High-Resolution Noisy Signal and Image Processing

High-Resolution Noisy Signal and Image Processing

By

Igor Zurbenko, Devin Smith,
Amy Potrzeba-Macrina, Barry Loneck,
Edward Valachovic and Mingzeng Sun

Cambridge
Scholars
Publishing

High-Resolution Noisy Signal and Image Processing

By Igor Zurbenko, Devin Smith, Amy Potrzeba-Macrina, Barry Loneck, Edward Valachovic and Mingzeng Sun

This book first published 2021

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-6293-X
ISBN (13): 978-1-5275-6293-6

All of the chapters of this book were written by I. Zurbenko, University at Albany and the following co-authors:

Chapters 1, 2, 3, 4, & 5: **Barry Loneck**, University at Albany

Chapter 6: **Devin Smith**, Rensselaer Polytechnic Institute, Albany, NY

Chapters 7 & 8: **Amy Potrzeba-Macrina**, George Mason University, VA

Chapter 9: **Edward Valachovic**, University at Albany

Chapters 10 & 11: **Mingzeng Sun**, University at Albany

CONTENTS

INTRODUCTION

Kolmogorov's Theory of Turbulence is considered to be among the greatest achievements in the history of science. This theory has made an essential contribution to the safety of airplanes, ships, and submarines. In the 1970s, Kolmogorov embarked on an oceanic expedition in the Pacific to investigate turbulence. During the expedition, he could not prove his results because heavy noise was obscuring the turbulence measurement signal. By removing noise from a signal, he was able to provide proof for aspects of this theory. This idea is the genesis of what is now widely known as a KZ filter. Widespread applications of KZ filters have provided great results by suppressing noise and focusing on a signal. Successful applications of KZ filters include analysis in atmospheric sciences, pollution, climate research, geology, medicine, and public health, as well as numerous other fields. KZ filter applications are able to discern essential reality within raw data by zooming into a signal that is frequently obscured to the point of invisibility.

The great naturalist, Sydney Chapman (1935), was the first to discover the periodic oscillations in atmospheric pressure due to the gravity of the moon. However, he was unable to reproduce those oscillations because of the complexity of reconstructing the signal embedded in the synoptic fluctuations. These atmospheric tidal waves were reconstructed by applying KZ filters. This reconstruction of atmospheric pressure changes, which are due to atmospheric tidal waves, provides an opportunity to predict dangerous moments in extreme weathers events. These results on the reconstruction of tidal waves have been widely cited by specialists in atmospheric science. The same ideas were extended to illustrate how solar activity fluctuations contribute to essential climate changes on earth.

When examining any type of image analysis, problems in terms of the separation of scales significantly multiply. Spatial KZ filters provide a new opportunity to discern previously unseen patterns and images. The transfer function of the KZ filter provides the closest analytical approximation to a rectangular shape. This approximation is accomplished by working asymptotically at the limit of the Heisenberg uncertainty principle. This filter operation uses finite data support in time, thereby allowing for numerous practical applications. The order of the analytic approximation of a rectangular spectral window is equal to the number of iterations within the

KZ filter. The researcher controls the tradeoff between the accuracy of the result and the required computer capacity. As the number of iterations “k” in the filter goes to infinity, the KZ filter converges to the well-known Gaussian filter. While the Gaussian filter is optimal, it is unachievable because of the need for infinite information support. The KZ filter provides a finite-order analytic approximation of the Gaussian window with the same required order of accuracy. The appropriate choice of filtration parameters has a simple physical sense and is easily interpreted in each problem. The basic computer operation of averaging is the central piece of the algorithm, which allows for the construction of extraordinarily fast and effective computer algorithms. This computational efficiency is vital for the interpretation of the massive databases in contemporary spatial data analysis. All of the software algorithms for KZ filtration in time and space are available in R-software and Python.

When data is obscured in time or space due to some anomalies, the parameters of the KZ filter can be adaptively determined by computer logic. The adaptive KZ filter can zoom into the location of a break in data as the filter approaches it. This important feature was developed when investigating data that contains abrupt changes in the equipment used for collecting time series data and fuzzy images in tomography. Another important feature for spatial data analysis is the ability of the KZ filter to work successfully when there is missing data. The KZ filter has provided successful results in an environment when there is over 90% of the data missing.

The construction of filtration is designed in the spectral domain where natural laws occur. This feature allows us to interpret and use the physical sense of natural phenomena. The development of this book was germinated from the ideas of the great scientists, N. Wiener and A. Kolmogorov, as well as H. Cramer, L. Koopmans, E. Parzen, R. Shumway, T.W. Anderson, D. Brillinger, M. Rosenblatt, and P. Bloomfield.

The first chapter is a short survey of the mathematical concepts required to build the theoretical and practical developments in time and space that follow in the later chapters. The second chapter provides a brief list of some time-space applications for developed methods. The third chapter introduces adaptive procedures for the specific features in analyzed data. For example, these adaptive procedures detect sudden breaks in analyzed data where the value of the break is much lower than the quick fluctuations of the data. The fourth chapter introduces a filter capable of zooming to a specific frequency in the signal. We are able to accurately investigate the

details of the specific "voice" at that frequency while, at the same time, removing all other frequencies. The fifth chapter investigates the properties of the DiRienzo-Zurbenko semi-adaptive spectral algorithm and provides a discussion of diagnostic examples derived from the algorithm. The sixth chapter provides both the theoretical foundations and practical examples of spatial developments in time. The seventh chapter describes the detection and reproduction of atmospheric tidal waves caused by the gravity of the moon on earth. These periodic atmospheric pressure fluctuations are above regular synoptic pressure fluctuations. These pressure oscillations can have an impact on big disturbances in atmospheric pressure, such as hurricanes. Examples where the most dangerous moments in hurricanes were linked to tidal waves in the atmosphere include Hurricanes Katrina and Sandy. The eighth chapter investigates the periodicities of sunspot numbers. This investigation shows that these periodicities are equal to some of the specific frequencies of the Sun's planetary system. This investigation also finds that levels of sunspot activity are strictly proportional to the sun radiation levels recorded at satellites. The ninth chapter is an examination of the relationship between sun radiation levels and skin cancer rate in the US, which shows a high correlation in time and space. The tenth chapter provides solutions for spatial boundary detections. One application of these solutions is to detect the boundaries of atmospheric jet streams. These boundaries can be key factors in the prediction of the occurrence of tornados. The eleventh chapter provides a practical guide to KZA (Kolmogorov-Zurbenko algorithms) software in R.

All of the current developments originated from my lecture notes for the Time and Space analysis course, which has been offered in numerous universities through the world. These lecture notes were first edited by Laura Close and they were then later extended to the current text by many of my colleagues.

Igor Zurbenko

CHAPTER 1

BASIC DEFINITIONS AND KEY CONCEPTS

The purpose of this chapter is to provide a brief overview of time series analysis. It presents the basic definitions and key concepts used throughout this book, including stochastic processes, aliasing effect, spectral representation, the covariance function, the Parseval Equality, linear filters, autoregression, and an introduction to cross-spectrum analysis. More details can be found in primary works by Bloomfield (2004), Brillinger (2001), Koopmans (1995), Rosenblatt (2000), Shumway and Stoffer (2014), and Zurbenko (1986).

1.1. Stochastic Processes

A stochastic process, X_t, is defined as

$$X_t \equiv X(t, \omega)$$

where

X_t = the value of a random variable X at time t, for some value of ω;
t = time, either discrete or continuous; and
ω = a random element from a sample space Ω.

If t is fixed, say $t = t_0$, then we have a random variable over the sample space. For example, if Ω is a population of people, $X_{t_0}$are the values of variable X for all of the individuals in the population at the moment $t = t_0$. If ω is fixed, say $\omega = \omega_0$, then we have a sample trajectory over time for individual ω_0.

Time series data consists of a series of real numbers. However, all real numbers can be cast in the imaginary plane. In addition to its standard form (i.e., $a + bi$), a complex number can be placed in either polar form or exponential form, both of which have distinct advantages.

Setting $a = X\,Cos\,\lambda t$ and $b = X\,Sin\,\lambda t$, we have $a + bi = (X\,Cos\,\lambda t) + (X\,Sin\,\lambda t)i$, where X is the amplitude, λ is the frequency, and t is the time. The chief advantage with polar form is that it captures periodicity in time series data because of its use of the sine and cosine functions. By virtue of Euler's formula, the polar form can be recast in exponential form as $Xe^{i\lambda t} = (X\,Cos\,\lambda t) + (X\,Sin\,\lambda t)i$. The chief advantage of the exponential form is its greater facility in manipulating formulas within derivations and proofs. Both polar and exponential forms are used throughout the following discussion, with the particular form dependent on whichever has the greatest advantage in a particular context.

Some stochastic processes are *stationary*. However, the term stationary can be defined in a wide sense or in a strict sense. A stochastic process is stationary in a *wide sense* if the expected value is constant and the covariance of observations at two points in time, t_1 and t_2, depends only on the time difference: that is,

$$C(t_1, t_2) = E((X_{t_1} - EX_{t_1}) \cdot (X_{t_2} - EX_{t_2}))$$

depends only on the difference t_2 - t_1. Given that the expected value of X_t is often 0 (or can be set to 0 by utilizing mean centered data) and given an increment of time, τ, then for real and complex values

$$C(\tau) = E(X_{t+\tau} \cdot X_t) = E(X_{t+\tau} \cdot \overline{X_t})$$

For a complex variable $a + bi$ with complement $\overline{a + bi} = a - bi$, then

$$C(\tau) = \overline{C(-\tau)}$$

A stochastic process is stationary in a *strict sense* if its distribution F is such that

$$F_{X_{t1},\cdots,X_{tn}}(x_1, \cdots, x_n) = F_{X_{t1+\tau},\cdots,X_{tn+\tau}}(x_1, \cdots, x_n)$$

Unlike a stochastic process that is stationary in a wide sense where the covariance of observations at two points in time can depend on the time difference, a stochastic process is stationary in a strict sense when F is independent of any shift in time.

Example 1: Constant Time Series and White Noise

In a constant time series, a random variable, X_t, is constant over time and does not depend on *t*, with $X_t = X$ and $C(\tau) = VarX = \sigma^2$. For example, Figure 1.1.1 depicts three trajectories of one time series with no dependence on time as, for example, the heights of three different adults over time.

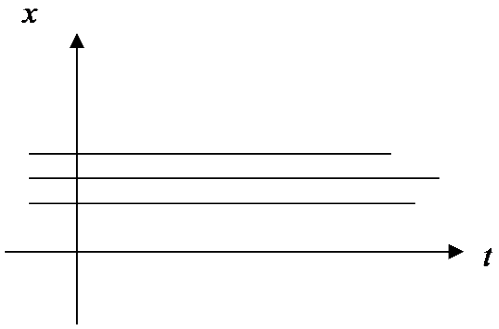

Figure 1.1.1

White noise is defined as a random variable, X_t, for which the values of X_t are uncorrelated over time and the variance is constant. Because the variance is constant, white noise is stationary. If white noise is stationary in a wide sense, then all X_t are simply uncorrelated and have the same variance, $Var(X_t) = \sigma^2$; with $E(X_t) = 0$, we have

$$C(\tau) = Cov(X_{t+\tau}, X_t) = \begin{cases} \sigma^2, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases}$$

However, when white noise is stationary in a strict sense, then all X_t are independent and not only have the same variance, but also have the same distribution. One example of white noise is the number of high energy particles from space that strike a particular geographic location on earth over time. Computer generated random numbers are considered as white noise.

Example 2: Periodic Time Series

We are given a periodic time series with $E(X_t) = 0$ frequency λ, and a phase shift of θ that has a uniform distribution such that $\theta \in [-\pi, \pi]$, with the value of a variable X at time t as

$$X_t = Cos(\lambda t + \theta)$$

This process is depicted in Figure 1.1.2 for several trajectories, each with a different value of θ. If both the starting point, x_0, and the phase shift, θ, are known, then the predicted sample trajectory is complete. Clearly, the period is $T = \frac{2\pi}{\lambda}$ and the harmonic oscillation, with θ having a phase distribution that is random uniformly distributed, does not depend on the shift.

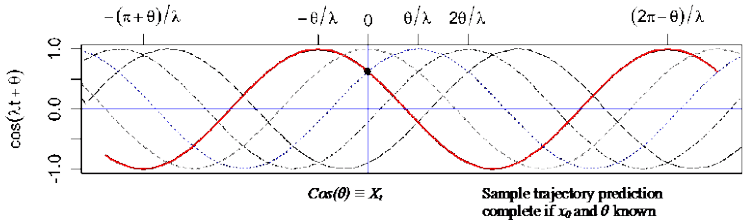

Figure 1.1.2

The covariance between two observations in time, $X_{t+\tau}$ and X_t, is

$$C(\tau) = E(X_{t+\tau} \cdot X_t) = E(Cos(\lambda(t+\tau)+\theta) \cdot Cos(\lambda t + \theta)) = \frac{1}{2} Cos\lambda\tau$$

Proof:

We know that

$$e^{i(\alpha+\beta)} = Cos(\alpha+\beta) + i\, Sin(\alpha+\beta)$$

or, through factoring,

$$e^{i\alpha}e^{i\beta} = (Cos\alpha + i\,Sin\alpha)\cdot(Cos\beta + i\,Sin\beta)$$

$$= (Cos\alpha\cdot Cos\beta - Sin\alpha\cdot Sin\beta) + i(Sin\alpha\cdot Cos\beta + Cos\alpha\cdot Sin\beta)$$

However,

$$Cos(\alpha+\beta) = Cos\alpha\cdot Cos\beta - Sin\alpha\cdot Sin\beta$$

and

$$Cos\alpha\cdot Cos\beta = \frac{1}{2}(Cos(\alpha+\beta) + Cos(\alpha-\beta))$$

Setting $t_1 = t + \tau$ and $t_2 = t$, we can rewrite $C(\tau)$ as

$$E\big(Cos(\lambda t_1+\theta)\cdot Cos(\lambda t_2+\theta)\big) = \frac{1}{2}ECos(\lambda(t_1+t_2)+2\theta) + \frac{1}{2}ECos(\lambda(t_1-t_2))$$

$$= \frac{1}{2}EX_t + \frac{1}{2}Cos\lambda(t_1-t_2)$$

$$= 0 + \frac{1}{2}Cos\lambda(t_1-t_2)$$

Thus, the covariance only depends on $t_1 - t_2$ and because $\tau = t_1 - t_2$, and we have

$$C(\tau) = \frac{1}{2}Cos\lambda\tau$$ Q.E.D.

Example 3: Random Amplitude

We are given a random variable X with $E(X_t) = 0$, $Var(X) = \sigma^2$, and frequency λ, which is defined as

$$X_t = Xe^{i\lambda t}$$

This time series has a period $T = \frac{2\pi}{\lambda}$ and random amplitude X because

$$ae^{ibx} = a\ Cos\, bx + i\ a\ Sin\, bx$$

then

$$Xe^{i\lambda t} = X\ Cos(\lambda t) + i\ X\ Sin(\lambda t)$$

Further,

$$C(\tau) = E(X_{t+\tau} \cdot \overline{X_t}) = E(Xe^{i\lambda(t+\tau)} \cdot Xe^{-i\lambda t}) = E(X^2)e^{i\lambda\tau} = \sigma^2 \cdot e^{i\lambda\tau}$$

Because $C(\tau)$ is determined by the given variance, X_t is stationary in a wide sense.

1.2. Aliasing Effect

Whenever two time series are indistinguishable, there is said to be a possible *aliasing effect*. If time is discrete, with $t = \ldots -2, -1, 0, 1, 2, \ldots$, the time series in Example 3 for λ_2 and λ_1 are indistinguishable if $\lambda_2 = \lambda_1 + 2\,\pi\,k$, where k is some integer. In general, $X_t = Xe^{i\lambda t}$ with amplitude X. If $X_t = Xe^{i\lambda_2 t}$, by substitution we have

$$\begin{aligned} X_t &= Xe^{i\lambda_2 t} \\ &= e^{i(\lambda_1 + 2\pi k)t} \\ &= e^{i\lambda_1 t}\ e^{i2\pi(kt)} \\ &= e^{i\lambda_1 t} \cdot 1 \\ &= e^{i\lambda_1 t} \end{aligned}$$

In general, if $t = k \cdot \Delta t$, with k an integer, then for λ_1, λ_2

$$|\lambda_1 - \lambda_2| = \frac{2\pi}{\Delta t} \cdot m$$

and there is an aliasing effect; thus, the time series is indistinguishable because

$$Xe^{i\lambda_1 t} = Xe^{i\lambda_2 t}$$

for $t = k \cdot \Delta t$. Consequently, observation of a discrete time series does not distinguish two alias frequencies.

Example 4: Aliasing

Assume that we have a time series observed both on a daily basis; that is, $\Delta t = 1\ day$. For this time series, $T = 7\ days$; that is, it has a weekly period. Thus,

$$\lambda_1 = \frac{2\pi}{T_1} = \frac{2\pi}{7}\ radians/day$$

For the second period, $\lambda_2 = \lambda_1 + 2\pi/1$ and

$$\lambda_2 = \left(\frac{2\pi}{7} + \frac{2\pi}{1}\right) radians/day = \frac{2\pi \cdot 8}{7}\ radians/day$$

Thus, we have

$$T_2 = \frac{2\pi}{\lambda_2} = \frac{2\pi}{2\pi \cdot \frac{8}{7}} days = \frac{7}{8} day = 21 hours$$

Consequently, one cannot distinguish the first period of 1 week from the period of 21 hours when observations are done on a daily basis (i.e., one observation done once each day at exactly the same time). As a result, there is an aliasing effect for the first and second periods. This means the 21-hour period may look to the observer as a weekly period. Consequently, $T_1 = 7\ days$ and $T_2 21\ hours$ by daily observation are indistinguishable.

Example 5: Covariance when Aliasing is Present

Assume that time is discrete with $= \ldots -2, -1, 0, 1, 2, ...$, and that we have

$$X_t = \sum_k X_k e^{i\lambda_k t}$$

Further, assume that $E(X_t) = 0$, $E(X_k^2) = \sigma_k^2$, and, for $k \neq p$, $E(X_k\ X_p) = 0$ and $\lambda_k \neq \lambda_p$ for $-\pi \leq \lambda_k \leq \pi$. As a result, the covariance of X_t at two points in time, X_t and $X_{t+\tau}$, is

$$C(\tau) = E(X_{t+\tau} \cdot \overline{X_t}) = E(\sum_{k,p} X_k X_p e^{i\lambda_k (t+\tau) - i\lambda_p t})$$
$$= \sum_{k=p} X_k X_p e^{i\lambda_k \tau} = \sum_k \sigma_k^2 e^{i\lambda_k \tau}$$

1.3. Spectral Representation

Any time series that is stationary in a wide sense can be represented as

$$X_t = \int_{-\pi}^{\pi} e^{i\lambda t} a(\lambda) d\lambda$$

where

X_t = the value of variable X at time t;
$e^{i\lambda t} = Cos(\lambda t) + i\, Sin(\lambda t)$, with frequency λ; and
$a(\lambda)$ = random amplitude of the time series.

with $Cov\big(a(\lambda_1), a(\lambda_2)\big) = 0$, where $\lambda_1 \neq \lambda_2$. Because X_t is a real number, $a(\lambda) = \overline{a(-\lambda)}$.

The *spectral density*, $f(\lambda)$, is

$$f(\lambda) = E\,(|a(\lambda)|^2$$

In other words, the spectral density is the variance of the amplitude, $a(\lambda)$, over a given range of frequencies. Further, for $E(X_t) = 0$, the variance of X_t is

$$Var X_t \overset{for\ EX_t=0}{=} EX_t \overline{X}_t = \int_{-\pi}^{\pi} f(\lambda) d\lambda$$

(see Figure 1.3.1).

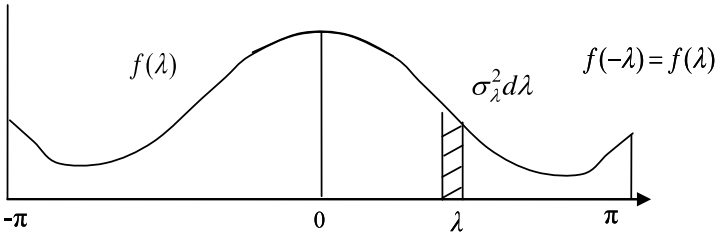

Figure 1.3.1

Figure 1.3.2 shows the spectral density when X_t is almost periodic with frequency λ and period $T_0 = \frac{2\pi}{\lambda_0}$.

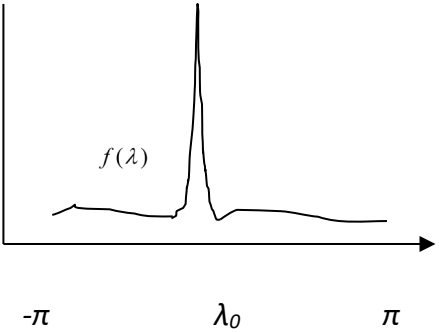

Figure 1.3.2

In a similar fashion, Figure 1.3.3 depicts an approximation to white noise where the energy (i.e., the variance of the amplitude) of all of the frequencies is equal.

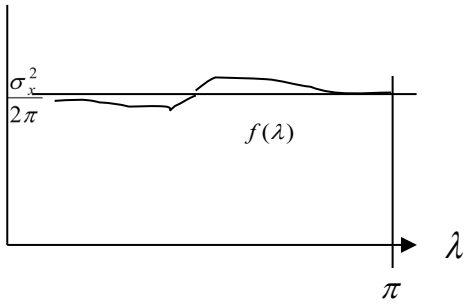

Figure 1.3.3

1.4. Covariance Function

Assuming that $EX_t = 0$ and because $X_t = \int_{-\pi}^{\pi} e^{i\lambda t} a(\lambda) d\lambda$, we have

$$C(\tau) = E(X_{t+\tau} \cdot \overline{X_t}) = E(\int_{-\pi}^{\pi} \int_{-\pi}^{-\pi} e^{i\lambda_1 (t+\tau)} e^{-i\lambda_2 t} a(\lambda_1) \overline{a(\lambda_2)} d\lambda_1 d\lambda_2)$$

$$= \int_{-\pi}^{\pi} \int_{-\pi}^{-\pi} e^{i(\lambda_1 - \lambda_2)t} e^{i\lambda_1 \tau} E(a(\lambda_1) \overline{a(\lambda_2)}) d\lambda_1 d\lambda_2$$

Because $a(\lambda_1)\overline{a(\lambda_2)} = 0$ if $\lambda_1 \neq \lambda_2$, we have

$$C(\tau) = \int_{-\pi}^{\pi} 1 \cdot e^{i\lambda_1 \tau} f(\lambda_1) d\lambda_1 = \int_{-\pi}^{\pi} 1 \cdot e^{i\lambda\tau} f(\lambda) d\lambda$$

or

$$f(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} e^{-i\lambda\tau} C(\tau)$$

To verify, it was noted above that $C(\tau) = \int_{-\pi}^{\pi} e^{i\lambda\tau} f(\lambda) d\lambda$. Thus, by substitution for $f(\lambda)$ with τ_1 and $\tau \neq \tau_1$, we have

$$\int_{-\pi}^{\pi} e^{i\lambda\tau} \left(\frac{1}{2\pi} \sum_{\tau_1=-\infty}^{\infty} e^{-i\lambda\tau_1} C(\tau_1) \right) d\lambda$$

$$= \frac{1}{2\pi}\sum_{\tau_1=-\infty}^{\infty} C(\tau_1)\int_{-\pi}^{\pi} e^{i\lambda\tau - i\lambda\tau_1} d\lambda$$

$$= \frac{1}{2\pi}\sum_{\tau_1=-\infty}^{\infty} C(\tau_1)\int_{-\pi}^{\pi} e^{i\lambda(\tau-\tau_1)} d\lambda$$

and because $\int_{-\pi}^{\pi} e^{i\lambda(\tau-\tau_1)} d\lambda = 2\pi$ if $\tau = \tau_1$, then the result is

$$= \frac{1}{2\pi} C(\tau) \cdot 2\pi = C(\tau)$$ Q.E.D.

1.5. Linear Filters and the Parseval Equality

Given a stationary time series, X_t, with covariance, $C_X(\tau) = Cov(X_{t+\tau}, X_t)$, and spectral density, $f(\lambda)$, if b_k, with $k = \cdots, -1, 0, 1, \ldots$ and $\sum_k b_k^2 < \infty$, then b_k is the impulse characteristics of filter L, with filter L defined as

$$Y_t = \sum_{-\infty}^{\infty} b_k X_{t+k}$$

Such that

$$X_t \overset{L}{\rightarrow} Y_t$$

As shown in Figure 1.5.1, the time series X_{t+k} is plotted concurrent with b_k over time.

$$Y_t = \sum_{-\infty}^{\infty} b_k X_{t+k} \quad X_t \xrightarrow{L} Y_t$$

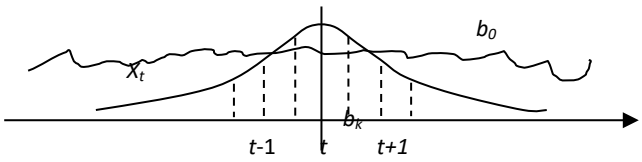

Figure 1.5.1

If $X_t = \int_{-\pi}^{\pi} e^{i\lambda t} a_X(\lambda) d\lambda$ with λ in radians per unit time and the corresponding $\omega = \frac{\lambda}{2\pi}$ in cycles per unit time, then

$$Y_t = \sum_{k=-\infty}^{\infty} b_k X_{t+k} = \sum_k b_k \int_{-\pi}^{\pi} e^{i\lambda(t+k)} a_X(\lambda) d\lambda$$
$$= \int_{-\pi}^{\pi} e^{i\lambda t} (\sum_k b_k e^{i\lambda k}) a_X(\lambda) d\lambda$$

Setting $B(\lambda) = \sum_k b_k e^{i\lambda k}$, we call this the *spectral characteristic*, or the *spectral transfer function*, of filter L and

$$Y_t = \int_{-\pi}^{\pi} e^{i\lambda t} B(\lambda) a_X(\lambda) d\lambda$$

such that Y_t is stationary with $a_Y(\lambda) = B(\lambda) a_X(\lambda)$. Consequently,

$$f_Y(\lambda) = E \mid a_Y(\lambda) \mid^2 = \mid B(\lambda) \mid^2 \; E \mid a_X(\lambda) \mid^2 = \mid B(\lambda) \mid^2 \; f_X(\lambda)$$

and, as a result, $|B(\lambda)|^2$ is the *energy gain* or the *energy transfer function* of the filter L.

Because $B(\lambda) = \sum_k b_k e^{i\lambda k}$, we see that

$$\int_{-\pi}^{\pi} \mid B(\lambda) \mid^2 d\lambda = \int_{-\pi}^{\pi} \sum_k b_k \cdot e^{i\lambda k} \sum_s \bar{b}_s \cdot e^{-i\lambda s} d\lambda$$
$$= \sum_{k,s} b_k \, \bar{b}_s \underbrace{\int_{-\pi}^{\pi} e^{i\lambda(k-s)} d\lambda}_{=0 \; if \; k \neq s}$$
$$= \sum_k \mid b_k \mid^2 \cdot 2\pi \, d\lambda = 2\pi \sum_k b_k^2$$

As a result, the total energy gain is

$$\int_{-\pi}^{\pi} \mid B(\lambda) \mid^2 d\lambda = 2\pi \sum_k b_k^2 < \infty$$

1.6. Examples of Linear Filters

The Parseval Equality lays the foundation for a range of filters which are used to identify cycles and trends in time series datasets. The following are examples of several filters: the uniform transfer filter, the difference filter, the physically realizable filter, the linear filter, the moving average filter, the seasonal filter, the double linear transformation filter, and the finite Fourier transform, which is a linear filter.

The Uniform Transfer Filter

Setting $Y_t = CX_t + A$, there is a uniform energy transfer across all frequencies because

$$|B(\lambda)|^2 = C^2$$

The Difference Filter

If we let $Y_t = X_t - X_{t-1}$, then $b_0 = 1$ and $b_{-1} = -1$ and we obtain

$$B(\lambda) = \sum_k b_k e^{i\lambda k}$$
$$= 1 - e^{-i\lambda}$$

and thus,

$$|B(\lambda)|^2 = (1 - Cos\lambda)^2 + Sin^2\,\lambda = 4\,Sin^2\,\lambda/2$$

This is referred to as a *high pass filter* because it permits signals higher than a particular cutoff frequency to pass, while suppressing frequencies that are moderate to low in a given domain of frequencies.

The Physically Realizable Filter

If $b_k = 0$ for $k > 0$, then $Y_t = \sum_{k=0}^{-\infty} b_k X_{t+k}$ is said to be a *physically realizable filter*. In other words, Y_t only uses information from the present to the past (depicted in Figure 1.6.1).

Physically realizable filter:

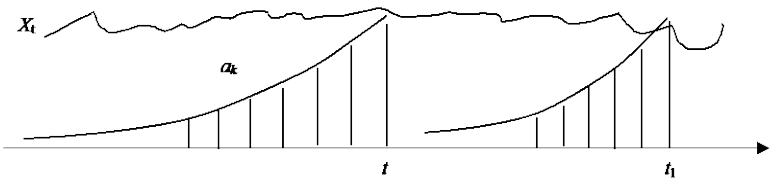

Figure 1.6.1

Symmetry in Transfer Functions

Symmetry. Recalling that $B(\lambda) = \sum_k b_k e^{i\lambda k}$, then $B(-\lambda) = \overline{B(\lambda)}$ for any filter because all b_ks are real coefficients. Thus, we have

$$|B(\lambda)|^2 = B(\lambda) \cdot \overline{B(\lambda)} = B(\lambda) \cdot B(-\lambda)$$

Consequently, $|B(\lambda)|^2$ is always symmetric around $\lambda = 0$, with $|B(\lambda)|^2 = |B(-\lambda)|^2$.

The Linear Filter Operations

Let $B(\lambda)$ be a linear function of b_k. As a result, $Y_t = X_t$ when $B(\lambda) = 1$. By extension, if we can construct a sum of linear filters such that $Y_t = A_1(X_t) + A_2(X_t)$, with $b_k = b_k^{(1)} + b_k^{(2)}$, then

$$B_Y(\lambda) = B_1(\lambda) + B_2(\lambda)$$

Given X_t and linear filter LX_t with b_k and $B(\lambda) = 1$, Figure 1.6.2 depicts the fact that $X_t - LX_t$ is a mirror image of LX_t rotated around the t-axis.

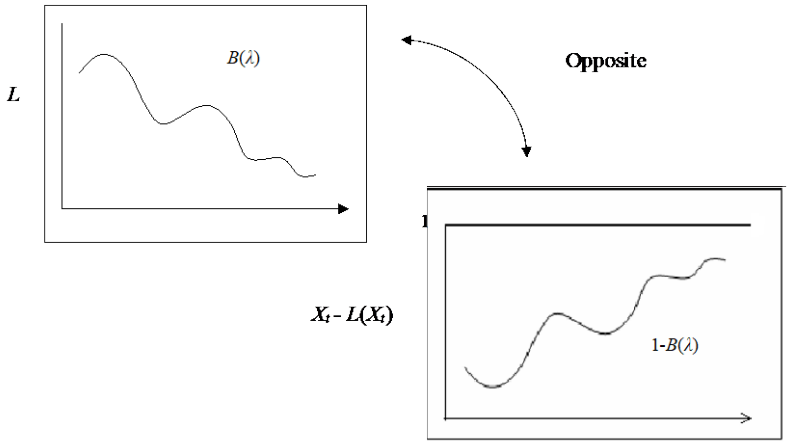

Figure 1.6.2

The Moving Average Filter

The *moving average filter* is defined as

$$Y_t = \frac{1}{m}\sum_{k=0}^{m-1} X_{t-k}$$

with $b_k = \frac{1}{m}$ for $k = 0, \ldots, m-1$ and is depicted in Figure 1.6.3.

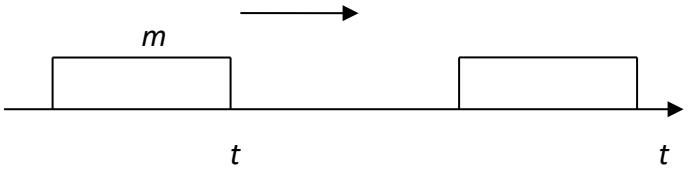

Figure 1.6.3

As a result, we have

$$B(\lambda) = \frac{1}{m}\sum_{k=0}^{m-1} e^{i\lambda k} = \frac{1}{m}\sum_{k=0}^{m-1}\left(e^{i\lambda}\right)^k = \frac{1}{m}\cdot\frac{1-e^{i\lambda m}}{1-e^{i\lambda}}$$

and, thus,

$$|B(\lambda)|^2 = \frac{1}{m^2}\frac{\left|1-e^{i\lambda m}\right|^2}{\left|1-e^{i\lambda}\right|^2} = \frac{1}{m^2}\frac{(1-Cos\,\lambda m)^2+Sin^2\lambda m}{(1-Cos\,\lambda)^2+Sin^2\lambda}$$

$$= \frac{1}{m^2}\frac{2-2\,Cos\,\lambda m}{2-2\,Cos\,\lambda} = \frac{1}{m^2}\frac{Sin^2\frac{\lambda m}{2}}{Sin^2\frac{\lambda}{2}}$$

or

$$|B(\lambda)|^2 = \frac{1}{m^2}\frac{\left(\frac{\lambda m}{2}\right)^2}{\left(\frac{\lambda}{2}\right)^2} \approx 1$$

The moving average filter is a low pass filter because it permits signals lower than a particular cutoff frequency to pass while suppressing frequencies that are moderate to high. A graph of the resulting spectral density energy gain $|B(\lambda)|^2$, is depicted in Figure 1.6.4.

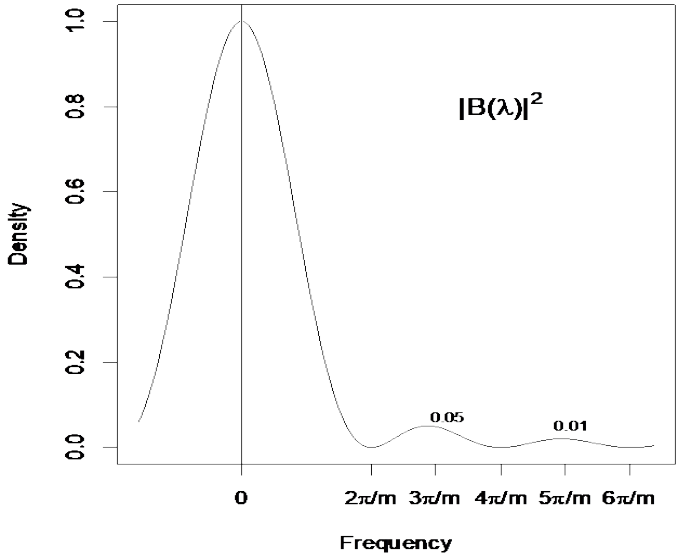

Figure 1.6.4

The filtering of a time series X_t, which includes a number of moderate and high frequencies, to form Y_t using a moving averages filter is presented in Figure 1.6.5. It can be seen that the moderate and high frequencies present in X_t have been attenuated in Y_t.

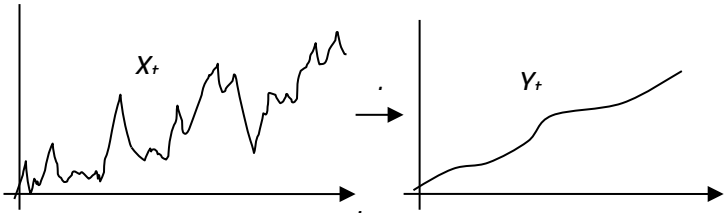

Figure 1.6.5

The Seasonal Filter

A *seasonal filter* is a variation of the moving averages filter and is defined as

$$Y_t = \sum_{k=0}^{m-1} X_{t-k}$$

with period p. A depiction of b_k is shown in Figure 1.6.6.

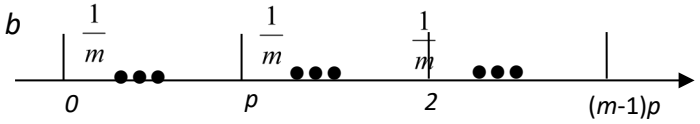

Figure 1.6.6

Because of the seasonal period, p, we have

$$B(\lambda) = \frac{1}{m}\sum_{k=0}^{m-1} e^{i\lambda pk}$$

and, thus,

$$|B(\lambda)|^2 = \frac{1}{m^2}\frac{Sin^2\frac{\lambda mp}{2}}{Sin^2\frac{\lambda p}{2}}$$

A graph of the resulting spectral density is shown in Figure 1.6.7 and one can note the repetition through each interval of 2π. For each frequency, $\frac{2\pi}{p}$, $\frac{4\pi}{p}$, …, there is a corresponding *pass period* given as

$$T = \frac{2\pi}{2\pi/p} = p,\ \frac{2\pi}{2\cdot\frac{2\pi}{p}} = \frac{p}{2} = \frac{T}{2}, \ldots$$