned by the scaling function $\varphi(t-n)$.

The scale of the initial space $\mathbf{V}_0$ is arbitrary and can be chosen at any available resolution. Thus also

$$\mathbf{W}_{-\infty} \oplus \cdots \oplus \mathbf{W}_{-1} = \mathbf{V}_0 \tag{7.65}$$

which again shows the arbitrariness of the scale of the scaling space. In practice, the scale of the scaling space is chosen to represent the coarsest detail, or largest observation window, of interest in the time series $x(t)$. This will become clearer when we exhibit some of the empirical MRA examples.

The MRA equation for scaling functions is complemented by the MRA equation for wavelets.

Definition 300 *The* MRA equation for wavelets *is the weighted sum of shifted scaling functions*

$$\psi(t) = \sum_{n=-\infty}^{+\infty} h_1(n)\sqrt{2}\phi(2t-n), \quad n \in \mathbb{Z} \tag{7.66}$$

for some set of wavelet (generation) coefficients $h_1(n)$*, since the wavelets reside in the space spanned by the next narrower scaling function,* $\mathbf{W}_0 \subset \mathbf{V}_1$.

This MRA equation for wavelets can equivalently be presented as

$$\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{+\infty} h_1(n)\phi(t-n), \quad n \in \mathbb{Z} \tag{7.67}$$

Its FT is

$$\Psi(\omega) = \frac{1}{\sqrt{2}}\mathbf{H}_1\left(\frac{\omega}{2}\right)\Phi\left(\frac{\omega}{2}\right) \tag{7.68}$$

It can be easily proved, that, because of the MRA requirements and because of the orthogonality of the translates of the scaling function, these wavelet generation coefficients $h_1(n)$ (modulo translates by integer multiples of two) are required by orthogonality to be related to the scaling function coefficients by the following equation

$$h_1(n) = (-1)^n h(1-n) \tag{7.69}$$

The MRA equation for the wavelet $\psi(t)$ gives the prototype or *mother wavelet* for a class of expansion functions of the form

$$\psi_{j,n}(t) = 2^{j/2}\psi(2^j t - n) \tag{7.70}$$

7.3.3.1 Examples of MRA wavelet filters

The following are again examples of the wavelet resonance coefficients $h_1(n)$ that satisfy the wavelet coefficient equation, and show how easy it is to generate wavelets from sets of particular scaling functions:

(1) For the *Haar wavelet*, the MRA equation for wavelets is

$$\psi(t) = \phi(2t) - \phi(2t-1) \tag{7.71}$$

and the two wavelet generation coefficients are

$$\begin{aligned} h_1(0) &= -h_1(1) \\ &= \frac{1}{\sqrt{2}} = 0.70711 \end{aligned} \tag{7.72}$$

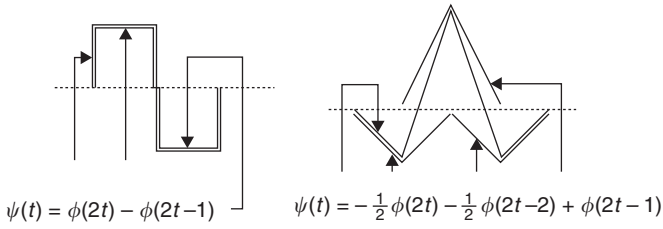


Figure 7.14 Haar (left) and triangle (right) wavelets and their respective MRA equations.

Figure 7.14 (left) shows how the Haar MRA equation for wavelets corresponds with the graph of its wavelet.

(2) For the *triangle wavelet*, the MRA equation for wavelets is

$$\psi(t) = -\frac{1}{2}\phi(2t) + \phi(2t-1) - \frac{1}{2}\phi(2t-2) \tag{7.73}$$

and the three wavelet generation coefficients are

$$h_1(0) = -\frac{1}{\sqrt{2}} = -0.70711 \tag{7.74}$$

$$h_1(1) = -0.35355 \tag{7.75}$$

and

$$h_1(2) = 0.35355 \tag{7.76}$$

Figure 7.14 (right) shows how the triangle MRA equation for wavelets corresponds with the graph of its wavelet.

(3) For the *Daubechies (4) wavelet* (*cf.* Figure 7.2), the four wavelet generation coefficients are

$$h_1(0) = 0.12941 \tag{7.77}$$

$$h_1(1) = 0.22414 \tag{7.78}$$

$$h_1(2) = -0.83651 \tag{7.79}$$

and

$$h_1(3) = 0.48296 \tag{7.80}$$

The wavelet resonance coefficients summarize the detailed transient information in time series $x(t)$.

7.3.4 Design properties of MRA systems

Nowadays the optimal design of MRA systems consists of the choosing of the scaling coefficients $h(n)$ according to particular signal processing design criteria to improve the resolution of a single time series $x(t)$, or of a 2D time series, *c.q.,* digital pictures and movies, like the JPEG and MPEG criteria (Donoho, 1993a,b; Strichartz, 1993; Daubechies, 1996).[8] Let's have a look at some of the specific MRA design properties, such as the order of the vanishing moments, support, regularity and symmetry of the wavelet and scaling functions.[9]

7.3.4.1 Vanishing moments

Criterion 301 *A wavelet $\psi(t)$ has p* vanishing moments

$$\int_{-\infty}^{+\infty} t^k \psi(t) dt = 0 \quad \textit{for } 0 \leq k < p \tag{7.81}$$

if and only if its scaling function can generate polynomials of degree smaller than or equal to p.[10]

When the wavelet's p moments are equal to zero, all the polynomial time series

$$x(t) = \sum_{0 \leq k < p} a_k t^k \tag{7.82}$$

have zero wavelet resonance coefficients and their details are also zero. This property ensures the suppression of time series that are polynomials, since the systematic polynomials are *exactly* captured by the scaling functions, just like cosine and sine waves are exactly captured by Fourier series.

Remark 302 *While this property of vanishing moments is described by the approximating power of scaling functions, for wavelets it has also the possibility to characterize the order of isolated singularities. The order of vanishing moments of a wavelet is entirely determined by the coefficients $h(n)$ of the filter h featured in the scaling function.*

If the FT of the wavelet is p times continuously differentiable, then the three following conditions are exactly equivalent:

(1) The wavelet ψ has p vanishing moments.
(2) The scaling function φ can generate polynomials if degree smaller than or equal to p.
(3) The FT of the filter h and its first $p-1$ first derivatives vanish at $\omega = \pi$.

All these conditions can rather easily be checked.

Remark 303 *Daubechies (1988) proved that, to generate an orthogonal wavelet with p vanishing moments, a filter of minimum length $2p$ has to be used. Daubechies filters, which generate Daubechies wavelets, have a length of $2p$.*

7.3.4.2 Compact support

The following property of compact support has to do with the speed of convergence to zero at infinity of the scaling and wavelet functions, when either time or frequency goes to infinity.

Criterion 304 *If the support of the scaling function is* $[N_1, N_2]$*, then the wavelet support is* $[(N_1 - N_2 - 1)/2, (N_2 - N_1 + 1)/2]$*. The scaling function is* compactly supported *if and only if the filter h has finite support, and the support boundaries of h are not the same* $N_1 \neq N_2$*, i.e., the support is asymmetric.*

Remark 305 *Daubechies (1988, 1992) also showed that it is possible for the scaling function and the wavelets to have both compact support (i.e., to be nonzero only over a finite region) and to be orthonormal. This made possible the desired exhaustive and complete time–scale analysis. Thus, all risk of a financial time series can be completely presented by a* dyadic tiling *of its whole time–scale domain.*

7.3.4.3 Irregularity

We have discussed irregularity of time series somewhat in Chapter 4, and we will discuss it in more detail in Chapter 8. Wavelet regularity is less important than their vanishing moments. Wavelets can be very irregular. Consequently, the following two wavelet properties are crucial:

Criterion 306 *There is* no *compactly supported orthogonal wavelet which is* indefinitely differentiable.

Criterion 307 *For Daubechies wavelets with a large number of vanishing moments p, the scaling function and wavelet are* α-Lipschitz*, where* α *is of the order of* $0.2p$*. For large classes of orthogonal wavelets, more regularity implies more vanishing moments.*[11]

To represent $x(t)$ with K derivatives, one can choose a wavelet $\psi(t)$ that is K (or more) times continuously differentiable. The penalty for imposing greater *smoothness* in this sense is that the supports of the basis functions, the filter lengths and, hence, the computational complexity all increase. The most remarkable property is that smooth bases are also the "best" bases for representing time series with arbitrarily many singularities. This is a property that may become essential for the ongoing research of singularity spectra, as discussed in Chapter 8 (Donoho, 1993a).

7.3.4.4 Symmetry

Symmetric scaling functions and wavelets are important, because they are used to build bases of regular wavelets over any interval, rather than over only the real axis. Daubechies (1988, 1993) proved that, for a wavelet to be symmetric

or anti-symmetric, its filter must have a linear complex phase and the following property.

Criterion 308 *There is no* symmetric *compactly supported orthogonal wavelet other than the Haar wavelet, which corresponds to a discontinuous wavelet with one vanishing moment. Thus, all compactly supported orthogonal wavelets other than the Haar wavelet are asymmetric.*

7.3.5 Usefulness of wavelets

Burrus *et al*. (1998, p. 216) summarize why wavelets are so useful for (financial) time series analysis:

(1) Wavelets, even very irregular ones, can represent *smooth* time series, in particular, those series which exhibit some form of scaling behavior, since wavelets themselves are self-similar.
(2) Wavelets can also represent series of *singularities* of various kinds.
(3) Wavelets are *local*. This makes most coefficient-based algorithms naturally adaptive to the heterogeneities in the time series.
(4) Wavelets have the *unconditional basis* property for a great variety of time series, implying that if one knows very little about a time series, as is often the case in the financial markets, the wavelet basis is usually a reasonable choice for measurement and analysis.

7.4 Wavelet Parseval Risk Decomposition Theorem

If the scaling functions and wavelets form an orthonormal basis (= a tight frame), there is a Parseval's (tiling) Theorem that relates the risk of the financial time series $x(t)$ exactly to the risk in each of the components and their wavelet resonance coefficients (Herley *et al*., 1993).

Theorem 309 (Parseval's tiling) *For the general wavelet expansion the risk or variance of the financial time series can be expanded (or analyzed) as*

$$\int_{-\infty}^{+\infty} |x(t)|^2 dt = \sum_{n=-\infty}^{+\infty} |c_n|^2 + \sum_{j=0}^{+\infty} \sum_{n=-\infty}^{+\infty} |d_{j,n}|^2 \tag{7.83}$$

with the risk in the expansion completely partitioned in time by n and in scale by j.

Parseval's Theorem for wavelets allows us to completely and uniquely decompose the overall risk of a financial market time series into subcomponent, i.e., into financial risk at any scale (frequency) and at any moment in time. Thus, we have achieved one of the major objectives of this book: to find a rigorous method to decompose the risk of a financial rate of return series into any time-localized scale or frequency of our choice.

Using Parseval's tiling one can check if volatility scales in foreign exchange markets, as Batten and Ellis (1999) claim, or if the volatility of credit spreads scales, as Batten *et al*. (1999) claim. In Chapter 8, we'll discuss how to use the wavelet MRA to identify the homogeneous Hurst exponent of the parametric scaling of the FBM.

7.5 Software

The computations of the following Exercises can be executed by using the MATLAB® Wavelet Toolbox, available from The MathWorks, Inc., 24 Prime Park Way Natick, MA 01760-1500, USA. Tel.: (508) 647-7000; Fax: (508) 647-7001; http://www.mathworks.com/products/wavelettbx.shtml.

A complete wavelet analysis toolkit called *Wavelab 802* can be obtained at no cost from Stanford University: http://playfair.stanford.EDU:80/~wavelab/.

This *Wavelab* is a very complete set of MATLAB® scripts that implement both the basic wavelet and related transforms and more advanced techniques. There is full documentation, a set of tutorials, and section of "Toons," short for *cartoons*. These Toon scripts reproduce from scratch the figures in many papers of Stanford's wavelet research group consisting of Dave Donoho, Ian Johnstone *et al*., describing the theoretical research underlying the algorithms in *Wavelab*. By studying these scripts and by experimenting with the data, the reader can learn all the details of the process that led to each figure. This forms part of the new discipline of *Reproducible Research*, i.e., the idea to provide the reader full access to all details (data, equations, code, etc.) needed to completely reproduce all the results normally presented only in summary form in scientific publications. For example, there are MATLAB® scripts to generate and exactly reproduce all the figures in the book by Mallat (1998). The pioneer of such *Reproducible Research* is Jon Claerbout of Stanford University's Geophysics Department (Claerbout, 1994; Buckheit and Donoho, 1995).

In addition, one can use Benoit 1.3: Fractal System Analysis (for Windows 95/98 or Windows NT), Trusoft International Inc., 204 37th Ave. N #133, St Petersburg, FL 33704. Tel.: (813) 925-8131; Fax: (813) 925-8141; sales@trusoft-international.com. See http://www.trusoft-international.com for details. This Benoit software enables you to measure the fractal dimension and/or Hölder-Hurst exponent of your data sets using your choice of method(s) for analysis of self-affine traces of speculative prices. However, astonishingly, the wavelet routine in Benoit 1.3 is incorrect, although the other routines to compute the Hurst exponent are correct.

7.6 Exercises

Exercise 310 *Study and execute MATLAB® Exercises – Chapter 1, Wavelet Display, in Strang and Nguyen (1997), p. 454. This assignment is designed to familiarize the reader with the MATLAB® Wavelet Toolbox and its GUI (= Graphical User Interface). The assignment includes Wavelet Display, Continuous Wavelet Transform, 1D and 2D DWT. It is assumed that the Wavelet Toolbox is installed. At the MATLAB® prompt, type* wavemenu. *A window should pop up with choices*

ranging from Wavelet 1D *to* Wavelet Packet 2D *to* Continuous Wavelet 1D*. Closely follow Strang and Nguyen's instructions.*

Exercise 311 *Study and execute MATLAB® Exercises – Chapter 1, CWT in Strang and Nguyen (1997), pp. 454–455. Under File menu option, select* Load Signal*. Choose the MATLAB® file MATLAB® on in* /toolbox/wavelet/wavedemo/freqbrk.mat.

Exercise 312 *and follow Strang and Nguyen's instructions. Notice the scalogram. What does the scalogram show us (in detail)? Use the File menu option to load another signal file of MATLAB® on 'Nbs1'* (F:)/toolbox/wavelet/wavedemo/qdchirp.mat*, and again follow Strang and Nguyen's instructions.*

Exercise 313 *Study and execute MATLAB® Exercises – Chapter 1, 1D DWT, Strang and Nguyen (1997), p. 455. Under File menu option, select* Load Signal*. Choose the MATLAB® file in* /toolbox/wavelet/wavedemo/noisdopp.mat.

Exercise 314 *and follow Strang and Nguyen's instructions. Study the various Decomposition and Statistics, Histogram, Compress and Denoise capabilities (Note: the MATLAB® GUI follows Daubechies' convention of giving the scaling function and the lowest frequency wavelet the highest scaling number. Strang and Nguyen gives them the lowest scaling number, i.e., the zero). Study in particular the automatic denoising and compression. Notice how few wavelet resonance coefficients are required for an acceptable synthesis and reconstruction of the original data series (= "signal").*

Exercise 315 *Study and execute MATLAB® Exercises – Chapter 6, Multiresolution Analysis in 1D, Strang and Nguyen (1997), p. 466.*

Exercise 316 *Study and execute MATLAB® Exercises – Chapter 6, Wavelet Packet in 1D, Strang and Nguyen (1997), p. 466.*

Notes

1 In Chapter 8, we'll provide an empirical analysis of Latin American financial markets using Morlet (6) wavelets.

2 An *electroencephalogram* (EEG) is a recording of the electrical activity of the brain, and an *electroencephalograph* is the instrument used for making the recording. The technique, called electroencephalography, was first reported in 1929 by Hans Berger, a German psychiatrist. The complexity of the brain and the inability of the electrical recording apparatus to distinguish the direction of nerve impulses within the brain, because it identifies correlations and not causalities, make it very difficult to interpret the EEG. The frequency of these impulses also varies in different parts of the brain. But certain distinctive, abnormal patterns are clearly associated with such situations as epilepsy, stroke and brain tumors. Thus, the study of a patient's EEG can aid in medical diagnosis.

3 Indeed, Brian Yuhnke Jr, one of the new media software programmers with whom I work at Kent State University, interprets my current research as trying to find a real-time "doppler radar for financial markets." No doubt, his comment is inspired by the doppler radar used for the weather report of Channel 3, the t.v. station in Cleveland, Ohio, close to our university, and also used for tornado warnings.
4 Thus, a scalogram shows the localized conventional "R-squareds" between the time series and each of the wavelets in the tiling diagram.
5 The scalegram corresponds with the *average* "R-squareds."
6 Alfréd Haar (1885–1933) was born in Budapest, Hungary. In 1904, Haar travelled to Germany to study at Göttingen under Hilbert's supervision, obtaining his doctorate in 1909 with a dissertation entitled *Zur Theorie der Orthogonalen Funktionensysteme* ("The Theory of Systems of Orthogonal Functions"). Haar then taught at Göttingen until 1912, when he returned to Hungary and held chairs at the university in Kolozsvár (which is now Cluj in Romania), Budapest University and Szeged University. Haar, together with Riesz, rapidly made the new Szeged University a major mathematical center. Later he went on to study partial differential equations. In 1932, he introduced a measure on groups, now called the Haar measure, which allows an analogue of Lebesgue integrals to be defined on locally compact topological groups. Thus, he generalized classical measure theory! The Haar measure was used by both von Neumann and Pontryagin in 1934 and by Weil in 1940 to set up an abstract theory of commutative harmonic analysis. Only now, after Mallat created its tree-algorithm in 1988 and combined it with the current advanced state of computing power, find these powerful abstract ideas feasible applications in advanced time series analysis and in the analysis of financial market risk.
7 *Dyadic* = based on a geometric sequence of ratio 2.
8 There exist now already better standards for the design of the multiresolution of signals than MPEG, e.g., the design criteria for the archives of the digitized FBI fingerprints and the design criteria for recent compact digital cameras, of which the filter allows hundreds of pictures to be compressed and stored in a relatively small physical memory.
9 This section is for specialists and can be skipped in a first reading. It was Daubechies who mathematically developed and researched these four important design criteria for WTs.
10 For example, the Gaussian distribution kernel, which can be presented as a wavelet, has vanishing moments at $p = 3$ and for all $p > 4$.
11 Since the Gaussian has vanishing moments at $p = 3$ and for all $p > 4$, the Gábor wavelet, which has a Gaussian atom, is very regular. We'll discuss the concept of α-Lipschitz irregularity in Chapter 8.

Bibliography

Aldroubi, Akram, and Michael Unser (Eds) (1996) *Wavelets in Medicine and Biology*, CRC Press, Boca Raton, FL.

Antoniadis, A., and G. Oppenheimer (Eds) (1995) *Wavelets and Statistics*, Springer Verlag, New York, NY.

Ariño, M. A., and B. Vidakovic (1995) "On Wavelet Scalograms and Their Applications in Economic Time Series," Discussion Paper 95(21), ISDS, Duke University, Durham, NC.

Aussem, Alex, Jonathan Campbell and Fionn Murtagh (1998) "Wavelet-Based Feature Extraction and Decomposition Strategies for Financial Forecasting," *Journal of Computational Intelligence in Finance*, **6**(2), 5–12.

Batten, Jonathan, and Craig Ellis (1999) "Volatility Scaling in Foreign Exchange Markets," CREFS Working Papers #99-04, Nanyang Technological University, Singapore, April, 23 pages.

Batten, Jonathan, Craig Ellis and Warren Hogan (1999) "Scaling the Volatility of Credit Spreads: Evidence from Australian Dollar Eurobonds," CREFS Working Papers #99-03, Nanyang Technological University, Singapore, April, 23 pages.

Bendjoya, Ph., and E. Slézak (1993) "Wavelet Analysis and Applications to Some Dynamical Systems," *Celestial Mechanics and Dynamical Astronomy*, **56**, 231–262.

Benedetto, J. J., and M. W. Frazier (Eds) (1994) *Wavelets: Mathematics and Applications*, CRC Press, Boca Raton, FL.

Bruce, A., David Donoho and H. Y. Gao (1996) "Wavelet Analysis," *IEEE Spectrum*, **33**(10), 26–35.

Buckheit, J. B., and D. L. Donoho (1995) "Wavelab and Reproducible Research," in Antoniadis, A., and G. Oppenheim (Eds) *Wavelets and Statistics*, Springer Verlag, Berlin, pp. 53–81.

Burke-Hubbard, Barbara (1994) "The Mathematical Microscope: Waves, Wavelets, and Beyond," in M. Bartusiak *et al*. (Ed.) *A Positron Named Priscilla, Scientific Discovery at the Frontier*, National Academy Press, Washington, DC, pp. 196–235.

Burke-Hubbard, Barbara (1998) *The World According to Wavelets: The Story of a Mathematical Technique in the Making*, 2nd edn, A K Peters, Wellesley, MA.

Burrus, Sidney, Ramesh A. Gopinath and Haitao Guo (1998) *Introduction to Wavelets and Wavelet Transforms: A Primer*, Prentice Hall, Upper Saddle River, NJ.

Chui, Charles K. (1992a) *An Introduction to Wavelets*, Academic Press, San Diego, CA.

Chui, Charles K. (Ed.) (1992b) *Wavelets: A Tutorial in Theory and Applications*, Academic Press, New York, NY.

Claerbout, J. (1994) "Reproducible Electronic Documents," http://sepwww.stanford.edu/research/redoc.

Coifman, R. R., and M. V. Wickerhauser (1992) "Entropy-based Algorithms for Best Basis Selection," *IEEE Transactions of Information Theory*, **38**(2), 713–718.

Cohen, Albert, and Jelena Kovacevic (1996) "Wavelets: The Mathematical Background," *Proceedings of the IEEE*, **84**(4), 514–522.

Daubechies, Ingrid (1988) "Orthonormal Bases of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics*, **41**(7), 909–996.

Daubechies, Ingrid (1992) *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA (Notes from the 1990 CBMS-NSF Conference on Wavelets and Applications, at Lowell, MA).

Daubechies, Ingrid (1993) Orthonormal Bases of Compactly Supported Wavelets II. Variations on a Theme, *SIAM Journal on Mathematical Analysis*, **24**(2), 499–519.

Daubechies, Ingrid (1996) "Where Do Wavelets Come From? – A Personal Point of View," *Proceedings of the IEEE*, **84**(4), 510–513.

Donoho, David L. (1993a) "Unconditional Bases are Optimal Bases for Data Compression and for Statistical Estimation," *Applied and Computational Harmonic Analysis*, **1**(1), 100–115.

Donoho, David L. (1993b) "Nonlinear Wavelet Methods for Recovery of Signals, Densities, and Spectra from Indirect and Noisy Data," *Proceedings of Symposia in Applied Mathematics*, **47**, 173–205.

Fama, Eugene F., and Kenneth R. French (1988) "Permanent and Transitory Components of Stock Prices," *Journal of Political Economy*, **96**(2), 24–73.

Flandrin, Patrick (1989) "On the Spectrum of Fractional Brownian Motion," *IEEE Transactions on Information Theory*, **35**(1), January, 197–199.

Flandrin, Patrick (1992) "Wavelet Analysis and Synthesis of Fractional Brownian Motion," *IEEE Transactions on Information Theory*, **38**(2), March, 910–917.

Frazier, Michael W. (1999) *An Introduction to Wavelets Through Linear Algebra*, Springer Verlag, New York, NY.

Gençay, Ramazan, Faruk Selçuk and Brandon Whitcher (2001) *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, Academic Press, San Diego, CA.

Grossman, A., and J. Morlet (1984) "Decomposition of Hardy Functions Into Square Integrable Wavelets of Constant Shape," *SIAM Journal of Mathematical Analysis*, **15**(4), 723–736.

Haar, Alfréd (1910) "Zur Theorie der Orthogonalen Funktionensysteme ('A Theory of Systems of Orthogonal Functions')," *Mathematical Analysis*, **69**(1), 331–371.

Herley, C., J. Kovacevic, K. Ramachandran and M. Vetterli (1993) "Tilings of the Time–Frequency Plane," *IEEE Transactions in Signal Processing*, **41**(12), 3341–3359.

Holschneider, M. (1995) *Wavelet: An Analysis Tool*, Clarendon Press, Oxford, UK.

Jawerth, Björn, and Wim Sweldens (1994) "An Overview of Wavelet Based Multiresolution Analysis," *SIAM Reviews*, **36**(3), 377–412.

Jensen, Mark J. (1997) "Making Wavelets in Finance," *Financial Engineering News*, **1**(1), 1 and 9–10.

Kaplan, Lance M., and C.-C. Jay Kuo (1993) "Fractal Estimation From Noisy Data, Via Discrete Fractional Gaussian Noise (DFGN) and the Haar Basis," *IEEE Transactions in Signal Processing*, **41**(12), 3554–3562.

Lau, K.-M., and H.-Y. Weng (1995) "Climate Signal Detection Using Wavelet Transform: How to Make a Time Series Sing," *Bulletin of the American Meteorological Society*, **76**, 2391–2402.

Lawton, W. (1990) "Tight Frames of Compactly Supported Wavelets," *Journal of Mathematical Physics*, **31**, 1898–1901.

Mallat, Stéphane G. (1989a) "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**(7), 674–693.

Mallat, Stéphane G. (1989b) "Multiresolution Approximation and Wavelet Orthonormal Bases of L^2," *Transactions of the American Mathematical Society*, **315**(1), 1989, 69–87.

Mallat, Stéphane G. (1989c) "Multifrequency Channel Decomposition of Images and Wavelet Models," *IEEE Transactions on Acoustics, Speech and Signal Processing*, **37**(12), 2091–2110.

Mallat, Stéphane (1999) *A Wavelet Tour of Signal Processing*, 2nd edn, Academic Press, Boston, MA.

Massopust, P. (1994) *Fractal Fluctuations, Fractal Surfaces and Wavelets*, Academic Press, San Diego, CA.

Meyer, Yves (1985) "Principe d'Incertitude, Bases Hilbertienne et Algèbres d'Opérateurs" ("Uncertainty Principle, Hilbert Bases and Operator Algebra"), in *Séminaire Bourbaki*, **145–146**, 209–223, Astérisque, Paris.

Meyer, Yves (1990) *Wavelets and Operators*, Cambridge University Press, Cambridge, UK.

Meyer, Yves (1993) *Wavelets: Algorithms and Applications* (translated and revised by Robert D. Ryan), Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Meyers, S. D., B. G. Kelly and J. J. O'Brien (1993) "An Introduction to Wavelet Analysis in Oceanography and Meteorology: With Application to the Dispersion of Yanai Waves," *Monthly Weather Review*, **121**(10), 2858–2866.

Morlet, J., G. Arens, I. Fourgeau and D. Giard (1982) "Wave Propagation and Sampling Theory," *Geophysics*, **47**(2), 203–236.

Mulcahy, Colm (1996) "Plotting and Scheming with Wavelets," *Mathematics Magazine*, **69**(5), 323–343.

Mulcahy, Colm (1997) "Image Compression Using the Haar Wavelet Transform," *Spelman Science and Mathematics Journal*, **1**(1), 22–31.

Ogden, R. T. (1997) *Essential Wavelets for Statistical Applications and Data Analysis*, Birkhäuser, Boston, MA.

Ogden, R. T., and E. Parzen (1996) "Change-Point Approach to Data Analytic Wavelet Thresholding," *Statistics and Computing*, **6**, 93–99.

Ramsey, James B., and Zhifeng Zhang (1996) "The Application of Waveform Dictionaries to Stock Market Index Data," in Kravtsov, Yurii A., and James B. Kadtke (Eds) *Predictability of Complex Dynamical Systems*, Springer Verlag, New York, NY, 1996, pp. 189–205.

Ramsey, James B., and Zhifeng Zhang (1997) "The Analysis of Foreign Exchange Data Using Waveform Dictionaries," *Journal of Empirical Finance*, **4**, 341–372.

Ramsey, James B., and Camille Lampart (1998a) "Decomposition of Economic Relationships by Time Scale Using Wavelets: Money and Income," *Macroeconomic Dynamics*, **2**(1), 49–71.

Ramsey, James B., and Camille Lampart (1998b) "Decomposition of Economic Relationships by Time Scale Using Wavelets: Expenditure and Income," *Studies in Nonlinear Dynamics and Econometrics*, **3**(1), 23–42.

Ramsey, James B., Daniel Usikov and George M. Zaslavsky (1995) "An Analysis of U.S. Stock Price Behavior Using Wavelets," *Fractals*, **3**(2), 377–389.

Rioul, O., and M. Vetterli (1991) "Wavelets and Signal Processing," *IEEE Signal Processing Magazine*, October, 14–38.

Ruskai, M. B., Gregory Beylkin, Ronald Coifman, Ingrid Daubechies, Stéphane Mallat, Yves Meyer and Louis Raphael (Eds) (1992) *Wavelets and Their Applications*, Jones and Bartlett Publishers, Boston, MA.

Scargle, Jeffrey D. (1997) "Wavelets, Scaling, and Chaos," in Buchler, J. Robert, and Henry Kandrup (Eds) *Nonlinear Signal and Image Analysis*, Annals of the New York Academy of Sciences, **808**, January 30, 125–138.

Strang, Gilbert (1989) "Wavelets and Dilation Equations: A Brief Introduction," *SIAM Reviews*, **31**(4), 614–627.

Strang, Gilbert (1993) "Wavelet Transforms vs. Fourier Transforms," *Bulletin of the American Mathematical Society*, **28**(2), 288–305.

Strang, Gilbert (1994) "Wavelets," *American Scientist*, **82**, 250–255.

Strang, Gilbert, and Truong Nguyen (1997) *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA.

Strichartz, R. (1993) "How to Make Wavelets," *American Mathematical Monthly*, **100**(6), 539–556.

Vetterli, Martin (1986) "Filter Banks Allowing Perfect Reconstruction," *Signal Processing*, **10**(3), 219–244.

Walker, James S. (1997) "Fourier Analysis and Wavelet Analysis," *Notices of the American Mathematical Society*, **44**(6), 658–670.

Wang, Y. (1995) "Jump and Sharp Cusp Detection by Wavelets," *Biometrika*, **82**, 385–397.

Wickerhäuser, M. V. (1994) *Adapted Wavelet Analysis: From Theory to Software*, A K Peters, Boston, MA.

Wornell, G. W. (1995) *Signal Processing with Fractals: A Wavelet-Based Approach*, Prentice-Hall, New York, NY.

8 Multiresolution analysis of local risk

"*Natura Saltus Facit*"

(= "Nature Jumps")[1]

8.1 Introduction

For the first time in history, huge quantities of high-frequency financial data are currently being recorded and stored. Both financial price and volume data have been recorded (Gopikrishnan *et al.*, 1998):

- on a daily basis since the nineteenth century;
- with a sampling rate of one minute or less since 1984; and
- on a transaction-by-transaction (tick-by-tick) basis since 1993.

The first to collect and archive high-frequency, intraday foreign exchange (FX) data from Reuters composite FXFX page were the researchers of the institute of Olsen and Associates in Zürich, Switzerland (Müller *et al.*, 1990; Dacorogna *et al.*, 2001). This was quickly followed by the massive data archiving project under the directorship of Dr Würtz at the Eidgenössische Technische Hochschule (ETH: Federal Technical University) in Zürich, who collected high-frequency data from Reuters data selection feeds, mainly from RIC data records. His research group collected the series of quoted prices of 355 major financial instruments, including FX spot rates, forward rates, deposit rates, currency and deposit fixings, treasury market yields and FX cross rates at a rate of 60 megabytes per month. Financial futures, options and financial news subsequently followed.

Most of the series collected by both the Olsen and Associates and the ETH groups contain unequally spaced prices in the time domain. Such unequal spacing, or *time warping* of prices, produces a new research challenge in finance. Wavelet multiresolution analysis (MRA) can very effectively deal with such time warping. This ability of the wavelet MRA is one of the many reasons why this book advocates its use as a major research tool in finance.

The currencies involved in the instruments of the ETH project include those of the G10 countries, Switzerland, the European Community, Hong Kong and Australia. In Asia's financial *annus horibilis* 1997, when the Asian Financial Crisis erupted, I serendipitously collected, archived and analyzed a complete year of seven minute-by-minute Asian FX data series from Reuters FXFX pages with the assistance of three undergraduate students at the Nanyang Technological University in Singapore. Our Asian high-frequency FX series had the advantage of being equally spaced in time, like conventional time series. But we've already noticed the advantage of wavelet MRA in dealing with equally and unequally spaced, smooth and irregular data, in particular, with both discontinuous and turbulent pricing series.[2]

High-frequency records of financial prices, or of rates of return, in competitive markets exhibit three striking characteristics:

(1) They are conspicuously *discontinuous*, i.e., they are singular at almost every point, because the financial supply and demand curves move in unequal discrete steps, in instantaneous response to discrete news events. For example, Figure 8.1 shows on a time scale of 20 minutes the US Dollar/Deutsche Mark (USD/DEM) exchange rates as mid-prices and as associated logarithmic differences or rates of return (Schnidrig and Würtz, 1995, p. 2, figures 1 and 2). In the right panel the quiet sections represent the two days of the weekends, while there are daily fluctuations in the volatility of the log returns for the five-day trading week.
(2) They are *strictly non-stationary*. However, they adhere to stable scaling or power laws and they are stationary at particular scales. For example, Figure 8.2 shows the scaling law behavior for the USD/DEM exchange rate in a double-logarithmic plot. This scaling law is independent from the source of data (in the period 1993–1994) and holds over several orders of magnitude. The scaling exponent $H = 0.58$ is significantly different from the Gaussian process $H = 0.5$ (Schnidrig and Würtz, 1995, p. 4, figure 3).
(3) They show *aperiodic cyclicity*, i.e., they show intermittent periods of condensation, succeeded by periods of rarefaction. Figures 8.3 and 8.4 demonstrate the impact of the intraday cycles of average trading activity on the intensity of the price changes in the global FX market (Dacorogna *et al.*, 1993). Although the FX market is active 24 hours per day, the social organization of business, combined with the circadian cycle, forces the market activity to experience temporal constraints in each financial region of the world. This impacts the price formation. Similar day and weekend effects can be observed in the stock market returns (French, 1980). Figure 8.3 (left and right panel) shows the average hourly trading transaction density in the global FX market as measured by the number of transactions per hour (Schnidrig and Würtz, 1995, p. 5, figure 6). Figure 8.4 (left and right panel) shows the mean absolute hourly log-price change $E\{|\Delta \ln P_t|\}$ for the USD/DEM rate as a measure for weekly averaged price risk or volatility per hour (Schnidrig and Würtz, 1995,

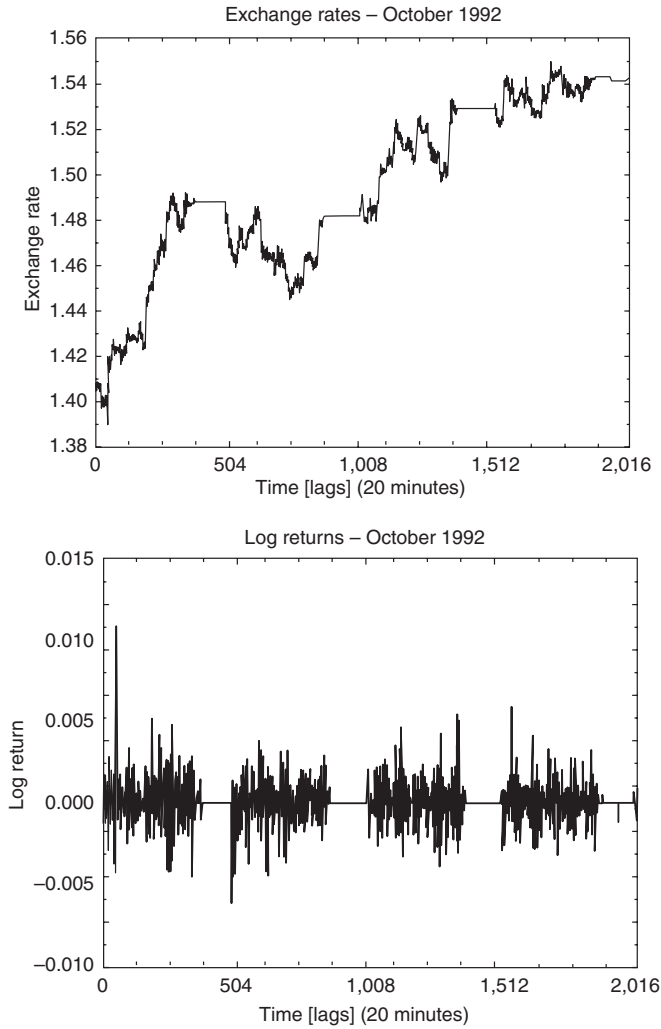


Figure 8.1 USD/DEM exchange rate, computed as a mid-price (upper) and as the associated logarithmic difference (lower), on a time scale of $\Delta t = 20$ minutes, October 5–November 2, 1992.

Source: USD/DEM from Reuters FXFX pages 5.10.1992–2.11.1992.

p. 6, figure 8). The time on the abscissa of the weekly figures in Figure 8.3 (lower) and 8.4 (lower) is measured in $7 \times 24 = 168$ hours per week. Time is measured in Greenwich Mean Time (GMT) and starts on Monday 0:00 GMT.

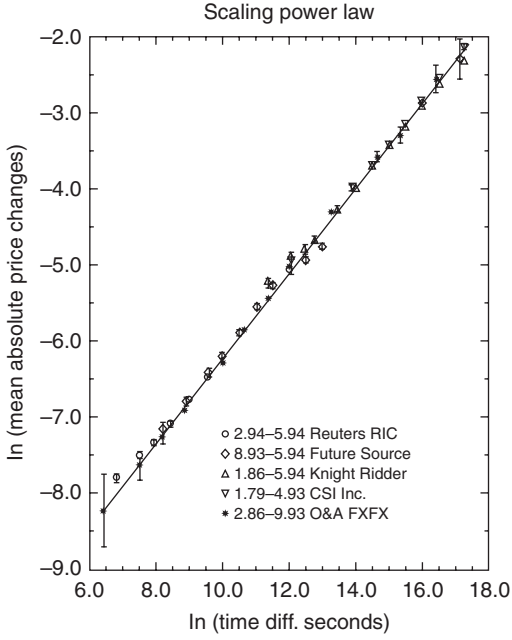


Figure 8.2 Scaling law behavior of the USD/DEM exchange rate in the period 1993–1994, for various subperiods: ln(mean absolute price changes) versus ln(time difference in seconds): $\ln E\{|\Delta P_t|\}/\ln \Delta t$.

Source: USD/DEM various vendors and time periods.

Notice that the three peaks in both Figures 8.3 (right) and 8.4 (right) relate to the maximum market activity in America (main peak), Asia (smallest peak), and Europe (small peak), respectively. There is very little trading during the weekends. The peaks in the trading activity in Figure 8.3 correspond with the peaks in price volatility or risk in Figure 8.4.

It is also observed that stock prices or foreign exchange rates are *singular at almost every point*, since their transaction records are essentially represented by step functions over time. The prices "jump" in small steps, because of small shifts in their respective supply and demand curves. The typical mechanism in price formation involves both knowledge of the present and expectations about the future. Even when the exogenous physical determinants of prices vary continuously, expectations can change drastically and instantaneously.

Such discontinuous price data are similar to particular physiological measurement data, such as heart records, electromagnetic fluctuations in galactic radiation noise, textures in images of natural terrain, variations of electric grid or traffic flows, etc. However, *not all singularities are alike!* Knowing the degree of irregularity

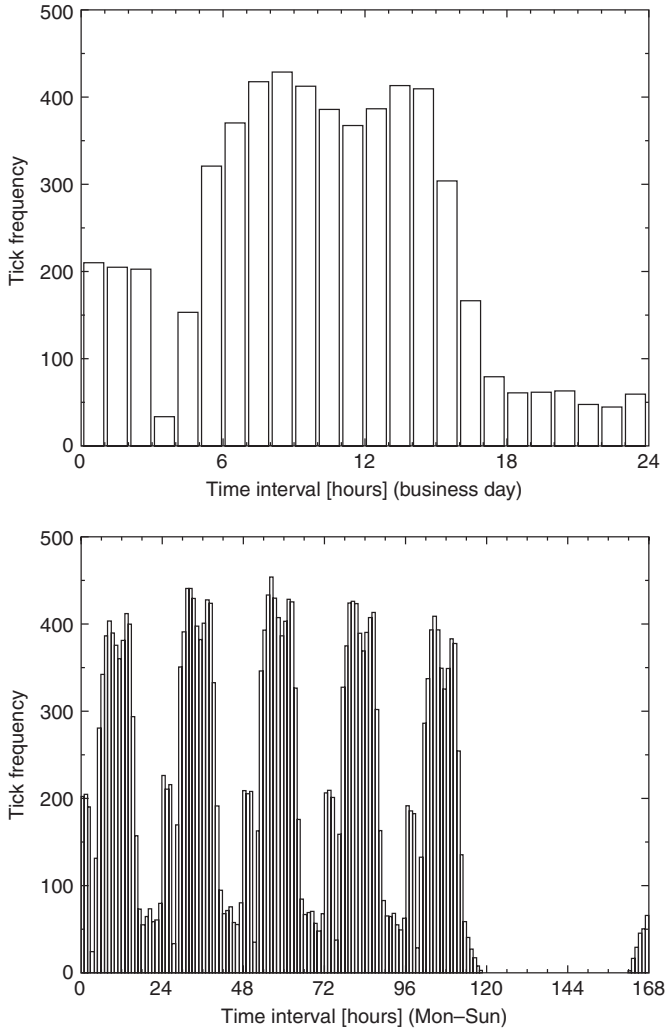


Figure 8.3 Trading transaction density: daily (upper) and weekly (lower) averaged number of ticks per hour. The tick labels are in GMT and start on Monday 0:00 GMT (October 5, 1992–September 29, 1993).

Source: USD/DEM from Reuters FXFX page 5.10.1992–26.9.1993.

of such discontinuities, or singularities, is important in analyzing their properties. In finance, knowing the distributions of the degrees of irregularity of financial time series is necessary for a correct analysis and valuation of the non-stationary, aperiodic, but cyclic financial risk.

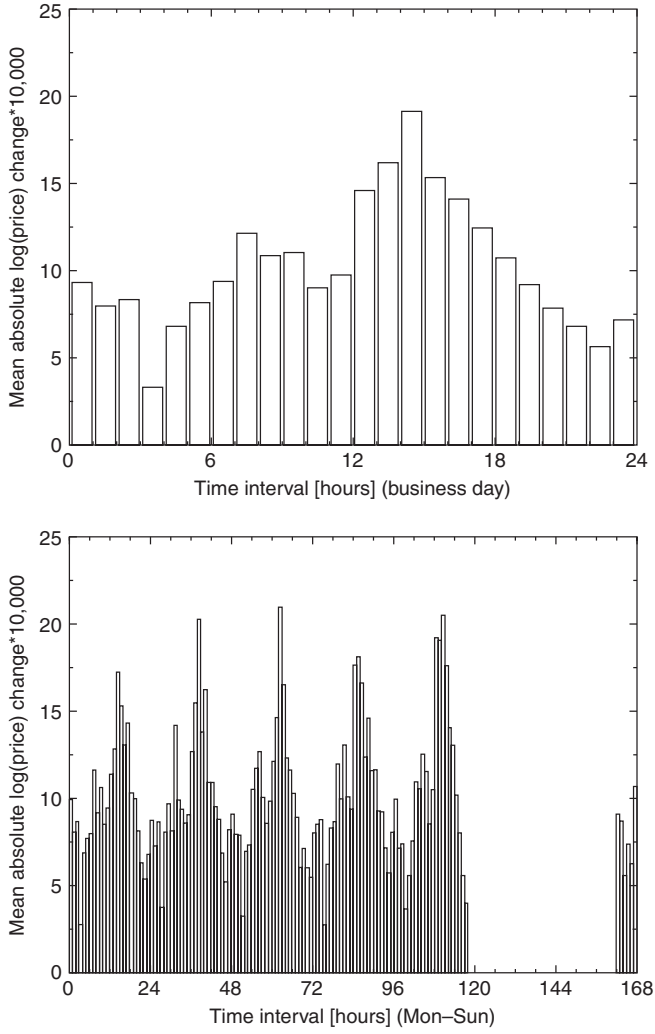


Figure 8.4 Daily (upper) and weekly (lower) averaged volatility per hour of the USD/DEM exchange rate. Tick labels are in GMT and start on Monday 0:00 GMT (October 5, 1992–September 26, 1993).

Source: USD/DEM from Reuters FXFX page 5.10.1992–26.9.1993

For this purpose, we return in this chapter to our original informal definition of irregularity or "randomness" of Chapter 1. This time we provide a proper formal definition of measurable irregularity, as measured by the Lipschitz (ir-) regularity exponent α_L.[3]

Pointwise measurements of Lipschitz regularity exponents, which measure the degree of irregularity of singularities, are not possible, because of finite numerical resolution of the empirical data. After discretization, each data set corresponds to a time interval where the time series has a very large (but finite) number of singularities. These singularities may show similarities, but they may also all be different. Such singularity distributions, or *singularity spectra* must therefore be computed from global measurements, which take advantage of *multifractal self-similarities* inherent in the financial data.

In the preceding chapters, we found that the Fractional Brownian Motion (FBM) provides a convenient uni-parametric model for such self-similar time series (cf. Chapter 4). FBMs are statistically self-similar i.i.d. processes, which exhibit long-term dependencies. Despite their non-stationarity, one can define a power spectrum, based on their stationary increments, that exhibits power decay. Thus, FBMs exhibit $1/\omega$-type spectral behavior over wide ranges of radian frequencies ω (cf. Chapter 4). Realizations of FBMs are almost everywhere singular, with the same *homogeneous* α_L-Lipschitz regularity at all points.

On the other hand, unlike FBMs, there exist fractal random processes that are not homogeneous α_L-Lipschitz irregular, although their power spectrum shows power decay. Empirical realizations of these processes may include increments of various types other than the familiar i.i.d. innovation processes of the classical Geometric Brownian Motion (GBM). Therefore, the computation of a complete singularity spectrum with a fractal dimension dependent on the Lipschitz regularity exponent α_L is important for such non-FBM cases, as we'll discuss at the end of this chapter.

To give a preview of the following topics: after discussing how to measure the irregularity or multifractal spectrum of time series $x(t)$, by implementing the wavelet MRA of Chapter 7, we'll discuss in the next chapter the phenomena of deterministic financial chaos. In Chapters 9 and 11 we'll present the current efforts to provide mathematical theories for financial turbulence. In Chapter 10 we'll give a simple example of a nonlinear dynamic interest rate term structure that demonstrates how financial intermittence and complete chaos can occur.

Financial turbulence theories are in debt to the original theory of physical turbulence of Kolmogorov of 1941, to his later amendment in 1962, and to the corrections by Mandelbrot in the 1970s and 1980s. Such theories are currently progressing far beyond Kolmogorov's fundamental insights, thanks to the analytic measurements provided by the wavelet MRA. One of the new insights by Frisch, Parisi and Farge is that turbulence is a heterogeneous multifractal phenomenon of which we can determine a multifractal spectrum of singularities from wavelet MRA. In addition, by using Galerkin's method of finite elements, wavelet MRA helps to numerically solve the dynamic Navier–Stokes nonlinear diffusion equations, which represent still the best dynamic model to explain turbulence, since it is based on fundamental mathematical and physical system laws.

8.2 Measurement of local financial market risk

8.2.1 Time-scale analysis of FBM

Due to the self-similarity, or scaling property of the FBM, Flandrin (1989, 1992) and Mallat (1989a) examined the FBM's behavior relative to different observational time scales, using self-similar wavelet MRA. A second-order moment analysis of the wavelet resonance coefficients of the FBM reveals a stationary structure at each scale and a power-law behavior of the wavelet coefficient's variance, from which the average Lipschitz exponent α_L of the FBM can be computed.

Reviewing the various aspects of the FBM that we discussed in Chapters 4 and 5, we first established that the autocovariance function of the FBM $x(t)$ is represented by:

$$\begin{aligned}\gamma(\tau) &= E\{x^2(\tau)\} \\ &= \sigma_\varepsilon^2 \tau^{2d-1} \\ &= \sigma_\varepsilon^2 \tau^{2H-2}\end{aligned} \tag{8.1}$$

which shows it to be *non-stationary*, and *self-similar*, since the second moment is a scaling law of the time lag τ. Next, we established that the average power spectral density of the FBM is:

$$\begin{aligned}P(\omega) &= \sigma_\varepsilon^2 \omega^{-2(d+1)} \\ &= \sigma_\varepsilon^2 \omega^{-(2H+1)}\end{aligned} \tag{8.2}$$

which is also a scaling law, this time of the frequency ω, or scale $a \sim 1/\omega$. Furthermore, the FBM is *statistically self-similar* in the sense that for any constant $c > 0$, and with the convention that $x(0) = 0$, we find the distributional scaling

$$\begin{aligned}x(c\tau) &\stackrel{d}{=} c^{d+0.5} x(\tau) \\ &= c^H x(\tau)\end{aligned} \tag{8.3}$$

where $\stackrel{d}{=}$ means equality in distribution, as discussed in Chapter 3. This means in frequency terms that the power spectrum of the FBM is represented by

$$\begin{aligned}\mathcal{F}[\gamma(c\tau)] &= \frac{1}{|c|^2} P\left(\frac{\omega}{c}\right) \\ &= \frac{\sigma_\varepsilon^2}{|c|^2} \left(\frac{\omega}{c}\right)^{-2(d+1)} \\ &= c^{2H-1} \sigma_\varepsilon^2 \omega^{-(2H+1)}\end{aligned} \tag{8.4}$$

Thus, any portion of a given FBM can be viewed as a scaled version of a larger part of the same process, both in the time domain and in the frequency domain.

Consequently, an individual realization of the FBM is a fractal time series and has a unique *fractal dimension* D, which is related to the Hurst H-exponent, as follows

$$D = 2 - H \tag{8.5}$$

In summary, the FBM has two important features:

(1) *non-stationarity*, which requires time-dependency analysis; and, more specifically,
(2) *self-similarity*, which requires time-scale power law analysis.

Since Mallat's wavelet MRA provides such a localized time-scale analysis, it is the natural tool to examine empirical FBMs.

8.2.2 Lipschitz analysis of local financial risk

The Fourier Transform analyzes the global regularity of a financial time series $x(t)$. The Wavelet Transform analyzes the pointwise irregularity of a financial time series $x(t)$. FBM traces are locally very irregular: they are continuous time series, but their first derivatives exist almost nowhere, i.e., their increments consist of singularities almost everywhere.

Definition 317 *A time series is called* regular *if it can be locally approximated by a polynomial, i.e., a particular mathematical system. If not, it is called* completely irregular.

Therefore, we must now introduce the formal definition of irregularity or "randomness."[4] It appears that there are degrees of local irregularity, from highly regular to highly irregular (Pincus and Singer, 1996). These degrees of irregularity are measured by the Lipschitz regularity exponent α_L.[5] If $x(t)$ has a singularity at time τ, which means that it is not differentiable at τ, the Lipschitz regularity exponent α_L characterizes this singular behavior at time τ. When we measure the Lipschitz α_L of a singularity, we assess how irregular or random such a singularity is. Consequently, we no longer have to assume the *degree of randomness* of a time series. We can measure the degree of its randomness by determining its Lipschitz α_L! In this section, we'll develop an apparatus to measure Lipschitz α_L, using Mallat's MRA from the preceding chapter.

The Lipschitz regularity exponent α_L is based on the approximation error of the Taylor expansion formula, which relates the differentiability of the continuous time series $x(t)$ to a local polynomial approximation.

Definition 318 *Suppose that $x(t)$ is d times differentiable in the bounded interval $[\tau - \epsilon, \tau + \epsilon]$ for a small ϵ. Then we can expand $x(t)$ in a Taylor expansion*

as follows

$$\begin{aligned} x(t) &= x(\tau) + x^{(1)}(\tau)(t-\tau) + \frac{x^{(2)}(\tau)}{2!}(t-\tau)^2 + \cdots \\ &\quad + \frac{x^{(d-1)}(\tau)}{(d-1)!}(t-\tau)^{d-1} + \epsilon_\tau(t) \\ &= \left[\sum_{k=0}^{d-1} \frac{x^{(k)}(\tau)}{k!}(t-\tau)^k\right] + \epsilon_\tau(t) \\ &= \hat{x}_\tau(t) + \epsilon_\tau(t) \end{aligned} \tag{8.6}$$

where $x^{(k)}(t)$ is the k^{th} derivative of $x(t)$. The $\hat{x}_\tau(t) = [\cdots]$ part is the exact polynomial Taylor expansion *of $x(t)$ at time τ, or* systematic *component, and the $\epsilon_\tau(t) = x(t) - \hat{x}_\tau(t)$ is the approximation error, or* unsystematic *component, of this Taylor expansion.*

Remark 319 *Statisticians often call the approximation error $\epsilon_\tau(t)$: the* residual. *It is clear that the character of this residual depends on the number of differentiation terms included in the linear Taylor expansion. Therefore, one cannot ascribe inherent characteristics like "Gaussian distribution" to this residual, since such characteristics are* not sui generis. *Still, this is what statisticians conventionally (conveniently, but unfortunately) do!*

The Taylor expansion proves that the approximation error

$$\epsilon_\tau(t) = x(t) - \hat{x}_\tau(t) \tag{8.7}$$

satisfies

for all $t \in [\tau - \epsilon, \tau + \epsilon]$,

$$|\epsilon_\tau(t)| \leq \sup_{u \in [\tau-\epsilon, \tau+\epsilon]} \left| x^{(d)}(u) \right| \frac{|u-\tau|^d}{d!} \tag{8.8}$$

The dth-order differentiability of $x(t)$ in the neighborhood of τ yields an upper bound on the approximation error $\epsilon_\tau(t)$ when t tends to τ, i.e., when the time interval becomes smaller. The following Lipschitz regularity refines this upper bound with the fractional Hölder exponent d, introduced in Chapter 4.

Definition 320 (Lipschitz) *A time series $x(t)$ is* pointwise α_L-Lipschitz regular, *with regularity exponent $\alpha_L \geq 0$ at point τ, if there exists a $K > 0$, and a polynomial $\hat{x}_\tau$ of degree $\lfloor \alpha_L \rfloor$ such that for all real time $t \in \mathbb{R}$, the absolute value*

of the error is bounded by:

$$\begin{aligned} |\epsilon_\tau(t)| &= |x(t) - \hat{x}_\tau(t)| \\ &\leq K|t-\tau|^d \\ &= K|t-\tau|^{\alpha_L} \end{aligned} \tag{8.9}$$

or, equivalently:

$$|\epsilon_\tau(t)|^{1/\alpha_L} = |x(t) - \hat{x}_\tau(t)|^{1/\alpha_L} \leq K'|t-\tau| \tag{8.10}$$

where $K' = K^{1/\alpha_L}$.

Definition 321

- *A time series* $x(t)$ *is* uniformly α_L-Lipschitz regular *over the interval* $[a, b]$ *if it is pointwise Lipschitz* α_L *for all* $\tau \in [a, b]$*, with a constant* K *that is independent of* τ.
- *The Lipschitz regularity exponent of* $x(t)$ *at point* τ *or over the interval* $[a, b]$ *is the* supremum *of* α_L *such that* $x(t)$ *is* α_L*-Lipschitz regular (pointwise or uniformly).*

This is a (very) technical definition of irregularity and a new definition of local (financial) risk, which requires some additional explication. At each time point τ, the polynomial $\hat{x}_\tau(t)$ is uniquely defined. If $x(t)$ is $d = \lfloor \alpha_L \rfloor$ times continuously differentiable in the neighborhood of τ, then $\hat{x}_\tau(t)$ equals the linear Taylor expansion of $x(t)$ at τ. Thus, when α_L is an integer, the regularity at point τ is defined as usual, with α_L indicating the order of differentiability of $x(t)$.

When α_L is not an integer, but a fraction, let d be an integer such that $d < \alpha_L < d + 1$, then $x(t)$ has an α_L- Lipschitz regularity at τ, if its derivative $x(t)^{(d)}$ of order d resembles $|t - \tau|^{\alpha_L - d}$ locally around point τ. Furthermore, the degree of regularity of $x(t)$ in a time domain is that of its least regular point. The greater α_L, the more regular is the time series $x(t)$. The smaller α_L, the more irregular, or "risky," is the time series $x(t)$.

Remark 322 *There exist* multifractal *time series* $x(t)$ *with non-isolated singularities, where* $x(t)$ *has a different Lipschitz* α_L *at each point* τ*. In contrast, uniform Lipschitz* α_L *exponents provide a more global measurement of regularity, which applies to a whole interval. If a time series* $x(t)$ *is uniformly Lipschitz* $\alpha_L > d$*, or monofractal, where* d *is an integer, then one can verify that* $x(t)$ *is* d *times continuously differentiable in that neighborhood.*

What values of the Lipschitz α_L exponent should one expect for the various kinds of singularities? If $0 \leq \alpha_L < 1$, then $\hat{x}_\tau(t) = x(\tau)$ and the Lipschitz condition becomes:

$$\text{for all } t \in \mathbb{R}, \quad |x(t) - x(\tau)| \leq K|t-\tau|^{\alpha_L} \tag{8.11}$$

A time series $x(t)$ that is bounded, but discontinuous at time τ is Lipschitz $\alpha_L = 0$ at the time of the discontinuity τ. If the Lipschitz regularity is $0 < \alpha_L < 1$ at

τ, then $x(t)$ is continuous, but not differentiable at τ and the α_L characterizes the degree, or type, of irregularity. If $\alpha_L = -1$, the discontinuity "flip-flops" (cf. Figure 8.7).

What is the Lipschitz regularity condition for the Fourier Transform and for the Wavelet Transform, respectively?

8.2.2.1 Fourier regularity condition

The precise definition of the Lipschitz-α_L regularity in the frequency domain, in addition to the one we already have in the time domain, is provided by the following theorem.[6]

Theorem 323 *A function $x(t)$ with Fourier Transform $F(\omega)$ is* bounded and uniformly Lipschitz-α_L *over the domain of real numbers $\mathbb{R}$, if*

$$\int_{-\infty}^{+\infty} |F(\omega)|(1 + |\omega|^{\alpha_L})d\omega < +\infty \tag{8.12}$$

Remark 324 *This uniform regularity condition is obviously a* global *regularity condition, since it holds true over the whole $(-\infty, +\infty)$ frequency domain.*

Next, we will discuss the required regularity condition of wavelets.

8.2.2.2 Wavelet regularity condition

The basic wavelet regularity condition is that it is a *fast decaying* wavelet with p vanishing moments.[7] In fact, if a function is continuous, has vanishing moments, decays quickly towards 0 when $t \to \infty$, or equals 0 outside a particular interval, it is already a likely candidate for a wavelet!

Theorem 325 *A wavelet $\psi(t)$ with a* fast decay *has p vanishing moments, if and only if there exists a function $\theta(t)$ with a fast decay such that*

$$\begin{aligned} \psi(t) &= (-1)^p \theta^{(p)}(t) \\ &= (-1)^p \frac{d^p \theta(t)}{dt^p} \end{aligned} \tag{8.13}$$

As a consequence, the CWT or resonance coefficient is equivalent to the following *multiscale differential operator*

$$\begin{aligned} W(\tau, a) &= \int_{-\infty}^{\infty} x(t)\psi^*_{\tau,a}(t)dt \\ &= a^p \frac{d^p}{d\tau^p}\{[x(t) \star \theta_a(t)](\tau)\} \end{aligned} \tag{8.14}$$

with the scaled wavelet

$$\theta_a(t) = \frac{1}{\sqrt{a}}\theta\left(\frac{t}{a}\right) \tag{8.15}$$

Table 8.1 Degree of Lipschitz irregularity of Daubechies wavelets

ψ	DB1(= *Haar*)	DB2	DB3	DB4	DB5	DB7	DB10
α_L	0	0.5	0.91	1.27	1.69	2.15	2.90

where the $\star$ sign indicates again the convolution operator (cf. Chapter 5). This provides a test whether the wavelet ψ has more than p vanishing moments. Compute

$$\int_{-\infty}^{+\infty} t^p \psi(t)\, dt = (i)^p \hat{\psi}^{(p)}(0) = (-i)^p p! \hat{\theta}(0) \tag{8.16}$$

Clearly, the wavelet $\psi(t)$ has no more than p vanishing moments if and only if:

$$\hat{\theta}(0) = \int_{-\infty}^{+\infty} \theta(t) dt \neq 0 \tag{8.17}$$

Remark 326 *An example of such a fast decaying wavelet is the Gábor's Gaussian wavelet, or chirp, discussed in Chapter 6.*

The degree of irregularity of certain wavelets is known. Table 8.1 gives some indications of the Lipschitz-α_L irregularity of Daubechies wavelets indexed by DBN.

Selecting an irregularity and a wavelet to measure this irregularity is useful for estimations of the local properties, like the *intrinsic* or *local risk*, of a financial time series. From a practical point of view, these questions arise in finance in dealing with financial markets for fine microstructure studies of high-frequency (= very fast) trading transactions. Let's now see, how we can approach the measurement of the irregularity or intrinsic risk of such high-frequency financial transactions.

If $x(t)$ is a financial time series which is a little bit more than p times differentiable at point τ, then it can be approximated by a polynomial of degree p, as we've already seen in Chapter 4. For example, it can be approximated by a Markov process of order p, which can represent trends and regular periodic oscillations. As we noticed earlier, the Wavelet Transforms of such exact polynomials are zero. But around point τ, its order is that of the *error* between the polynomial and the time series $|x(t) - \hat{x}_\tau(t)|$. If this error can be uniformly estimated on an interval $[a, b]$, this insight yields a tool for irregularity or local risk analysis on that interval and we can estimate the *fractal order* d of the financial time series $x(t)$.

8.2.3 Asymptotically decaying wavelet amplitudes

The decay of the Wavelet Transform amplitude across all scales is related to the uniform and pointwise Lipschitz regularity of the financial time series $x(t)$. Measuring this asymptotic decay is equivalent to zooming into the time series structure with a scale that goes to zero. The following theorems relate the uniform and pointwise Lipschitz regularity of $x(t)$ on an interval to the amplitude of its Wavelet Transform at the very fine scales. If we then measure the amplitude of the measured decaying wavelet resonance coefficients, we can find out what the Lipschitz α_L regularity of $x(t)$ is and thus its fractal difference order d.

Theorem 327 (Mallat) *If $x(t) \in L^2(\mathbb{R})$, an element of the Hilbert space, and is uniformly Lipschitz $\alpha_L \leq p$ over the interval $[a, b]$ then there exists a constant A such that*

$$\begin{aligned} &\text{for all } (\tau, a) \in [a, b] \times \mathbb{R}^+, \\ &\quad |W(\tau, a)| \leq A a^{\alpha_L + 0.5} \\ &\qquad\quad = A a^{d + 0.5} \end{aligned} \tag{8.18}$$

Conversely, if $|W(\tau, a)|$ satisfies this last inequality and if α_L is not an integer, but a fraction $\alpha_L < p$, then the time series $x(t)$ is uniformly α_L-Lipschitz on the interval $[a + \epsilon, b - \epsilon]$, for any $\epsilon > 0$.

Remark 328 *The inequality is really a condition on the asymptotic decay of the absolute value of the wavelet resonance coefficient $|W(\tau, a)|$, when its scale a goes to zero. At larger scales it does not introduce any constraints since the Cauchy–Schwarz inequality guarantees that the Wavelet Transform is always bounded:*

$$\begin{aligned} |W(\tau, a)| &= |\langle x(t), \psi_{\tau,a}(t) \rangle| \\ &\leq \|x(t)\| \|\psi_{\tau,a}(t)\| \end{aligned} \tag{8.19}$$

where the norms (= risk contents) $\|x(t)\| < \infty$ and $\|\psi_{\tau,a}(t)\| < \infty$.

Jaffard (1991) generalized Mallat's Theorem to pointwise Lipschitz regularity, while Mallat's Theorem can be viewed as a corollary of Jaffard's Theorem (Jaffard, 1989; Farge, *et al.*, 1993). Jaffard's Theorem provides a necessary and a sufficient condition on the modulus of the Wavelet Transform for computing the Lipschitz regularity of $x(t)$ at point τ.

Theorem 329 (Jaffard) *If $x(t) \in L^2(\mathbb{R})$ is Lipschitz $\alpha_L \leq p$ at point υ in time, then there exists a constant A such that*

$$\text{for all } (\tau, a) \in \mathbb{R} \times \mathbb{R}^+,$$
$$|W(\tau, a)| \leq A a^{\alpha_L + 0.5} \left(1 + \left|\frac{\tau - \upsilon}{a}\right|^{\alpha_L}\right) \tag{8.20}$$

Conversely, if $\alpha_L < p$ is not an integer and there exist A and $\alpha_L^j < \alpha_L$ such that

$$\text{for all } (\tau, a) \in \mathbb{R} \times \mathbb{R}^+,$$
$$|W(\tau, a)| \leq A a^{\alpha_L + 0.5} \left(1 + \left|\frac{\tau - \upsilon}{a}\right|^{\alpha_L}\right) \tag{8.21}$$

then $x(t)$ is Lipschitz α_L at υ.

To interpret more easily the necessary and sufficient conditions of Jaffard's Theorem, suppose that the wavelet $\psi_{u,a}(t)$ has a compact support equal to $[-C, C]$. Then we can formulate the following definition of the *cone of influence* of a particular point υ on the time line.

Definition 330 *The* cone of influence *of υ in the scale-time plane is the set of points (τ, a) such that υ is included in the support of the CWT wavelet*

$$\psi_{\tau,a}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t - \tau}{a}\right) \tag{8.22}$$

Since the support of this wavelet is equal to $[\tau - Ca, \tau + Ca]$, the cone of influence of υ is defined by

$$|\tau - \upsilon| \leq Ca \tag{8.23}$$

or, equivalently,

$$\frac{|\tau - \upsilon|}{a} \leq C \tag{8.24}$$

For example, Figure 8.5 shows the cone of influence for a time abscissa at $t = v$ in a scalogram.

Example 331 *Figure 8.6 shows the* regions of influence *in scalograms (a and c) for the CWT and spectrograms (b and d) for Gábor's Short-Term or Windowed Fourier Transform (STFT) for the singularity of a Dirac pulse $\delta(\omega)$ at time $t = t_0$, as well as three sinusoids of frequencies $\omega_0 = \omega_0$, $\omega_1 = 2\omega_0$, $\omega_3 = 4\omega_0$,*

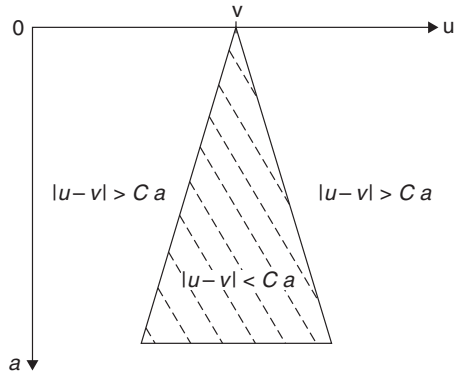


Figure 8.5 The cone of influence of an abscissa singularity v consists of the time-scale points (u, a) for which the support of the wavelet $\psi_{u,a}$ intersects the time point $t = v$.

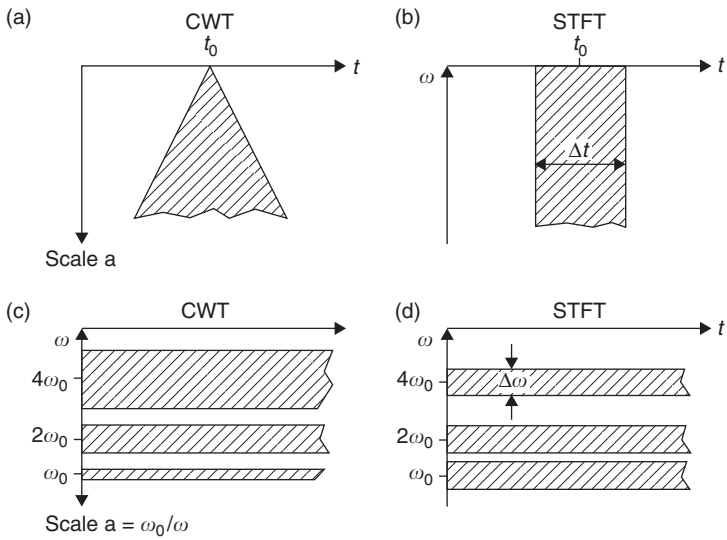


Figure 8.6 Singularity cones of influence of a Dirac pulse at $t = t_0$ for (a) the CWT and for (b) the STFT; versus monochromatic bands of influence of three sinusoids for (c) the CWT and for (d) the STFT.

respectively (cf. Chapter 6). The scale is $a = \omega_0/\omega$. Notice that for the CWT in (a) the width of the cone of influence of the Dirac pulse is scale-frequency dependent, while for the STFT in (b) it remains constant and is scale-frequency independent. Furthermore, for the CWT in (c) the monochromatic resonance bands differ in

width according to the scale-frequency, while for the STFT in (d) the width of the monochromatic resonance bands are scale-frequency independent.

Using a reformulation of the cone of influence, Jaffard's necessary and sufficient conditions can be considerably simplified, as follows:

$$\begin{aligned} |W(\tau, a)| &\leq A a^{\alpha_L + 0.5} \left(1 + \left|\frac{\tau - \upsilon}{a}\right|^{\alpha_L}\right) \\ &\leq A a^{\alpha_L + 0.5}(1 + C^{\alpha_L}) \\ &= A' a^{\alpha_L + 0.5} \end{aligned} \tag{8.25}$$

which is identical to the uniform Lipschitz condition of Mallat's Theorem!

Jaffard's Theorem relates the pointwise irregularity, e.g., the singularities, of a time series to the decay of the *modulus maximum* of its Wavelet Transform $|W(\tau, a)|$. This will assist us with the measurement of the degree of local financial risk.

8.2.4 Measuring various price singularities

We will now discuss the innovative concepts of modulus maxima and the maxima line to detect and measure the various kinds of singularities (Hwang and Mallat, 1994).

Definition 332 *The* modulus maximum *of a Wavelet Transform is any point* (τ_0, a_0) *such that* $|W(\tau, a)|$ *is locally maximum at* $\tau = \tau_0$. *This implies*

$$\frac{\partial |W(\tau_0, a_0)|}{\partial \tau} = 0 \tag{8.26}$$

Definition 333 *The* maxima line *is any connected curve* $a(\tau)$ *along the scale ordinate in the time-scale plane* (τ, a) *along which all points are modulus maxima.*

Singularities can thus be detected by finding the abscissa where the wavelet modulus maxima converge at the very fine scales of, say, high-frequency financial data. The following Theorem by Hwang and Mallat proves that if the Wavelet Transform $W(\tau, a)$ has no modulus maxima at fine scales, then the time series $x(t)$ is locally regular. Otherwise stated, there cannot be a singularity without a local maximum of the Wavelet Transform at the very fine scales.

Theorem 334 (Hwang, Mallat) *Suppose that the wavelet* $\psi(t)$ *is* C^p *(= continuous of order p) with a fast decay, has p vanishing moments with*

compact, finite support, and

$$\psi(t) = (-1)^p \theta^{(p)}(t) \tag{8.27}$$

with the Gaussian wavelet $\theta(t)$ *such that*

$$\int_{-\infty}^{+\infty} \theta(t)dt \neq 0 \tag{8.28}$$

Let $x(t) \in L^1[c, d]$. *If there exists* $a_0 > 0$ *such that* $|W(\tau, a)|$ *has no local maximum for* $\tau \in [c, d]$ *and* $a < a_0$, *then* $x(t)$ *is uniformly Lipschitz* p *on* $[c + \epsilon, d - \epsilon]$ *for any* $\epsilon > 0$.

This important and insightful theorem implies that $x(t)$ can be singular (= not Lipschitz-1) at a point υ only if there is a sequence of wavelet maxima points that converges towards the time point υ at very fine scales:

$$\lim_{n \to +\infty} \tau_n = \upsilon \quad \text{and} \quad \lim_{n \to +\infty} a_n = 0 \tag{8.29}$$

This sequence of wavelet maxima indicates the presence of a maximum modulus of the Wavelet Transform at the very fine scales where a singularity occurs.

In the general case, a sequence of modulus maxima, or maxima line, may be detected, which converges to the particular singularity. When the wavelet is the pth derivative of a Gaussian wavelet $\theta(t)$ (= Gábor wavelet), these maxima lines are connected and go through all of the finer scales. The decay rate of the maxima along the maxima ridges indicates the order of the isolated singularities. This can be easily shown, since from the Jaffard Theorem for $\tau = \upsilon$, we have for the log–log inequality:

$$\log_2 |W(\tau, a)| \leq (\alpha_L + 0.5) \log_2 a + \log_2 A' \tag{8.30}$$

Thus, one should display the modulus maxima of the Wavelet Transform as a function of scale a in a log–log plot, and its computed slope will be $b = \alpha_L + 0.5$, from which we then can immediately identify the Lipschitz α_L. For example, when this slope is $b = \alpha_L + 0.5 = 0.5$, the time series is Lipschitz $\alpha_L = 0$ and, thus, exhibits a discontinuity. But when the slope $b = \alpha_L + 0.5 = 1$, the time series is Lipschitz $\alpha_L = 0.5$. In other words, the degree of irregularity (= "randomness") of each singularity can be separately assessed. No longer have we to assume that some price singularity is random. We can now precisely locally measure its degree of irregularity, randomness or riskiness, as the following examples demonstrate!

Example 335 *Figure 8.7 provides the first example of this kind of financial risk analysis. In the top panel (a) we see a time series or signal* $S_1 = x(t)$, *which*

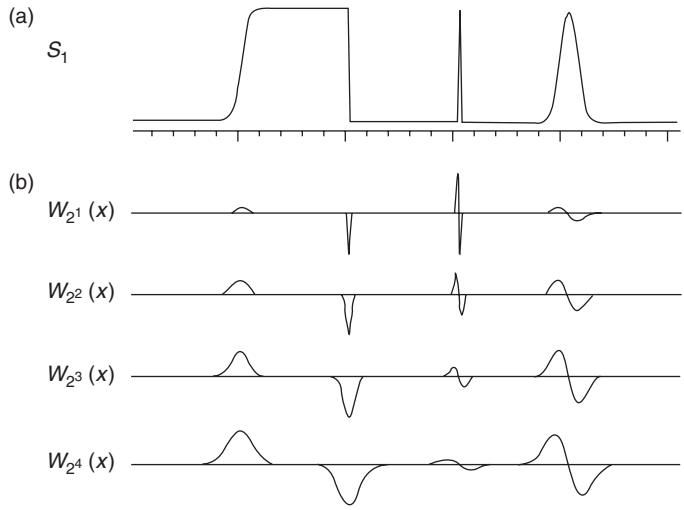


Figure 8.7 Wavelet decomposition of a time series with singularities. In the top panel (a) we see a time series or signal $S_1 = x(t)$, which contains four singularities, characterized, respectively by the Lipschitz α_L and the smoothing scale s: $(\alpha_L = 0, s = 0)$, $(\alpha_L = 0, s = 3)$, $(\alpha_L = -1, s = 0)$ and $(\alpha_L = -1, s = 4)$. In the lower panel (b) we see four scales of wavelet decomposition.

contains four singularities, characterized, respectively by the Lipschitz α_L and the smoothing scale s: $(\alpha_L = 0, s = 0)$, $(\alpha_L = 0, s = 3)$, $(\alpha_L = -1, s = 0)$, and $(\alpha_L = -1, s = 4)$. In the lower panel (b) we see four scales of wavelet decomposition (Mallat and Zhong, 1992, p. 86). It is clear that the behavior of the local maxima across the wavelet scales depend on the Lipschitz α_L and the smoothing scale s. Notice that the "sharpest" singularity, $(\alpha_L = 0, s = 3)$, and $(\alpha_L = -1, s = 0)$ are best detected at scale level $a = 1$, while the "softest" singularity, $(\alpha_L = 0, s = 0)$ and $(\alpha_L = -1, s = 4)$ are best detected at scale level $a = 4$. But each scale level provides its own piece of information about each of the different singularities.

Example 336 *Figure 8.8 provides a more complex financial risk analysis. At the top we have 256 observations of the irregular time series $x(t)$, which shows different kinds of singularities: from step functions at the left, to a sharp peak in the middle, followed by a discontinuity and a very "random" looking series. The question is how we can characterize these singularities of $x(t)$ using the scalogram based on the CWT $W(\tau, a)$. Panel (a) shows the scalogram $P_W(\tau, a)$. The horizontal and vertical axes measure t and $\log_2 a$, where a is the dyadic scale. Panel (b) shows the modulus maxima of the $W(\tau, a)$. Notice that the "random" looking series part is represented by a series of modulus maxima. The continuous*

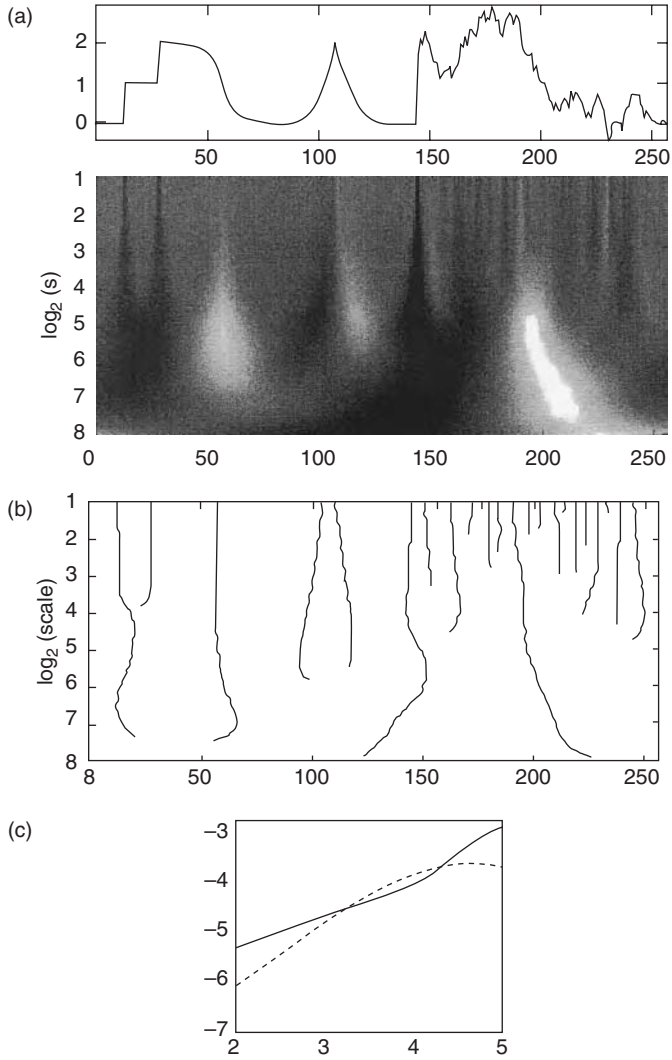


Figure 8.8 How to measure the degree of irregularity of local risk of a series of price singularities $x(t)$: (a) the scalogram $P_W(\tau, a)$, (b) the modulus maximae of the $W(\tau, a)$, (c) continuous line: the decay of $\log_2 |W(\tau, a)|$ as a function of $\log_2 a$ along the most left maxima line that converges on the abscissa point $t = 14$. Dashed line: the decay of $\log_2 |W(\tau, a)|$ along the maxima line that converges to time point $t = 108$.

line in panel (c) gives the decay of $\log_2 |W(\tau, a)|$ *as a function of* $\log_2 a$ *along the most left maxima line that converges to the abscissa* $t = 14$. *The dashed line in panel (c) gives the decay of* $\log_2 |W(\tau, a)|$ *along the maxima line that converges to* $t = 108$ *(Mallat, 1998, p. 180).*

8.3 Homogeneous Hurst exponents of monofractal price series

Now, let's discuss how this mathematical apparatus of wavelet regularity conditions can be used: first, to compute the *uniform* Lipschitz regularity exponent α_L, or, equivalently, the *homogeneous* Hurst H-exponent for our main model for financial risk, the FBM. This is followed by computation of the pointwise Lipschitz α_L to characterize the *singularity* or *local risk spectrum* of multifractal non-FBM processes. We begin with simple computations of homogeneous H-exponents for the FBM, using the wavelet detail coefficients $d_{j,n}$ from Mallat's MRA. Next, we will compute a multifractal spectrum of *heterogeneous* H-exponents or α_L's for non-FBM processes.

The heterogeneous or multifractal local risk spectrum characterizes the scaling and singularity structures of time series and has already proved to be a useful tool for numerous applications, from (electric) network traffic analysis to the analysis of turbulence in high-frequency financial time series. The computation of the complete local multifractal risk spectrum from a finite data record has long escaped the capability of the turbulence researcher, but the preceding general irregularity analysis shows that this kind of local risk analysis is now completely possible.

8.3.1 Logarithmic scalegram based on discrete wavelet MRA

Let's begin with the computation of the homogeneous Hurst exponent of global financial risk. When we discussed the MRA in Chapter 7, we stated that the *Discrete wavelet Transform (DWT) coefficient* of the FBM $x(t)$ is computed as the inner product of $x(t)$ and the basic discrete, dyadic, orthonormal wavelet ψ, e.g., the Haar wavelet, by the usual approach, as follows

$$d_{j,n} = 2^{-j/2} \int_{-\infty}^{+\infty} x(t)\psi(2^{-j}t - n)dt, \quad \text{with} j, n \in \mathbb{Z} \tag{8.31}$$

The tiling of the time-scale space by the resulting wavelet Heisenberg boxes is shown in Figure 8.9. For example, the $d_{0,0}$ coefficient represents the mean of $x(t)$. These wavelet resonance coefficients $d_{j,n}$ of an FBM have the following four properties, as proved by Flandrin (1992) and by Flandrin and Gonçalvès (1996):

(1) The wavelet resonance coefficients are *stationary in distribution*, i.e., they are *stably distributed:*

$$d_{j,n} \stackrel{d}{=} d_{j,0} \quad \text{for all } n \tag{8.32}$$

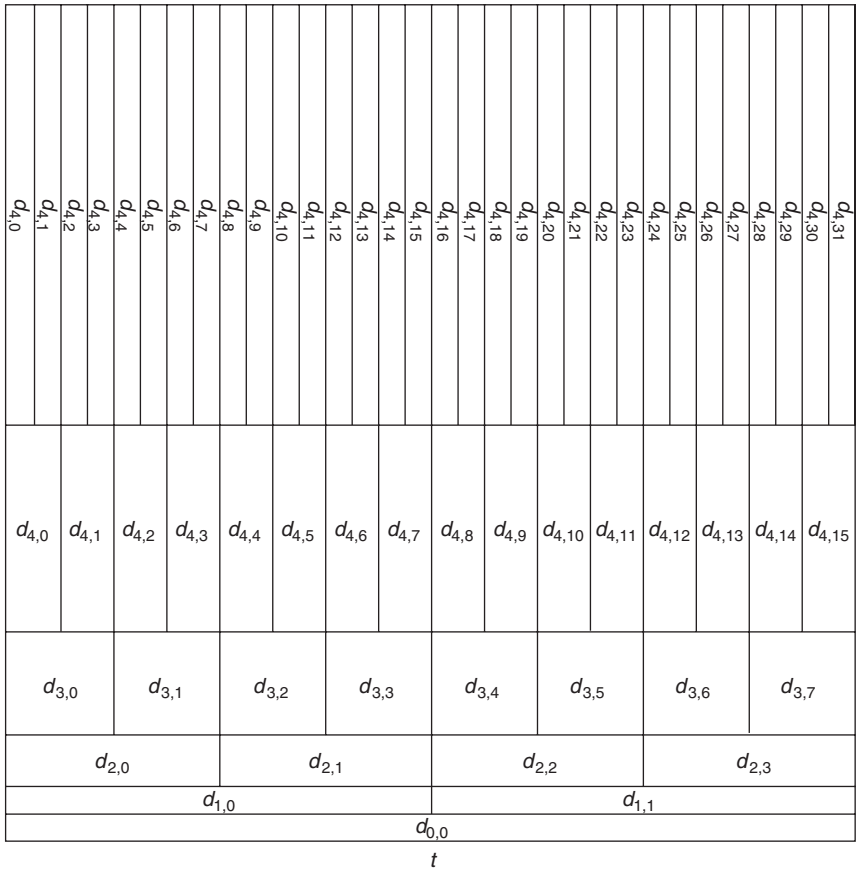


Figure 8.9 Complete wavelet tiling: the relation of the DWT coefficients $d_{j,n}$ to the time-scale tiles. This demonstrates the completeness of a financial risk analysis by wavelet MRA.

(2) The wavelet resonance coefficients are *Gaussian distributed:*

$$d_{j,n} \sim N(0, \mathrm{Var}(d_{j,n})) \tag{8.33}$$

Remark 337 *Consequently, the squared resonance coefficients, or local wavelet risk, are Chi-squared distributed:*

$$|d_{j,n}|^2 \sim \chi^2 \tag{8.34}$$

This statistical property allows for significance testing of the wavelet coefficients.

Remark 338 *If one wants to verify that the statistical distribution for the local wavelet risk is, indeed, Chi-squared distributed, one can apply the following Monte Carlo method: (1) create a large number, say 100,000 random time series, each with as many points as the financial time series* $x(t), t = 1, \ldots, T$ *to be analyzed; (2) then take the Wavelet Transform for each of the random time series and compute all the local wavelet risks* $|d_{j,n}|^2$*; next, (3) take a time slice from the middle (time* $n = T/2$*); (4) at each scale* j*, sort all selected 100,000 local wavelet risks into increasing order; (5) then make a plot of the local wavelet risk versus the sorted index number; (6) look at what the local wavelet risk is for number 95,000 out of 100,000, then 95 percent of the local wavelet risk is below that value, and only 5 percent is above; (7) this 95 percent level is the staistician's conventional 95 percent confidence level (or 5 percent significance level). This Monte Carlo method can be generalized to any process where the statistical distribution is unknown, yet one wants to determine statistical confidence, or significance levels.*

Flandrin (1989, 1992) and Kaplan and Kuo (1993) also proved that the variance of these wavelet resonance coefficients $d_{j,n}$ of the FBM is represented by the following scaling law, which is the integration of the Chi-squared distribution of the squared resonance coefficients:

$$\begin{aligned} \text{Var}\{d_{j,n}\} &= E\{|d_{j,n}|^2\} \\ &= \frac{\sigma_\varepsilon^2}{2} V_\psi(H)(2^j)^{-(2H+1)} \end{aligned} \tag{8.35}$$

where the constant $V_\psi(H)$ depends on both the ACF $\gamma_\psi(\tau)$ of the chosen wavelet $\psi(t)$ and the H-exponent, as follows:

$$V_\psi(H) = -\int_{-\infty}^{+\infty} \gamma_\psi(\tau)|\tau|^{2H} d\tau \tag{8.36}$$

with

$$\gamma_\psi(\tau) = \int_{=\infty}^{+\infty} \psi(t)\psi(t-\tau)dt \tag{8.37}$$

Thus, by taking the dyadic logarithm of $\text{Var}\{d_{j,n}\}$, we find the linear relationship from which we can compute H

$$\log_2[\text{Var}\{d_{j,n}\}] = -(2H+1)j + \log_2\left[\frac{\sigma_\varepsilon^2}{2} V_\psi(H)\right] \tag{8.38}$$

Since the second (intercept) term is a constant, we plot $\log_2[\mathrm{Var}\{d_{j,n}\}]$ against the scale coefficient j to find the slope value $(2H + 1)$ and thus H (Wornell and Oppenheim, 1992; Wornell, 1993). This average or global wavelet (log) scalegram delivers the same analytic irregularity measurement result of the homogeneous Hurst exponent H as the (log) Fourier power spectrum.

The other two properties of importance for this MRA analysis:

(3) The wavelet resonance coefficients are *almost uncorrelated*:

$$E\{d_{i,n}d_{j,m}\} \simeq |2^{-i}n - 2^{-j}m|^{2(H-R)} \text{ and} \tag{8.39}$$

(4) The wavelet resonance coefficients *scale:*

$$d_{j,n} \stackrel{d}{=} 2^{jH} d_{0,n} \tag{8.40}$$

Remark 339 *Because of property (2) and (3) it is often asserted that the FBM wavelet resonance coefficients are exactly uncorrelated and hence independent (Gonçalvès* et al.*, 1998). But this is, strictly speaking, not true since they scale. They are stably distributed (cf. Chapter 4).*

8.3.2 Scalegrams of heart arrhythmias and stock prices

We will now discuss two applications of the preceding analysis to compute homogeneous Hurst exponents from the dyadic logarithmic plot of wavelet resonance coefficients: a wavelet MRA of the heartbeat of a healthy human and of the Dow Jones Industrial Average Index (DJIA). The heartbeat inter-arrival times resemble those of foreign exchange quotations. Both these exemplary analyses are from Flandrin (1992). The persistence of the DJIA has also been studied by Lo and Mackinlay (1999) by non-wavelet, econometric time series methods. They produced similar results. In addition, we include some interesting scalograms and scalegrams based on wavelet MRA of Latin American financial markets around the times of major trading regime changes.

8.3.2.1 Scalegram of heart beats

Figure 8.10 shows the computation of the global or homogeneous Hurst exponent for heartbeat inter-arrival times $X(t)$ in seconds for a healthy human patient, of which 65,536 heartbeats are shown in the top panel. It is clear that this heart is not strictly periodic and that it shows arrhythmias: it is aperiodic cyclical. Fractal analysis is thus warranted. The time series is again analyzed using a Daubechies(5) wavelet basis and MRA tiling. In panel (a) the $\log_2[\mathrm{Var}\{d_{j,n}\}]$ is plotted versus the scale $j = m$. The scale in this panel (a) is such that a low scale j means low frequency while a high scale means low frequency. The approximate slope of the

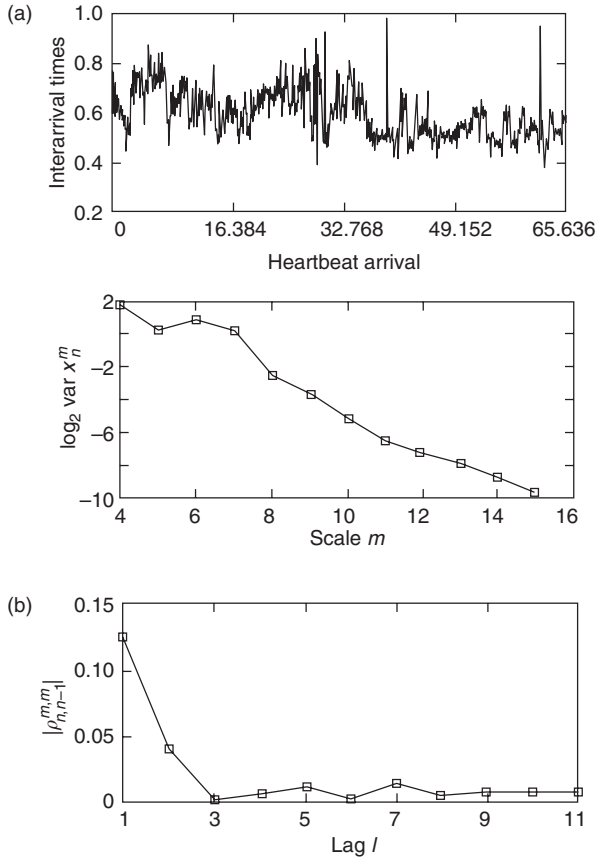


Figure 8.10 Wavelet-based persistence analysis of heartbeat interarrival times for a healthy patient with a Daubechies(5) wavelet. Top panel: data; (a) scale-to-scale wavelet coefficient variance progression; (b) average magnitude of normalized along-scale correlation between wavelet coefficients.

line that is not completely straight is:

$$\begin{aligned} b &= (2H + 1) \\ &= \frac{2 - (-10)}{15 - 4} \\ &= 1.090\,9 \end{aligned} \tag{8.41}$$

from which we derive the Hurst exponent $H = 0.045\,45$. This implies that the heartbeat inter-arrival times are almost blue noise, i.e., highly antipersistent. Differently stated, the human heart self-reverses or corrects itself (almost) immediately.

It is an extremely efficient pump. Notice, however, that the slope of the dyadic plot line is not strictly straight and thus the Hurst exponent is not strictly homogeneous, in particular around scales $a = 2^6 = 64$ seconds, or close to one minute, and $a = 2^7 = 128$ seconds, or close to two minutes. Panel (b) shows again the average magnitude of the normalized along-scale empirical correlation

$$\rho_{n,n-l}^{m,m} = \sqrt{E\{d_{i,n}d_{j,m}\}} \tag{8.42}$$

$$= \begin{cases} =0.125 & \text{for } l = 1 \\ =0.04 & \text{for } l = 2 \quad \text{and} \\ <0.01 & \text{for } l > 2 \end{cases} \tag{8.43}$$

This shows that there is virtually no serial correlation between the wavelet resonance coefficients although they clearly scale.

8.3.2.2 Scalegram of Dow Jones Industrial Average

Figure 8.11 shows the computation of the homogeneous Hurst exponent for 4,096 weekly DJIA data as follows. The time series $X(t)$ in the top panel is analyzed using a Daubechies(5) wavelet basis and MRA tiling. In panel (a) the $\log_2[\text{Var}\{d_{j,n}\}]$ is plotted versus the scale $j = m$. Consequently, the value of the slope of this (almost straight) line is:

$$\begin{aligned} b &= (2H + 1) \\ &= \frac{24 - 10}{11 - 4} \\ &= 2 \end{aligned} \tag{8.44}$$

from which we derive the Hurst exponent $H = 0.5$. This implies that the difference operator exponent $d = H - 0.5 = 0$ for the price increments $\Delta X(t) = \varepsilon(t)$ (cf. Chapter 4). Thus, the weekly DJIA series $X(t)$ follows a pure Random Walk (cf. Chapter 2). Panel (b) shows the average magnitude of the normalized along-scale empirical correlation between the wavelet resonance coefficients:

$$\rho_{n,n-l}^{m,m} = \sqrt{E\{d_{i,n}d_{j,m}\}} \tag{8.45}$$

$$= \begin{cases} =0.08 & \text{for } l = 1 \\ =0.03 & \text{for } l = 2 \quad \text{and} \\ <0.03 & \text{for } l > 2 \end{cases} \tag{8.46}$$

This shows again that there is virtually no serial correlation between the wavelet resonance coefficients.