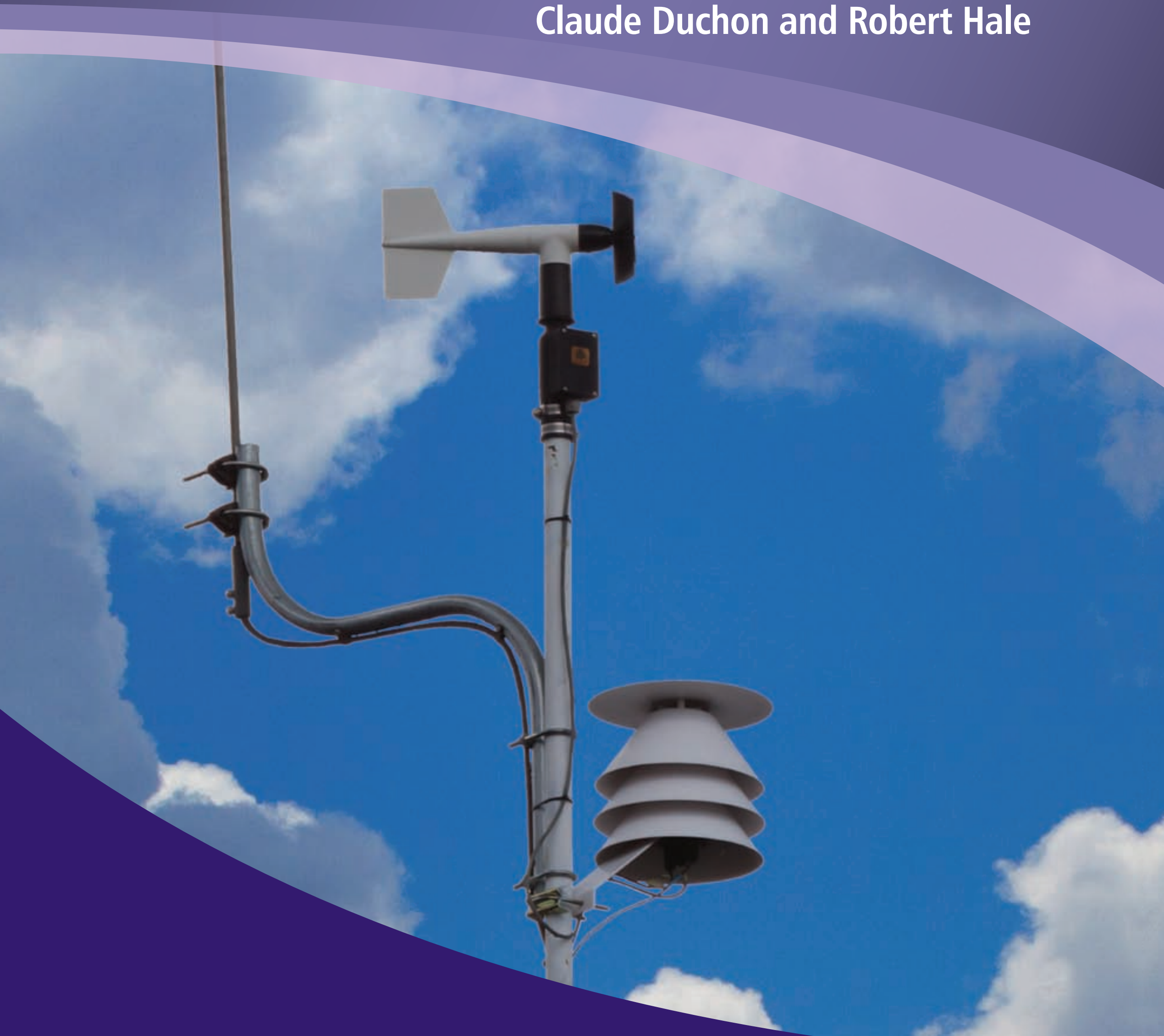
WILEY-BLACKWELL

RMetS
Royal Meteorological Society

Time Series Analysis in Meteorology and Climatology

An Introduction

Claude Duchon and Robert Hale



Advancing Weather and Climate Science

Time Series Analysis in Meteorology and Climatology

Companion Website

This book has a companion website www.wiley.com/go/duchon/timeseriesanalysis with datasets for the problems in each chapter.

Time Series Analysis in Meteorology and Climatology

An Introduction

Claude Duchon

University of Oklahoma, USA

Robert Hale

Colorado State University, USA

WILEY-BLACKWELL

A John Wiley & Sons, Ltd., Publication

Wiley-Blackwell is an imprint of John Wiley & Sons, formed by the merger of Wiley's global Scientific, Technical and Medical business with Blackwell Publishing.

Registered office: John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial offices: 9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK
111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Library of Congress Cataloging-in-Publication Data

Duchon, Claude E.
Time series analysis in meteorology and climatology : an introduction / Claude Duchon and Robert Hale.
p. cm. – (Advancing weather and climate science)
Includes bibliographical references and index.
ISBN 978-0-470-97199-4 (hardback)
1. Meteorology–Statistical methods. 2. Time-series analysis. I. Hale, Robert, 1940- II. Title.
QC874.5.D83 2011
551.501'51955–dc23

2011030428

A catalogue record for this book is available from the British Library.

This book is published in the following electronic formats: ePDF 9781119953098; Wiley Online Library 9781119953104; ePub 9781119960980; Mobi 9781119960997

Set in 10.5/13pt Minion by Thomson Digital, Noida, India

First Impression 2012

Contents

Series foreword

Advances in weather and climate

Meteorology is a rapidly moving science. New developments in weather forecasting, climate science and observing techniques are happening all the time, as shown by the wealth of papers published in the various meteorological journals. Often these developments take many years to make it into academic textbooks, by which time the science itself has moved on. At the same time, the underpinning principles of atmospheric science are well understood but could be brought up to date in the light of the ever increasing volume of new and exciting observations and the underlying patterns of climate change that may affect so many aspects of weather and the climate system.

In this series, the Royal Meteorological Society, in conjunction with Wiley–Blackwell, is aiming to bring together both the underpinning principles and new developments in the science into a unified set of books suitable for undergraduate and postgraduate study as well as being a useful resource for the professional meteorologist or Earth system scientist. New developments in weather and climate sciences will be described together with a comprehensive survey of the underpinning principles, thoroughly updated for the 21st century. The series will build into a comprehensive teaching resource for the growing number of courses in weather and climate science at the undergraduate and postgraduate levels.

Series Editors

Peter Inness
University of Reading, UK

William Beasley
University of Oklahoma, USA

Preface

Time series analysis is widely used in meteorological and climatological studies because the vast majority of observations of atmospheric and land surface variables are ordered in time (or space). Over the years we have found a continuing interest by both students and researchers in our profession (and those allied to it) in understanding basic methods for analyzing observations ordered in time or space and evaluating the results. The purpose of this book is to respond to this interest. We've done this by deriving and interpreting various equations that are useful in explaining the structure of data and then, using computer programs, applying them to meteorological data sets. Overall, the material we cover serves as an introduction in the application of statistics to the analysis of univariate time series. The topics discussed should be relevant to anyone in any science where events are observed in time and/or space. To demonstrate a procedure, we use scalar atmospheric variables, for example, air temperature. Anyone who completes the five chapters, including working the problems at the end of each chapter, will have acquired sufficient understanding of time series terminology and methodology to confidently deal with more advanced spectrum analysis, for example, that found in radar and atmospheric turbulence measurements, analysis, and theory.

Chapter 1 deals with Fourier analysis and is divided into five sections. In the first three sections, mathematical formulas for representing a time series by Fourier sine and cosine coefficients are developed and their inherent symmetry emphasized. These formulas are applied to three data sets, two of which are actual observations. The three sections provide the background necessary to apply Fourier analysis to a time series, and one of the end-of-chapter problems invites the reader to write a computer program designed to accomplish this.

In the fourth section of Chapter 1 we investigate statistical properties of the Fourier spectrum. These statistical properties arise because time series from the physical world are usually nondeterministic, that is, no two data sets are alike. We explore the concept of a random variable, a realization, a population, stationarity, expectation, and a probability density function. The goal is to understand how random data produce a distribution of variances at each harmonic frequency and the statistical properties of this distribution. Armed with this information, the last part of this section involves testing the hypothesis that a particular data set, as viewed through the Fourier spectrum, is a sample from a population of white noise, that is, random numbers.

The fifth section of Chapter 1 is an examination of various topics relevant to time series analysis. We discuss aliasing, spectrum folding, and spectrum windows, phenomena that are a direct consequence of digital sampling, and show examples of each. In addition, we develop the Fourier transform, the mathematical formula that in one step converts a time series into its Fourier components in the frequency domain.

Chapter 1 is the longest of the five chapters because it encompasses both theory and application of Fourier analysis, relevant statistical concepts, and the foundation of methods of time series analysis developed in the remaining chapters.

The subject of Chapter 2 is linear systems. This chapter is the study of the relationships between two time series, an input series and an output series, and the associated input and output spectra. What links the two time series is a physical system, as in the case of measurement of some physical variable (for example, a thermometer to measure temperature), or a mathematical system, as in the case of filtering an observed time series to remove unwanted noise in the data.

Fundamental to Chapter 2 is the convolution integral. Whether a system is physical or mathematical, the convolution integral provides the mathematical connection between the input and output series, and its Fourier transform provides the connection between the input and output spectra.

Most variables of interest in the physical sciences are continuous in time (or space). Nevertheless, we practically always analyze digital time series. We investigate the relationship between analog and digital time series using a generalized function called the Dirac delta function. Through its application we can explain how the structure of an output time series that has passed through a linear system is altered relative to the input time series in terms of modified Fourier coefficients and phase angles. Two examples are discussed, a first order linear system and an integrator, both of which have practical use in meteorology and climatology, and the physical sciences in general.

Chapter 3 is principally about nonrecursive data filtering; that is, a filtered output time series is related only to the input time series – there is no feedback (as in recursive filtering). Time series that are to be filtered are viewed as data that already have been collected as opposed to real time filtering.

The primary objective of this chapter is to design and apply a two-parameter filter called the Lanczos filter. The two design parameters are the number of weights and the frequency that separates the Fourier spectrum into harmonic variances that remain unchanged and those that are suppressed. This filter provides its designer much more control of the filtering process than simple one-parameter filters, for example, the running mean. The theory of Lanczos filtering is developed, examples of its use are shown, and a computer program is provided so that the reader can apply the procedure to a data set.

One of the goals of a physical scientist is to understand the morphology of natural events. An obvious step that must be taken is to obtain samples in time and/or space of variables that characterize the physical properties of an event over its lifetime. The fact that an event has a lifetime implies that it evolves in time and/or space, a consequence of which is that successive observations of its properties are related. This is called autocorrelation, the title of Chapter 4. To realize the importance of autocorrelation in analyzing time series, we compare the formula for calculating the variance of the mean of a random variable with autocorrelation to that without autocorrelation. The latter formula is the form seen in typical undergraduate statistics texts while the former formula takes into account the degree of dependence in the time series.

In Chapter 4 we are interested in finding the best formula for estimating the mean, variance, autocovariance function, and autocorrelation function of a population of time series based, typically, on a single observed time series taken from that population. We examine populations of independent as well as autocorrelated data. Among the five chapters, this one is the most statistically oriented.

The lagged-product method discussed in Chapter 5 is an alternative to Fourier analysis. Quite often, Fourier analysis of geophysical data yields noisy-looking spectra. When this occurs, it is common to smooth a spectrum to make it more visually interpretable. In the lagged-product method, a smoothed variance spectrum can be obtained directly from the Fourier transform of the product of the autocovariance function with another function that alters its shape. The degree of smoothing is controlled entirely by the latter function. The term lagged-product is used because the autocovariance function comprises time-lagged (or spatially-lagged) products and it is the autocovariance function that is being transformed.

This book was written for students and scientists who have a background in calculus and statistics, and familiarity with complex variables. Prior in-depth study of complex variables is not required.

The authors wish to thank the many students who have provided valuable comments and corrections over the years the material was used as lecture notes. Chapters 2, 4, and 5 were inspired by the book *Spectral Analysis and its Applications* (1968) by G.M. Jenkins and D.G. Watts, a classic volume in time series analysis.

Claude Duchon and Rob Hale
22 May 2011

1
Fourier analysis

It is often the case in the physical sciences, and sometimes the social sciences as well, that measurements of a particular variable are collected over a period of time. The collected values form a data set, or *time series*, that may be quite lengthy or otherwise difficult to interpret in its raw form. We then may turn to various types of statistical analyses to aid our identification of important attributes of the time series and their underlying physical origins. Basic statistics such as the mean, median, or total variance of the data set help us succinctly portray the characteristics of the data set as a whole, and, potentially, compare it to other similar data sets.

Further insight regarding the time series, however, can be gained through the use of *Fourier*, *harmonic*, or *periodogram analysis* – three names used to describe a single methodology. The primary aim of such an analysis is to determine how the total variance of the time series is distributed as a function of frequency, expressed either as ordinary frequency in cycles per unit of time, for example, cycles per second, or angular frequency in radians per unit of time. This allows us to quantify, in a way that the basic statistics named above cannot, any *periodic* components present in the data. For example, outside air temperature typically rises and falls with some regularity over the course of a day, a periodic component governed by the rising and setting of the sun as the earth rotates about its axis. Such a periodic component is readily apparent and quantifiable after applying Fourier analysis, but is not described well by the mean, median, or total variance of the data.

In the first two sections of Chapter 1, we will learn some essential terminology of Fourier analysis and the fundamentals of performing Fourier analysis and its inverse, Fourier synthesis. Example data sets and their analyses are presented in Section 1.3 to further aid in understanding the methodology.

As with other types of statistical analyses, statistical significance plays an important role in Fourier analysis. That is, after performing a Fourier analysis, what if we

Time Series Analysis in Meteorology and Climatology: An Introduction, First Edition. Claude Duchon and Robert Hale.

find that the variance at one frequency is noticeably larger than at other frequencies? Is this the result of an underlying physical phenomenon that has a periodic nature? Or, is the larger variance simply statistical chance, owing to the random nature of the process? To answer these questions, in Section 1.4 we examine how to ascribe confidence intervals to the results of our Fourier analysis.

In Section 1.5, we take a more detailed look at particular issues that may be encountered when using Fourier analyses. Although not generally requisite to performing a Fourier analysis, the concepts covered are often critical to correct interpretation of the results, and in some cases may increase the efficacy of an analysis. An understanding of these topics will allow an investigator to pursue Fourier analysis with a high degree of confidence.

1.1 Overview and terminology

1.1.1 Obtaining the Fourier amplitude coefficients

The goal of Fourier analysis is to decompose a data sequence into harmonics (sinusoidal waveforms) such that, when added together, they reproduce the time series. What makes sinusoidal waveforms an appropriate representation of the data is their orthogonality property, their ability to successfully model waves in the atmosphere, oceans, and earth, as well as phenomena resulting from solar forcing, and the fact that the harmonic amplitudes are independent of time origin and time scale (Bloomfield, 1976, p. 7).

Harmonic frequencies are gauged with respect to the *fundamental period*, the shortest record length for which the time series is not repeated. In most practical cases, this is the entire length of the available record, since the record typically does not contain repeated sequences of identical data. The harmonic frequencies include harmonic 1, which corresponds to one cycle over the fundamental period, and higher harmonics that are integer multiples of one cycle. Thus each harmonic is always an integer number of cycles over the length of the fundamental period.

To establish a sense of Fourier analysis, consider a simple example. The heavy line in Figure 1.1 connects the average monthly temperatures at Oklahoma City over the three-year period 2007–2009. By looking at the heavy line only, it is quite evident that there is a strong annual cycle in temperature. It is equally clear that one sinusoid will not exactly fit all the data, so other harmonics are required. The fundamental period, or period of the first harmonic, is the length of the record, three years. The third harmonic has a period one-third the length of the fundamental period, and consequently represents the annual cycle. The thin line in Figure 1.1 shows the third harmonic after it has been added to the mean of all 36 months, that is, the 0-th harmonic. As expected, the third harmonic provides a close fit to the observed time series.

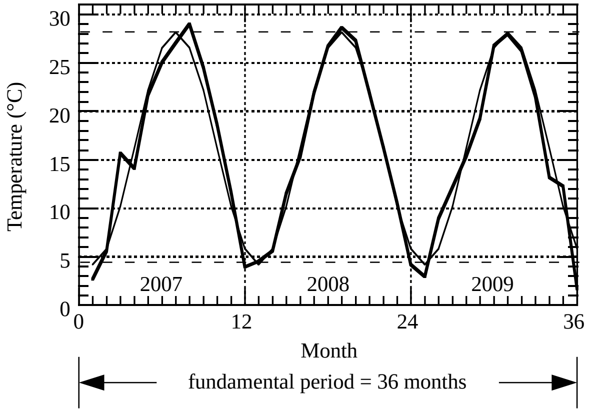


Figure 1.1 Mean monthly temperatures at Oklahoma City 2007–2009 (heavy line), and harmonic 3 (light line) of the Fourier decomposition.

1.1.2 Obtaining the periodogram

The computation of variance arises in elementary statistics as a defined measure of the variability in a data set. When the computation of variance is applied to a time series, it is similarly defined. Now, though, the variance in the data set can be decomposed into individual variances, each one related to the amplitude of a harmonic. Just as adding the sinusoids from all harmonics reproduces the original time series, adding all harmonic variances yields the total variance in the time series. How the decomposition is achieved and how variance is related to harmonic amplitude are discussed in Section 1.2.

A *periodogram* is a plot of the variance associated with each harmonic (usually excluding the 0-th) versus harmonic number and shows the contribution by each harmonic to the total variance in the time series. Henceforth, the term periodogram will be used to refer to the calculation of variance at the harmonic frequencies. The term *Fourier line variance spectrum* is synonymous with periodogram, while the generic term *spectrum* generally means the distribution of some quantity with frequency.

The variance at each harmonic frequency is given by the square of its amplitude divided by two, except at the last harmonic. Figure 1.2 shows the periodogram (truncated to the first 10 harmonics) of the data in Figure 1.1 where we see that harmonic 3 dominates the variability in the data. The small variances at harmonics 6 (period = 6 months) and harmonic 9 (period = 4 months) are easily observed in Figure 1.2, but, in fact, there are nonzero variances at all 18 possible harmonics (excluding the 0-th) and their sum equals the total variance of 75.23 $°C^2$ in the 2007–2009 Oklahoma City mean monthly temperature time series.

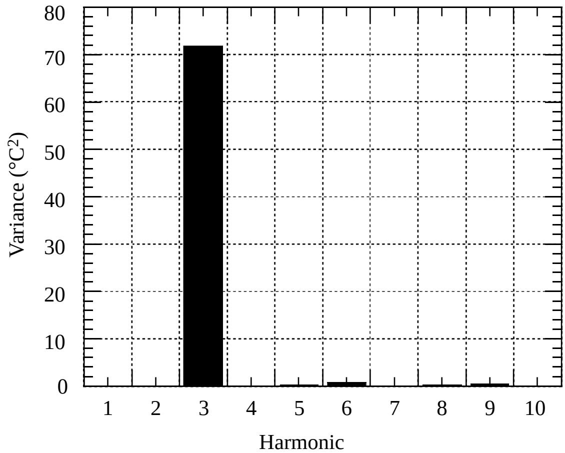


Figure 1.2 Variance at each harmonic through 10 for the data in Figure 1.1.

The periodogram in Figure 1.2 was computed using the computer program given in Appendix 1.A. This program, written in Fortran 77, performs a 'fast' Fourier analysis of any data set with an even number of data and has been used throughout this chapter to compute the periodograms we discuss.

1.1.3 Classification of time series

We can classify time series of data into four distinct types of records. The type of record determines the mathematical procedure to be applied to the data to obtain its spectrum.

The 36 values of temperature x_n, in Figure 1.1, connected by straight-line segments for ease in visualization, constitute a *finite digital* record. Digital time series arise in two ways (Box and Jenkins, 1970, p. 23): sampling an analog time series, for example, measuring continuously changing air temperature each hour on the hour; or accumulating or averaging a variable over a period of time, for example, the previous record of monthly mean temperatures at Oklahoma City. With respect to the latter case, if N is the number of months of data and Δt the time interval between successive values, the record length in Figure 1.1 is $N\Delta t = 36$ months. In this case, as well as with all finite digital records, all data points can be exactly fitted with a finite number of harmonics. This is in contrast to a *finite analog* record of length T, such as a pen trace on an analog strip chart, for example, a seismograph, for which an infinite number of harmonics may be required to fit the signal.

Figure 1.3 is an example of a finite analog record. Sampling the time series at intervals of Δt yields the finite digital record shown in Figure 1.4. The sample values again have been connected by straight-line segments to better visualize the variations

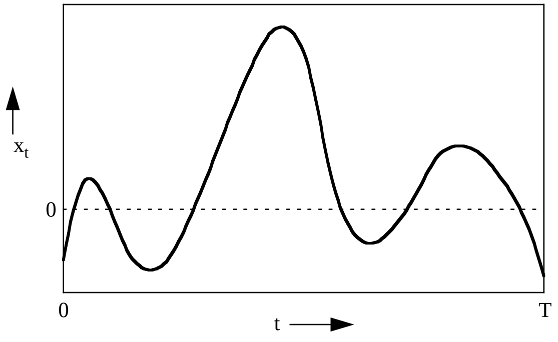


Figure 1.3 An example of a finite analog data record.

in x_n. The sampling interval, Δt, associated with each datum can be shown on a time series plot to the left or right of, or centered on, each datum – it is a matter of choice. In Figure 1.4, Δt is to the right of each datum. One might think that there should be a fifteenth sample point at the very end of the curve in Figure 1.3. However, because of the association of each sampled value with one Δt, the length of the digital record would be one sample interval longer than the analog record. Conceptually, the fifteenth sample point is the first value of a continuing, but unavailable, analog record.

The concept of an *infinite analog* record is often used in theoretical work. An example would be the trace in Figure 1.3 extended indefinitely in both directions of time. For this case a continuum of harmonics is required to fit the signal, thereby resulting in a *variance density spectrum*. Note, however, that a variance density spectrum can be created also with a finite digital record. How this comes about is

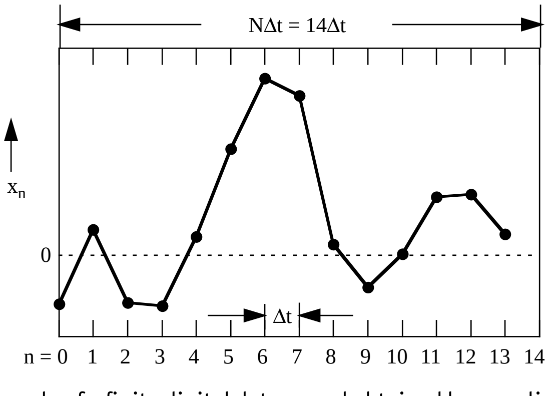


Figure 1.4 An example of a finite digital data record obtained by sampling the finite analog record in Figure 1.3. There are $N = 14$ data.

discussed in Chapter 5. An *infinite digital* record would be obtained by sampling the infinite analog record at intervals of Δt. We will use infinite analog and digital records in Section 3.1.4 (Chapter 3) to determine the effects on the mean value of a time series after it is filtered.

By far the type of record most commonly observed and analyzed in science and technology is the finite digital record. With a few exceptions, this is the type of data record we will deal with in the remainder of Chapter 1, and for which the formulas for computing a periodogram are presented.

1.2 Analysis and synthesis

1.2.1 Formulas

If one of the data sets collected in your research is a time series of atmospheric pressure, Fourier "analysis" can be used to derive its periodogram and to examine which harmonics dominate the series. Conversely, once the analysis has been done, the original time series of pressure can be reconstructed purely from knowledge of the harmonic amplitudes. Thus Fourier "synthesis" is the inverse process of analysis. Note that the title of this chapter employs the more generic meaning of analysis and includes both the analysis and synthesis terms just described.

The formulas in Table 1.1 are those needed to perform analysis and synthesis. The equations under Fourier Analysis are used to calculate the Fourier coefficients or harmonic amplitudes. The equations under Fourier Synthesis express the time series x_n as the sum of products of cosines and sines with amplitudes A_m and B_m, respectively, or, alternatively, the sum of products of cosines only with amplitudes R_m and phase angles θ_m. Notice that the expressions are slightly different depending on whether the time series has an even or an odd number of data. The synthesis equations are equivalent to the forms introduced by Shuster around 1900 (Robinson, 1982).

The arguments of the cosine and sine terms associated with the A_m and B_m coefficients are of the form

$$\frac{2\pi mn\Delta t}{N\Delta t}$$

where m is harmonic number and $n\Delta t$ a point in time along the time axis of total length $N\Delta t$. Thus, $2\pi m$ is the number of radians in the m-th harmonic over the total length of the time series. The product of $2\pi m$ and the ratio $n\Delta t/N\Delta t$ provide location along the sinusoid in radians. Because the time increments (Δt) cancel, they are not shown in Table 1.1. In Fourier synthesis, the summation is over all harmonics at a given location $n\Delta t$, while in Fourier analysis the summation is over all data locations for a given harmonic m.

Table 1.1 Formulas used in Fourier synthesis and analysis for an even or odd number of data.

Fourier Analysis

$$A_0 = \frac{1}{N}\sum_{n=0}^{N-1} x_n \qquad B_0 = 0$$

$$A_m = \frac{2}{N}\sum_{n=0}^{N-1} x_n \cos\frac{2\pi mn}{N} \qquad B_m = \frac{2}{N}\sum_{n=0}^{N-1} x_n \sin\frac{2\pi mn}{N}$$

$$m = \left[1, \tfrac{N}{2}-1\right] (\text{N even}); \quad m = \left[1, \tfrac{N-1}{2}\right] (\text{N odd})$$

$$A_{N/2} = \frac{1}{N}\sum_{n=0}^{N-1} x_n \cos(\pi n) \qquad B_{N/2} = 0 \quad (\text{N even})$$

$$R_m = \sqrt{A_m^2 + B_m^2} \qquad \theta_m = \tan^{-1}\left(\frac{B_m}{A_m}\right)$$

Fourier Synthesis

$$x_n = \sum_{m=0}^{N/2}\left(A_m \cos\frac{2\pi mn}{N} + B_m \sin\frac{2\pi mn}{N}\right) = \sum_{m=0}^{N/2} R_m \cos\left(\frac{2\pi mn}{N} - \theta_m\right), \quad n = [0, N-1] \ (\text{N even})$$

$$x_n = \sum_{m=0}^{\frac{N-1}{2}}\left(A_m \cos\frac{2\pi mn}{N} + B_m \sin\frac{2\pi mn}{N}\right) = \sum_{m=0}^{\frac{N-1}{2}} R_m \cos\left(\frac{2\pi mn}{N} - \theta_m\right), \quad n = [0, N-1] \ (\text{N odd})$$

Variance at Harmonic m

$$S_m^2 = \frac{A_m^2 + B_m^2}{2} \qquad m = \left[1, \tfrac{N}{2}-1\right] \ (\text{N even}); \quad m = \left[1, \tfrac{N-1}{2}\right] \ (\text{N odd})$$

$$S_{N/2}^2 = A_{N/2}^2 \quad (\text{N even})$$

Total Variance

$$S^2 = \sum_{m=1}^{N/2} S_m^2 \quad (\text{N even}) \qquad S^2 = \sum_{m=1}^{\frac{N-1}{2}} S_m^2 \quad (\text{N odd})$$

The variance at each harmonic for even and odd data lengths is given in Table 1.1 under the heading Variance at Harmonic m. Note that the only exception to the general formula for harmonic variance occurs at $m = N/2$ when N is even. The cosine coefficient at N/2 is squared but not divided by two (the sine coefficient is zero). The formulas for the total variance S^2 under the heading Total Variance yield the same variance estimates as the formula

$$S^2 = \frac{1}{N}\sum_{n=0}^{N-1} (x_n - \overline{x})^2 \tag{1.1}$$

for computing total variance directly from the data, in which $\overline{x}$ is the series mean. The two formulas in Table 1.1 are nearly the same, the only difference being that the expression for the upper limit of each summation depends on whether N is even or odd.

1.2.2 Fourier coefficients

The method for obtaining the Fourier coefficients is based on the *orthogonality* of cosine and sine functions at harmonic frequencies, where orthogonality means that the sum of the products of two functions over some interval equals zero. The method entails multiplying both sides of a Fourier synthesis equation by one of the cosine or sine harmonic terms, summing over all n, and solving for the coefficient associated with the harmonic term.

For example, consider multiplying both sides of the first Fourier synthesis equation in Table 1.1 (using the A_m, B_m form) by $\cos\frac{2\pi kn}{N}$ and summing over all n. The second summation on the right-hand side will have the form and result

$$\sum_{n=0}^{N-1} \sin\frac{2\pi mn}{N}\cos\frac{2\pi kn}{N} = 0 \tag{1.2}$$

where m and k are integers. That this sum is zero can be shown with two examples as well as mathematically. The sine and cosine terms for $m = k = 1$ are shown in Figure 1.5 and for $m = 1$ and $k = 2$ in Figure 1.6. The algebraic signs of the sum of cross products within each quadrant are shown at the base of each figure. Because of symmetry, the absolute magnitude of each sum is the same for each quadrant in

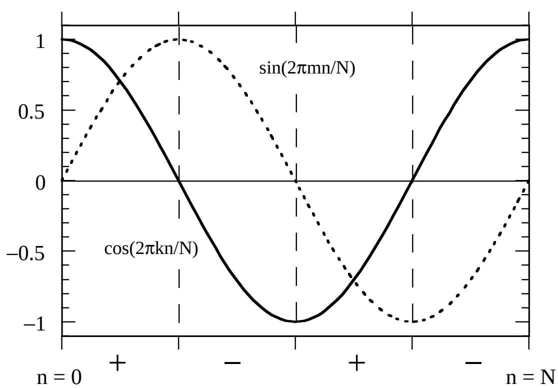


Figure 1.5 Signs of sums of cross products of cosine and sine terms for $m = k = 1$.

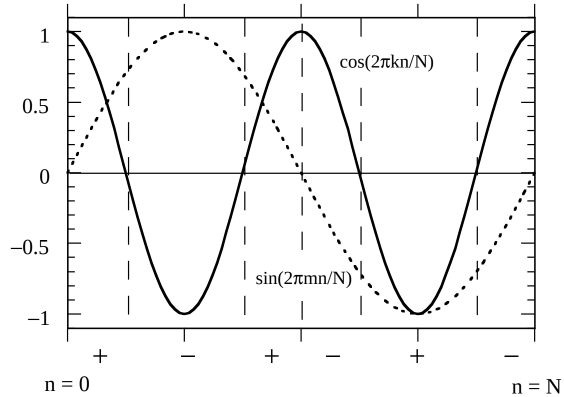


Figure 1.6 Signs of sums of cross products of cosine and sine terms for $m = 1$ and $k = 2$.

Figure 1.5 and similarly for Figure 1.6. Thus the waveforms are orthogonal because the sum of their cross products is zero over the interval 0 to N in each illustration.

It can be surmised from these figures that the sum of the cross products is zero over the fundamental period for any combination of the m and k integers. But how could this be shown mathematically? Firstly, we put the sine and cosine terms in complex exponential form, and then expand the summation above using Euler's formula to obtain

$$\begin{aligned}
&\sum_{n=0}^{N-1} \sin(2\pi mn/N)\cos(2\pi kn/N) \\
&\quad = \sum_{n=0}^{N-1} \frac{1}{2i}\left(e^{i2\pi mn/N} - e^{-i2\pi mn/N}\right)\frac{1}{2}\left(e^{i2\pi kn/N} + e^{-2\pi kn/N}\right) \\
&\quad = \frac{1}{4i}\sum_{n=0}^{N-1}\left(e^{i2\pi(m+k)n/N} + e^{i2\pi(m-k)n/N} - e^{-i2\pi(m-k)n/N} - e^{-i2\pi(m+k)n/N}\right).
\end{aligned} \tag{1.3}$$

A procedure is developed in Appendix 1.B for finding the sum of complex exponentials. The final two formulas, Equations 1.B.3 and 1.B.4, are very useful for quickly finding the sums of sines and cosines over any range of their arguments. An example of using the first formula follows.

Consider just the first summation on the right-hand side in Equation 1.3. Let

$$Q = \sum_{n=0}^{N-1} e^{i2\pi(m+k)n/N}. \tag{1.4}$$

Using Equation 1.B.3, Q becomes

$$
\begin{aligned}
Q &= \frac{1 - e^{i2\pi(m+k)}}{1 - e^{i2\pi(m+k)/N}} \\
&= \frac{1 - \cos[2\pi(m+k)] - i\sin[2\pi(m+k)]}{1 - \cos[2\pi(m+k)/N] - i\sin[2\pi(m+k)/N]} \\
&= 0, \quad m + k \neq 0,\ N. \qquad (1.5)
\end{aligned}
$$

The numerator is zero for all integer values of m and k while the denominator is nonzero except when $(m+k) = 0$ or N, in which cases the denominator is 0 and Equation 1.5 is indeterminate. To evaluate Equation 1.5 for these cases we can apply l'Hopital's rule. The result of taking the first derivative with respect to $(m+k)$ in both the numerator and denominator yields a determinate form with value N. That is

$$
\begin{aligned}
Q' &= \frac{2\pi\sin[2\pi(m+k)] - i\,2\pi\cos[2\pi(m+k)]}{(2\pi/N)\sin[2\pi(m+k)/N] - i\,(2\pi/N)\cos[2\pi(m+k)/N]} \\
&= N, \quad m + k = 0,\ N. \qquad (1.6)
\end{aligned}
$$

The same result also can be obtained by substituting 0 or N for (m+k) in Equation 1.4. We observe that the first and fourth summations in Equation 1.3 cancel each other for these values.

We can apply the above procedure to the second term in Equation 1.3. The summation will be zero again, except when $(m - k)$ is 0 or N. Employing l'Hopital's rule yields a determinate form with value N for these cases, similar to Equation 1.6. And again, the same results can be obtained from Equation 1.4. Accordingly, when $(m - k) = 0$ or N, the second and third summations in Equation 1.3 cancel. Thus Equation 1.2 is valid for any integer k or m. This includes the possibility that $(k + m)$ is an integer multiple of N.

Now that we have shown that the summed sine–cosine cross product terms akin to Equation 1.2 must be zero, let us consider the sums of sine–sine and cosine–cosine products resulting from multiplying the first Fourier synthesis equation by $\cos\frac{2\pi kn}{N}$ and summing over all n. Following the procedure in Appendix 1.B we find that

$$
\sum_{n=0}^{N-1} \sin\frac{2\pi mn}{N}\sin\frac{2\pi kn}{N} = \begin{cases} 0, & k \neq m \\ \frac{N}{2}, & k = m \neq 0, \frac{N}{2}\,(N\,\text{even});\ k = m \neq 0\,(N\,\text{odd}) \\ 0, & k = m = 0, \frac{N}{2}\,(N\,\text{even});\ k = m = 0\,(N\,\text{odd}) \end{cases} \qquad (1.7)
$$

and

$$\sum_{n=0}^{N-1} \cos\frac{2\pi mn}{N} \cos\frac{2\pi kn}{N} = \begin{cases} 0, & k \neq m \\ \frac{N}{2}, & k = m \neq 0, \frac{N}{2}\,(N\,\text{even}); k = m \neq 0\,(N\,\text{odd}) \\ N, & k = m = 0, \frac{N}{2}\,(N\,\text{even}); k = m = 0\,(N\,\text{odd}). \end{cases} \tag{1.8}$$

Thus multiplying the synthesis equation for N even by the k-th sine harmonic term and summing yields

$$\sum_{n=0}^{N-1} x_n \sin\frac{2\pi kn}{N} = \sum_{m=0}^{N/2} \left(A_m \sum_{n=0}^{N-1} \sin\frac{2\pi kn}{N} \cos\frac{2\pi mn}{N} + B_m \sum_{n=0}^{N-1} \sin\frac{2\pi kn}{N} \sin\frac{2\pi mn}{N} \right) \tag{1.9}$$

which reduces to

$$\sum_{n=0}^{N-1} x_n \sin\frac{2\pi kn}{N} = B_k N/2, \quad k = \left[1, \frac{N}{2} - 1\right] \tag{1.10}$$

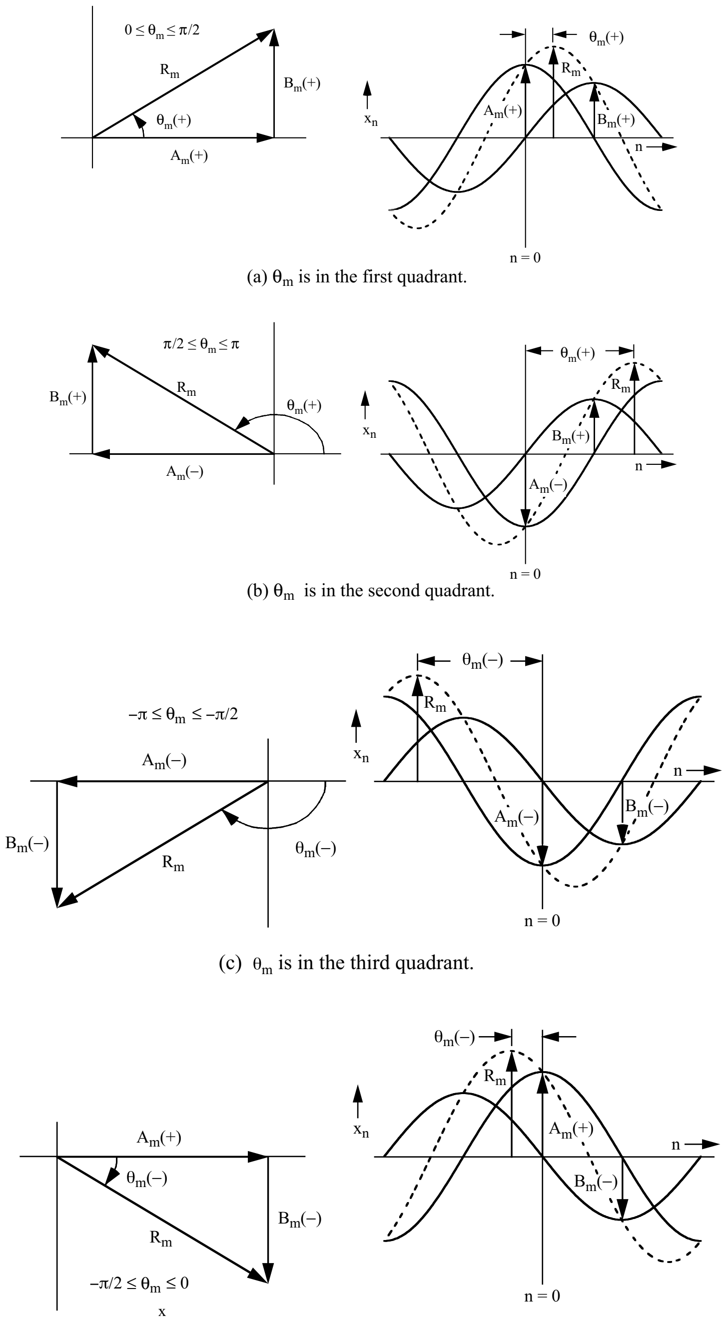
so that

$$B_k = (2/N) \sum_{n=0}^{N-1} x_n \sin\frac{2\pi kn}{N}, \quad k = \left[1, \frac{N}{2} - 1\right]. \tag{1.11}$$

Observe that the sine coefficients for $k = 0$, $N/2$ (N even) are always zero.

The Fourier cosine coefficients, A_k, are obtained in a similar manner, but A_0 and $A_{N/2}$ are, in general, nonzero. As is evident from Table 1.1, A_0 is the mean of the time series. For N odd, an expression similar to Equation 1.9 is used to obtain the Fourier coefficients, the only difference being that the range of harmonics extends from 0 to $(N - 1)/2$. Table 1.1 shows the resulting formulas for all Fourier coefficients.

The coefficients A_m and B_m represent the amplitudes of the cosine and sine components, respectively. As shown in the left-hand panel of Figure 1.7a, the cosine coefficient is always along the horizontal axis (positive to the right), and the sine coefficient is always normal to the cosine coefficient (positive upward). In the right-hand panel we see how the cosine and sine vector lengths determine the associated cosine and sine waveforms (ignore the dashed line for the moment). Figures 1.7b–d show various possibilities of waveform relationships depending on the sign of A_m and the sign of B_m. More discussion of Figure 1.7 is given in Section 1.2.4.



(a) θ_m is in the first quadrant.

(b) θ_m is in the second quadrant.

(c) θ_m is in the third quadrant.

(d) θ_m is in the fourth quadrant.

Figure 1.7 (a)–(d) The magnitude and sign of each Fourier coefficient determines the quadrant in which the phase angle lies. Geometric vector lengths in the left hand panels are twice the lengths of the Fourier coefficients in the right hand panels.