FREQUENCY-DOMAIN ANALYSIS OF MEMORYLESS NONLINEARITIES HAVING LARGE-SIGNAL, ALMOST PERIODIC EXCITATIONS

by

Donald Michael Keller

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Electrical Engineering

APPROVED:

W. A. Davis, Chairman

F. W. Stephenson

I. M. Besieris

L. G. Kraige

R. L. Moose

May, 1988

Blacksburg, Virginia

FREQUENCY-DOMAIN ANALYSIS OF MEMORYLESS NONLINEARITIES HAVING LARGE-SIGNAL, ALMOST PERIODIC EXCITATIONS

by

Donald Michael Keller

Committee Chairman: William A. Davis

Electrical Engineering

(ABSTRACT)

Numerical frequency–domain techniques are widely used for the a.c. steady–state analysis of nonlinear electric circuits. Such techniques require that one compute the Fourier series for the response of each nonlinear circuit element, given a known excitation.

Current approaches to this computation encounter difficulty when the response is almost periodic (that is, when the frequencies in its Fourier series are not all harmonically related), especially when the nonlinear characteristic is abrupt and the Fourier series for the response contains many significant terms.

This dissertation develops an alternative approach that is theoretically sound and computationally efficient, for the important special case of a memoryless nonlinearity described by a continuous, bounded function. To begin the development, basic properties of almost periodic functions are presented. It is proven that the response of a memoryless nonlinearity is almost periodic whenever the excitation is. Next, the concept of a basis for a set of frequencies is introduced. The frequency content of the response is investigated, and it is proven that the frequencies in the response have the same basis as those in the excitation.

The Fourier series for an almost periodic function is discussed, and its coefficients are expressed as mean values taken over an infinite interval. Results are given for the summability of the series.

Starting with a theorem from Diophantine Approximation, it is proven that the normalized (Hertzian) phases corresponding to a set of M basis frequencies have their fractional parts uniformly distributed in an M-dimensional unit cube. This property of uniform phase distribution is then used to convert the single-dimensional integral for the Fourier series coefficients into a multiple integral over the unit cube, with the dimension of the integral equal to the number of basis frequencies in the Fourier series.

A multi-dimensional extension of the Discrete Fourier Transform is used to evaluate the multiple integral, and expressions for aliasing are derived. It is shown that the multiple integral formulation compares favorably with existing approaches, and several numerical examples are presented to illustrate this formulation's capabilities.

ACKNOWLEDGMENTS

I believe that one of the signs of a good friendship between two individuals is that each person, over time, acquires some of the attitudes and opinions of the other. Such has been the case with my friend and advisor, Dr. William A. Davis, and me. We had much in common to begin with, but through countless lunches, classes, meetings, and trips, we have found our outlooks growing more and more similar. Which one of us has had the greater influence on the other, we have not yet decided. I can say with certainty, however, that he has been a fine example to me, both personally and professionally, and I feel very fortunate to have known him for the past eight years. As I complete my studies, I wish to thank him for his help, understanding, and enthusiasm for our profession, which have played a major part in whatever success I have achieved.

I also wish to thank the other members of my advisory committee—Drs. Bill Stephenson, Ioannis M. Besieris, L. Glenn Kraige, and Richard L. Moose—for their willing service. Each of them, in his own way, has also had a positive influence on me, and I consider my association with them to be among my many fine experiences as a student at this institution.

CONTENTS

CHAPTER 1
Introduction

Electrical engineers have been intrigued and perplexed by nonlinear electric circuits[1] since the beginning of the profession. Historically, the analysis and design of these circuits have been limited to a few canonic forms, often handled through the use of clever mathematics together with graphs and tables.

With the advent of the high–speed digital computer, nonlinear circuit problems increasingly have been subjected to numerical methods that are designed to solve for the response due to a particular excitation. While some of the insight provided by the old methods is lost, the numerical techniques generally provide better accuracy together with the ability to handle more complicated networks.

The nature of the nonlinear problem makes it unlikely that there will ever be a single best method for all nonlinear networks. As with linear circuits, there are both time–domain and frequency–domain formulations,[2] with some techniques well–suited to d.c. analysis, others to transient analysis, and still others to a.c. steady–state analysis.[3]

---

1 Nonlinear electric circuits are those whose voltages, currents, electric charges, and magnetic fluxes are related by nonlinear differential equations.

2 "Time domain," as usual, means that all circuit quantities such as voltages and currents are functions of time. "Frequency domain" means that all circuit quantities are functions of frequency; that is, they are the Fourier Transforms of the corresponding time quantities. We may also include phasor analysis under the heading of frequency–domain techniques, since it is merely a shortcut used when all sources and responses are pure complex exponentials.

3 D.C. analysis concerns circuits having only d.c. sources and responses. Transient analysis concerns the response of a circuit to a "recently–applied" excitation; the transient response by definition decays to zero for times infinitely far removed from the onset of the excitation. A.C. steady–state analysis concerns the response of a circuit to a sum of pure sinusoids applied at some "long–ago" time; by definition, the steady–state response is that part of the response that remains for times infinitely far removed from the onset of the excitation.

1.1 Frequency–Domain Methods for A.C. Steady–State Analysis

A.C. steady–state techniques are commonly used in the analysis and design of a variety of nonlinear electronic circuits, including audio– and radio–frequency amplifiers (both narrowband and wideband), sinusoidal oscillators, mixers, and detectors. Many computer–aided analysis packages use time–domain formulations, chiefly because these formulations are more general and more easily programmed; however, there are significant advantages to using frequency–domain formulations, both in hand computations and in computer–aided analysis.

Among these advantages, we note two. First, the frequency domain is inherently suited to steady–state problems, whether linear or nonlinear. In a steady–state problem, we typically know the input signals (excitations) through their Fourier series, and wish to compute the output signals (responses) as Fourier series. A frequency–domain solution works directly with the Fourier series components of all voltages and currents in the circuit, while a time–domain solution to the network requires the additional labor of computing the time–domain excitations from their Fourier series, then computing the Fourier series of the time–domain responses.

The second advantage comes in analyzing a network that contains many linear energy–storage elements but only a few nonlinear elements. Time–domain methods such as state equations are notoriously inefficient in this case because they typically generate a first–order differential equation for each energy–storage element. The system of all such differential equations must be solved to determine the response of the network, and as the order of the system increases, computational effort naturally increases. In contrast, frequency–domain methods analyze linear portions

of the network using their transfer–function and driving–point immittance properties, resulting in simple algebraic equations. Further, if it is necessary to solve a system of simultaneous equations, the order of the system depends on the number of nonlinear elements, not on the number of linear inductors and capacitors.

Apart from improving computational efficiency in computer analysis, the two aforementioned advantages allow us to analyze and design many practical nonlinear networks using approximate frequency–domain techniques. Some of the most successful work in this area was done by Clarke and Hess (1971), whose book *Communication Circuits: Analysis and Design* has become something of a classic among radio engineers. By treating a few canonic nonlinearities (piecewise–linear, square–law, and exponential) embedded in narrowband linear filters, the authors provide approximate analyses for narrowband amplifiers, oscillators, mixers, and detectors.

1.2 Frequency–Domain Analysis of a Memoryless Nonlinearity

Frequency–domain analysis of a nonlinear network brings with it the two advantages mentioned above; however, it also brings with it an important disadvantage: analysis of the nonlinear circuit elements themselves is usually more difficult in the frequency domain than in the time domain. For the case of a.c. steady–state analysis in the frequency domain, all circuit quantities (voltages, currents, electric charges, and magnetic fluxes) are described by their Fourier series. Analysis of a nonlinear element then means that we must compute the Fourier series of its "response" quantity, given the Fourier series of its "excitation" quantity. For example, a nonlinear resistor may have its terminal current as the excitation and its

terminal voltage as the response. (In many cases, including some nonlinear resistors, the roles of excitation and response could be interchanged.)

For an arbitrary nonlinear element, this frequency–domain analysis problem is quite difficult indeed. Henceforth, we will consider only an important special case, that of the so–called memoryless nonlinearity. Such a nonlinearity has a time–domain description in which the value of the response at any given instant of time depends only on the value of the excitation at the same instant. By appropriate selection of the excitation and response quantities, even such energy–storage elements as nonlinear capacitors can be described by memoryless nonlinearities. (For example, a nonlinear capacitor might have $q = \phi(v)$, where q is the instantaneous charge displacement, v is the terminal voltage, and ϕ is a single–valued nonlinear function. The terminal current can then be found by taking the time derivative of q, which corresponds to a simple multiplication by the transform variable in the frequency domain.) To simplify our work, we will further restrict our attention to memoryless nonlinearities that are described by continuous, bounded, real–valued functions.

The problem of analyzing an arbitrary memoryless nonlinearity in the frequency domain has been attacked in several ways. Most of the methods currently in use are computationally oriented, since the complexity of the problem generally precludes a symbolic analysis. We will now look at three of these methods. For the remainder of this chapter, we will let $x(t)$ be the excitation signal, $y(t)$ be the response signal, and $y = \phi(x)$ be the memoryless nonlinearity.

Discrete Fourier Transform

Since the Fourier series coefficients of $x(t)$ are assumed known, it is easy to calculate the value of $x(t)$, hence the value of $y(t)$, for any given value of t. Thus we can generate a set of N samples of $y(t)$, say $\{y(n\Delta t),\ n=0,1,...,N-1\}$, and by applying the Discrete Fourier Transform (DFT) to these samples, we can estimate the Fourier Transform of $y(t)$. Because $y(t)$ consists of a sum of sinusoids having distinct frequencies, its Fourier spectrum consists of impulse functions located at the frequencies of these sinusoids. The weight of each impulse function is equal, within a multiplicative constant, to the corresponding Fourier series coefficient.

This approach has been popular with researchers for many years, in part because software packages for DFT analysis are readily available, and in part because the DFT is conceptually straightforward and handles arbitrary nonlinearities. There are serious difficulties with this method, however, that make it impractical for many applications. In the first place, there is the well-known aliasing problem, which causes the corruption of each series coefficient due to the presence of higher-order coefficients (those corresponding to higher frequencies). If the Fourier series of $y(t)$ is finite, we can eliminate aliasing by sampling at a sufficiently rapid rate (according to the Sampling Theorem, at a rate exceeding twice the highest frequency in $y(t)$). This rate may lead to extremely large numbers of sample points, however. If the Fourier series of $y(t)$ is infinite, we can make the effect of aliasing small, but can never eliminate it.

A second, even greater difficulty comes when $y(t)$ is not purely periodic, that is, when the frequencies in its Fourier series are not all harmonically related. (We call such a function $y(t)$ an "almost periodic function.") In this case, a second

source of inaccuracy in a given coefficient arises due to "leakage" from other nearby coefficients. Even if we can eliminate aliasing, we can never completely eliminate this leakage. It can be made small by using a large number of sample points, but again, we may require extremely large numbers of points to obtain acceptable accuracy.

Perhaps the greatest objection to the DFT is that its computational effort is not always directly related to the complexity of the problem. For example, suppose that $y(t)$ is a sum of three sinusoids having frequencies of ω_0, $\omega_0 + \omega_1$, and $\omega_0 - \omega_1$. If $\omega_1 = \omega_0/2$, the problem is simply that of finding the coefficients of a three-term Fourier series. On the other hand, if $\omega_0 >> \omega_1$, the three frequencies will be spaced very closely relative to ω_0. Whether or not the three frequencies are harmonically related, the DFT will now require a much larger number of points than in the first case, even though the apparent complexity of the problem has not increased.

Despite the aforementioned shortcomings, the DFT continues to find advocates in the current literature [for example, Nakhla and Vlach (1976), Hicks and Khan (1982), Camacho-Peñalosa (1983), Gilmore (1986), and Gayral et al. (1987)]. Gilmore (1986) suggests a procedure for eliminating aliasing while minimizing the number of sample points, but his method requires the solution of a system of simultaneous equations.

Orthogonal Series

A second approach to the memoryless nonlinearity problem is to expand the nonlinear function $\phi(x)$ in a series of orthogonal basis functions $\Phi_n(x)$, so that $\phi(x) = \sum_{n=0}^{\infty} c_n \Phi_n(x)$. The coefficients c_n are computed in the usual way by making

use of the orthogonality properties of the Φ_n.

The set of basis functions is chosen such that the Fourier series of $\Phi_n[x(t)]$ is easily computed for each n. A common choice is to make $\{\Phi_n(x)\}$ a set of orthogonal polynomials, such as Legendre polynomials [see Beckmann (1973), for example]. With this choice, we can then rearrange the series of orthogonal functions into a power series $\phi(x) = \sum_{n=0}^{\infty} a_n x^n$, and each term $[x(t)]^n$ is easily expanded (at least in concept) using trigonometric identities for products of sines and cosines. Of course, for practical computation, we must take the upper limit in the summation to be finite.

This power series expansion of $\phi(x)$ is the one suggested most often in the literature [by Weiner and Spina (1980), for example], although Antonov and Ponkratov (1974) among others have studied the general case. The major shortcoming of such an expansion is that $\phi(x)$ may not be well represented by a low-order polynomial, and as the number of significant terms in the series expansion grows, so too does the computational effort. (Weiner and Spina present algorithms for organizing the computation.) A classic example of a troublesome $\phi(x)$ is the exponential nonlinearity used to describe a diode junction.

Another difficulty with the power series expansion is that the Fourier series coefficients of $y(t)$ cannot be found in closed form unless the series representation of $\phi(x)$ happens to terminate for finite n. Even if the series for $\phi(x)$ does terminate, a particular Fourier series coefficient can receive contributions from a number of terms $[x(t)]^n$, all of which must be expanded to find just this one coefficient. Despite these shortcomings, power series expansion of $\phi(x)$ continues to be a widely-used approach because of its simplicity and theoretical soundness.

In an attempt to handle more abrupt nonlinearities, some researchers have used sets of orthogonal real exponentials; often, they use the simplest case of a single real exponential to approximate a diode junction [Stoodley (1964), Gretsch (1966), Gardiner (1968), Bhatkar and Atre (1970), Surana and Gardiner (1971), Clarke and Hess (1971)]. The Fourier series coefficients arising from these formulations typically take the form of infinite series involving Bessel functions or modified Bessel functions.

Least Squares

Chua and Ushida (1981) propose a least squares formulation for dealing with the memoryless nonlinearity. They begin by approximating $y(t)$ with a function $\tilde{y}(t)$ that has a finite–length Fourier series. Next, they generate a set of N samples of $\tilde{y}(t)$ for N different values of t; they also generate a set of N samples of the actual output $y(t)$ at the same values of t, using the known $x(t)$ and $\phi(x)$. Finally, they choose the coefficients of the assumed series for $\tilde{y}(t)$ such that the mean–square error between the two sets of samples is minimized.

This process of choosing the coefficients can be carried out efficiently using matrix inversion and multiplication. For a given number of samples and a given set of assumed output frequencies, the matrix inversion need be carried out only once; thereafter, the coefficients corresponding to a particular set of output samples are found through a single matrix multiplication.

The true values of the Fourier series coefficients are found by minimizing the mean–square error between $y(t)$ and its Fourier series representation over all t; thus the least squares formulation is a discretized approximation to minimizing the

mean-square error. The approximation becomes exact under either of two conditions: 1) $y(t)$ is represented exactly by the assumed series, or 2) the number of samples approaches infinity.

When neither of the above two conditions is met, the error in the calculated Fourier series coefficients can be substantial. Success in using the least squares method requires assuming enough terms in the series for $y(t)$. Should results show that not enough terms were used, the analysis must be started over assuming more terms; iterative improvement is not possible.

1.3 Proposed Solution

In searching for an improved method to handle memoryless nonlinearities in the frequency domain, the author set forth the following requirements on such a method:

1) It should work well with arbitrary nonlinear characteristics, including gross nonlinearities such as exponentials; ideally, it should allow piecewise or even pointwise descriptions of the characteristic in addition to analytical ones.

2) It should be optimized for handling an input consisting of a sum of sinusoids whose frequencies may or may not be harmonically related. It should remain computationally efficient even if the input or output signals have very closely-spaced or widely-spaced frequencies.

3) It should be oriented toward numerical analysis with a digital computer, and it should be computationally efficient relative to existing methods. It should also allow for iterative improvement, achieving arbitrary accuracy for a sufficient number of terms or iterations.

The solution to be presented in the following chapters meets all of these requirements. It formulates an expression for the Fourier series coefficients of $y(t)$ as a multiple integral over finite limits, with the dimension of the integral equal to the number of independent frequencies in $x(t)$. (The meaning of "independent" will be made more precise later.) This formulation is quite amenable to numerical analysis using, for example, a multi-dimensional extension of the Discrete Fourier Transform.

This multiple integral formulation is formally the same as the expression for the coefficients of a multiple Fourier series. In fact, a few researchers have attempted to attack special cases of the memoryless nonlinearity problem using multiple Fourier series techniques. Among them are Bennett (1933), who analyzed linear and square-law half-wave rectifiers for the case of two input signals, Surana and Gardiner (1971), who analyzed a half-wave rectifier with its "on" characteristic expanded in a power series, and Bhatkar and Atre (1970), who analyzed an exponential nonlinearity driven by a finite number of sinusoids. Each of these analyses begins by expressing the Fourier series coefficients as multiple integrals, but reduces the expressions using transcendental functions such as Bessel, modified Bessel, and elliptic functions.

All of the aforementioned analyses suffer from a common deficiency: the authors fail to justify the multiple Fourier series as a valid representation of $y(t)$, and even assuming that this representation is valid, they fail to state clearly under what conditions it is valid.

Our upcoming work will not suffer from this deficiency. We will rigorously derive the multiple integral formulation from fundamental principles of Fourier series analysis for an almost periodic function. There will be no assumptions about

the existence of a multiple Fourier series, and such concepts will not be needed. Apart from its mathematical correctness, this approach will yield useful insights into the problem, insights that lead both to a deeper understanding of the underlying theory and to practicable computational techniques.

1.4 Organization of this Dissertation

Before deriving the multiple integral formulation, we will devote Chapters 2 through 4 to the necessary background material. In Chapter 2, we will study almost periodic functions, both in the classical sense and in one generalized sense. We will summarize results concerning their time–domain and frequency–domain properties, and we will prove that the output $y(t)$ of a memoryless nonlinearity is almost periodic whenever the input $x(t)$ is.

In Chapter 3, we will introduce the concept of a basis for a set of frequencies, and we will prove that the frequencies in $x(t)$ and $y(t)$ have the same basis. We will also derive an expression for the number of frequencies appearing in $y(t)$, and we will study the distribution of these frequencies on the real axis.

In Chapter 4, we will construct the Fourier series of an almost periodic function. We will express the Fourier series coefficients as mean values taken over an infinite interval, and we will present results from the literature concerning the existence and summability of the series.

We will carry out the development of the multiple integral formulation in Chapters 5 and 6. First, in Chapter 5, we will prove the central theorem of our work, namely, that the fractional parts of the vector $\overline{\mu}t$, $0 \leq t < \infty$, are uniformly distributed in the M–dimensional unit cube whenever $\mu_1, \mu_2, \ldots, \mu_M$ are independent

over the integers. Using this result in Chapter 6, we will convert the single integral form of the Fourier series coefficients into a multiple integral over the unit cube. We will find that the dimension of this integral is equal to the number of basis frequencies in $y(t)$.

In Chapter 7, we will approximate the multiple integral using a multi–dimensional extension of the Discrete Fourier Transform (DFT), an extension that we will call a Multiple Discrete Fourier Transform (MDFT). We will derive an expression for the error involved in using the MDFT approximation, and we will show that this error has the form of multi–dimensional aliasing. We will also write the MDFT in matrix notation, and compare it to the least–squares method for accuracy and computational efficiency.

In Chapter 8, we will give five simple examples of nonlinear element analysis to illustrate the potential of the MDFT for good accuracy and efficiency. Finally, in Chapter 9, we will present our conclusions, and discuss directions for further research.

CHAPTER 2
Almost Periodic Functions

In the early part of this century, the Danish mathematician Harald A. Bohr developed the theory of a class of everywhere–bounded and everywhere–continuous functions that he called "almost periodic functions." Bohr characterized these functions in two ways: first, by their structural properties (that is, their properties as functions of an independent variable such as time), and second, by their vibration properties (that is, their properties as sums of sinusoids). Almost periodic functions are an extension of the class of purely periodic functions, an extension that replaces harmonically related sinusoids with sinusoids that merely have distinct frequencies. It is precisely the class of almost periodic functions in which we are interested in our analysis of the memoryless nonlinearity.

We will now consider briefly some fundamental results in the theory of almost periodic functions. We will then look at one extension of almost periodic functions (due to Stepanoff), in which the requirement of continuity everywhere is relaxed, and replaced with a requirement of continuity almost everywhere. The former class will be called uniformly almost periodic functions (denoted $\{a.p.\}$), and the latter class Stepanoff almost periodic functions (denoted $\{S.a.p.\}$).

2.1 Uniformly Almost Periodic Functions

Throughout this section, we will require the concept of the uniform distance between two functions:

Definition 2.1.1: Let $f(t)$ and $g(t)$ be continuous, bounded functions that are defined for $-\infty < t < \infty$. Then we define the uniform distance between $f(t)$ and $g(t)$ as

$$D[f(t),g(t)] \equiv \underset{-\infty < t < \infty}{\text{upper bound}} |f(t) - g(t)|. \tag{2.1.1}$$

We further define

$$D[f(t)] \equiv D[f(t),0]. \; \Delta \tag{2.1.2}$$

This distance clearly obeys the Triangle Inequality; that is, for continuous, bounded functions $f(t)$, $g(t)$, and $h(t)$ defined on $-\infty < t < \infty$,

$$D[f(t),h(t)] \leq D[f(t),g(t)] + D[g(t),h(t)]. \tag{2.1.3}$$

Structural Properties of $\{a.p.\}$

We will now examine the structural properties of *a.p.* functions. To do so, we will need two additional definitions:

Definition 2.1.2: Let $f(t)$ be a function continuous for $-\infty < t < \infty$, and let ϵ be a positive number. If a real number τ satisfies

$$D[f(t + \tau),f(t)] < \epsilon, \tag{2.1.4}$$

then we call τ a uniform translation number of $f(t)$ corresponding to ϵ. We will denote the set of all such τ by $E_\epsilon[f(t)]$. Δ

Definition 2.1.3: Let E be a set of real numbers. If there exists a positive number l such that every interval of length l on the real number line contains at least one member of E, then we say that the set E is relatively dense. Δ

We may now use these results to give the definition of uniformly almost periodic functions:

Definition 2.1.4: If the set $E_\epsilon[f(t)]$ is relatively dense for all $\epsilon > 0$, then the function $f(t)$ is a uniformly almost periodic function. Δ

This definition says that a uniformly almost periodic function repeats itself, within an error of ϵ, after an interval of $t = \tau$; further, this error can be made arbitrarily small by choosing an appropriate τ. If $f(t)$ is not purely periodic, τ must become arbitrarily large as ϵ becomes sufficiently small (otherwise, there would be a finite point of accumulation for τ, and this point would be a period of $f(t)$, contrary to assumption).

Vibration Properties of $\{a.p.\}$

We will now characterize the vibration properties of $\{a.p.\}$. We will need the following three definitions:

Definition 2.1.5: Consider a sequence of functions $\{f_n(t),\ n=1,2,...\}$, and a function $f(t)$. We call $f(t)$ a uniform limit function of the sequence if

$$D[f(t),f_n(t)] \to 0 \quad \text{as } n \to \infty. \ \Delta \tag{2.1.5}$$

Definition 2.1.6: Consider a set of functions U. Then $f(t)$ is called a uniform limit function of the set U if it is the uniform limit function of some sequence of functions $\{f_n(t)\}$ contained in U. Δ

Definition 2.1.7: A set of functions U augmented by the set of all possible uniform limit functions of U is called the uniform closure of U, and will be denoted by $C(U)$. Δ

We can now state the main result of this section, due to Bohr:

Theorem 2.1.1: Let $s(t)$ denote a finite trigonometric polynomial, that is,

$$s(t) = \sum_{k=-n}^{n} a_k e^{j\lambda_k t} \tag{2.1.6}$$

where the exponents λ_k are any real numbers, and the coefficients a_k any complex

numbers. We call the class of all such polynomials A. Then the uniform closure of A is identical to the class of uniformly almost periodic functions. That is,

$$C(A) = \{a.p.\}. \tag{2.1.7}$$

Proof: See Bohr (1951), pp. 80–88. ∆

Theorem 2.1.1 tells us that uniformly almost periodic functions can be approximated to an arbitrary degree of accuracy, in the sense of the uniform distance, by finite trigonometric polynomials. (We say that *a.p.* functions can be "uniformly approximated" by such polynomials.) Hence, the theorem imparts vibration properties to the class $\{a.p.\}$ in addition to the structural properties studied earlier. In Chapter 4, we will complete the investigation of these vibration properties by learning how to construct the trigonometric polynomials of Theorem 2.1.1.

A Function of an *a.p.* Function

To conclude this section, we prove a result that we will need for our later work.

Theorem 2.1.2: Let $\phi(x)$ be a continuous, bounded, real-valued function defined over $x \subset [a,b]$. Let $x(t) \subset \{a.p.\}$ be real-valued and have $a < x(t) < b$ for all t. Then $y(t) = \phi[x(t)]$ is a uniformly almost periodic function.

Proof: Since $\phi(x)$ is bounded and continuous on $[a,b]$, it is uniformly continuous on any interior interval of $[a,b]$, that is, for each $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|\phi(x) - \phi(x+\delta)| < \epsilon \text{ for any } x,\ x+\delta \subset (a,b). \tag{2.1.8}$$

We may write $\delta = \delta(\epsilon)$ to show that δ depends on ϵ but not on x. Now let

$\tau \subset E_\delta[x(t)]$, that is, let τ be a translation number of $x(t)$ corresponding to an error δ. Then by Definition 2.1.2,

$$D[x(t) - x(t+\tau)] < \delta = \delta(\epsilon). \tag{2.1.9}$$

Now, using equations (2.1.9) and (2.1.8), we can write

$$D\{\ \phi[x(t)] - \phi[x(t+\tau)]\ \} < D\{\ \phi[x(t)] - \phi[x(t) + \delta]\ \} < \epsilon. \tag{2.1.10}$$

Thus, $E_\delta[x(t)] \subset E_\epsilon\{\phi[x(t)]\}$, that is, each translation number τ of $x(t)$ corresponding to an error δ is also a translation number of $\phi[x(t)]$ corresponding to an error ϵ. Since the set $E_\delta[x(t)]$ is relatively dense, so is $E_\epsilon\{\phi[x(t)]\}$. Further, we may make ϵ as small as we wish by choosing δ sufficiently small. Hence, by definition, $\phi[x(t)]$ is a uniformly almost periodic function. Δ

2.2 Stepanoff Almost Periodic Functions

The basic theory of *S.a.p.* functions closely parallels that of *a.p.* functions. The differences are chiefly contained in a new definition of distance that allows for discontinuities on a set of Lebesgue measure zero:

Definition 2.2.1: Let $f(t)$ and $g(t)$ be defined for $-\infty < t < \infty$, and let $f(t)$, $g(t) \subset L^p$, $p \geq 1$; that is, let $|f(t)|^p$ and $|g(t)|^p$ be integrable in the Lebesgue sense. Then the Stepanoff distance, corresponding to a length $l > 0$, is defined as

$$D_{S_l^p}[f(t),g(t)] \equiv \underset{-\infty < t < \infty}{\text{upper bound}} \left[\frac{1}{l}\int_t^{t+l} |f(\gamma)-g(\gamma)|^p\, d\gamma\right]^{1/p}. \tag{2.2.1}$$

When either l or p is equal to unity, we will drop it from the notation; hence, we will write $D_S[f(t),g(t)]$ instead of $D_{S_1^1}[f(t),g(t)]$. We will also use the notation

$$D_{S_l^p}[f(t)] \equiv D_{S_l^p}[f(t),0]. \ \Delta \tag{2.2.2}$$

This new definition of distance also obeys the Triangle Inequality, so that if $f(t)$, $g(t)$, and $h(t) \subset L^p$ are defined for $-\infty < t < \infty$,

$$D_{S_l^p}[f(t),h(t)] \leq D_{S_l^p}[f(t),g(t)] + D_{S_l^p}[g(t),h(t)]. \tag{2.2.3}$$

We need to prove another form of this rule for later use. By rearranging equation (2.2.3), we have

$$D_{S_l^p}[f(t),h(t)] - D_{S_l^p}[g(t),h(t)] \leq D_{S_l^p}[f(t),g(t)]. \tag{2.2.4}$$

We distinguish two cases in equation (2.2.4): either the left–hand side is non–negative or it is negative. If it is the former, the desired result will follow. If it is the latter, we write the Triangle Inequality (2.2.3) with a permutation of f, g, and h so that

$$D_{S_l^p}[g(t),h(t)] - D_{S_l^p}[f(t),h(t)] \leq D_{S_l^p}[f(t),g(t)], \tag{2.2.5}$$

and now the left–hand side of equation (2.2.5) is positive. Combining the results of equations (2.2.4) and (2.2.5) we find that

$$\left| D_{S_l^p}[f(t),h(t)] - D_{S_l^p}[g(t),h(t)] \right| \leq D_{S_l^p}[f(t),g(t)], \tag{2.2.6}$$

which is the desired result. In particular, with $h = 0$, we have

$$\left| D_{S_l^p}[f(t)] - D_{S_l^p}[g(t)] \right| \leq D_{S_l^p}[f(t),g(t)]. \tag{2.2.7}$$

We will now use our definition of distance to investigate the structural properties of *S.a.p.* functions.

Structural Properties of $\{S_l^p.a.p.\}$

Definition 2.2.2: Let $f(t) \subset L^p$ be defined for $-\infty < t < \infty$, and let ϵ be a positive number. If some real number τ satisfies

$$D_{S_l^p}[f(t),f(t+\tau)] < \epsilon, \tag{2.2.8}$$

then τ is called an S_l^p translation number of $f(t)$ corresponding to an error ϵ. The set of all such translation numbers will be denoted $S_l^p E_\epsilon[f(t)]$. ∆

Definition 2.2.3: If the set $S_l^p E_\epsilon[f(t)]$ is relatively dense for all $\epsilon > 0$, then $f(t)$ is an S_l^p almost periodic function. ∆

We see from these definitions that S_l^p.*a.p.* functions, like their *a.p.* counterparts, exhibit a repetition of values, within an error ϵ, for values of t increased by τ. Here, however, the repetition is not in a pointwise sense, as defined by D, but rather in an integral or local average sense, as defined by $D_{S_l^p}$. In other words, making $D[f(t) - f(t+\tau)]$ small for an *a.p.* function $f(t)$ implies that $f(t) - f(t+\tau)$ is small at every point on the t axis. On the other hand, making $D_{S_l^p}[g(t) - g(t+\tau)]$ small for an S_l^p.*a.p.* function $g(t)$ implies only that $g(t) - g(t+\tau)$ is small at most points on the t axis.

It is clear from the definition of $D_{S_l^p}$ that an S_l^p.*a.p.* function is also an S^p.*a.p.* function. For simplicity, then, we will often take $l = 1$. Besicovitch and Bohr (1931) also point out that $\{S^{p''}.a.p.\} \subset \{S^{p'}.a.p.\}$ for any $1 \leq p' < p''$. Thus, all $\{S^p.a.p.\}$ functions are contained in the class $\{S.a.p.\}$.

Vibration Properties of $\{S_l^p.a.p.\}$

In analogy with Definitions 2.1.5 through 2.1.7, we can define an S_l^p limit function of a sequence of functions, an S_l^p limit function of an entire class of functions, and an S_l^p closure of a class of functions. We omit the details, the only change being the substitution of the distance $D_{S_l^p}$ for the distance D. We will denote the S_l^p closure of a class U by $C_{S_l^p}(U)$.

We again denote by A the class of all finite trigonometric polynomials. The main theorem connecting the structural and vibration properties of $\{S_l^p.a.p.\}$ can now be stated in analogy with Theorem 2.1.1:

Theorem 2.2.1: The S_l^p closure of A is identical to the class of $S_l^p.a.p.$ functions. That is,

$$C_{S_l^p}(A) = \{S_l^p.a.p.\}. \tag{2.2.9}$$

Proof: See Besicovitch and Bohr (1931), pp. 228–231.

Corollary: For every $x(t) \subset \{S_l^p.a.p.\}$, and every $\epsilon > 0$, there exists a $\Psi(t) \subset \{a.p.\}$ such that

$$D_{S_l^p}[x(t),\Psi(t)] < \epsilon. \tag{2.2.10}$$

That is, there exists a continuous function that can be made arbitrarily close in S_l^p distance to any $S_l^p.a.p.$ function. ∆

From Theorem 2.2.1 and its corollary, we may deduce immediately that the class $\{a.p.\}$ is contained in the class $\{S^p.a.p.\}$. (In fact, by comparing the distance D of equation (2.1.1) with the distance $D_{S_l^p}$ of equation (2.2.1), we may conclude that $\{a.p.\}$ is simply a limiting case of $\{S_l^p.a.p.\}$ for $p \to \infty$.) Thus, in future chapters we will occasionally speak of $S^p.a.p.$ functions with the understanding that our work also applies to $a.p.$ functions.

A Function of an $S_l^p.a.p.$ Function

We wish to prove a result for $S_l^p.a.p.$ functions analogous to Theorem 2.1.2.

Theorem 2.2.2: Let $\phi(x)$ be a continuous, bounded, real–valued function defined over $x \subset [a,b]$. Let $x(t) \subset \{S_l^p.a.p.\}$ be real–valued and have $a < x(t) < b$ for

all t. Then $y(t) = \phi[x(t)]$ is also an S_l^p.*a.p.* function.

Proof: From the corollary to Theorem 2.2.1, we can choose a function $\Psi(t)$ C $\{a.p.\}$ that is as close to $x(t)$, in the sense of $D_{S_l^p}$, as we desire. Making use of equations (2.2.2), (2.2.3), and (2.2.7), we then can write

$$\begin{aligned}
&\left| D_{S_l^p}\{ \phi[x(t)], \phi[x(t+\tau)] \} - D_{S_l^p}\{ \phi[\Psi(t)], \phi[\Psi(t+\tau)] \} \right| \\
&\quad = \left| D_{S_l^p}\{ \phi[x(t)] - \phi[x(t+\tau)] \} - D_{S_l^p}\{ \phi[\Psi(t)] - \phi[\Psi(t+\tau)] \} \right| \\
&\quad \le D_{S_l^p}\Big[\{\phi[x(t)] - \phi[x(t+\tau)]\} - \{\phi[\Psi(t)] - \phi[\Psi(t+\tau)]\} \Big] \\
&\quad = D_{S_l^p}\Big[\{\phi[x(t)] - \phi[\Psi(t)]\} - \{\phi[x(t+\tau)] - \phi[\Psi(t+\tau)]\} \Big] \\
&\quad \le D_{S_l^p}\{ \phi[x(t)] - \phi[\Psi(t)] \} + D_{S_l^p}\{ \phi[x(t+\tau)] - \phi[\Psi(t+\tau)] \}. \qquad (2.2.11)
\end{aligned}$$

The corollary to Theorem 2.2.1 tells us that we can make $D_{S_l^p}[x(t) - \Psi(t)]$ as small as we wish by choosing an appropriate $\Psi(t)$. Consequently, since $\phi(x)$ is a uniformly continuous function, we can also make $D_{S_l^p}\{ \phi[x(t)] - \phi[\Psi(t)] \}$ as small as we wish, and we certainly can make

$$D_{S_l^p}\{ \phi[x(t)] - \phi[\Psi(t)] \} < \epsilon/3 \qquad (2.2.12)$$

for any $\epsilon > 0$. Hence, using equations (2.2.11) and (2.2.12), we have

$$\begin{aligned}
&\left| D_{S_l^p}\{ \phi[x(t)], \phi[x(t+\tau)] \} - D_{S_l^p}\{ \phi[\Psi(t)], \phi[\Psi(t+\tau)] \} \right| \\
&\quad \le D_{S_l^p}\{ \phi[x(t)] - \phi[\Psi(t)] \} + D_{S_l^p}\{ \phi[x(t+\tau)] - \phi[\Psi(t+\tau)] \} \\
&\quad < \epsilon/3 + \epsilon/3. \qquad (2.2.13)
\end{aligned}$$

From equation (2.2.13) we deduce that

$$-2\epsilon/3 < D_{S_l^p}\{ \phi[x(t)], \phi[x(t+\tau)] \} - D_{S_l^p}\{ \phi[\Psi(t)], \phi[\Psi(t+\tau)] \} < 2\epsilon/3. \qquad (2.2.14)$$

Rearranging equation (2.2.14) gives

$$D_{S_l^p}\{\ \phi[\Psi(t)],\ \phi[\Psi(t+\tau)]\ \} - 2\epsilon/3$$

$$< D_{S_l^p}\{\ \phi[x(t)],\ \phi[x(t+\tau)]\ \}$$

$$< D_{S_l^p}\{\ \phi[\Psi(t)],\ \phi[\Psi(t+\tau)]\ \} + 2\epsilon/3. \qquad (2.2.15)$$

We have already proven that $\phi[\Psi(t)]$ is *a.p.*, so that it is also S_l^p.*a.p.*, and we can now choose τ such that

$$D_{S_l^p}\ \{\ \phi[\Psi(t)] - \phi[\Psi(t+\tau)]\ \} < \epsilon/3. \qquad (2.2.16)$$

Introducing equation (2.2.16) into (2.2.15) results in

$$D_{S_l^p}\{\ \phi[x(t)],\ \phi[x(t+\tau)]\ \} < \epsilon/3 + 2\epsilon/3 = \epsilon. \qquad (2.2.17)$$

Thus,

$$S_l^p E_{\epsilon/3}\{\ \phi[\Psi(t)]\ \} \subset S_l^p E_\epsilon\{\ \phi[x(t)]\ \}, \qquad (2.2.18)$$

and since the first of these sets is relatively dense, so is the second. By definition, then, the function $\phi[x(t)]$ is S_l^p.*a.p.* ∆

CHAPTER 3
Frequency Content of the Nonlinearity Output

In the previous chapter, we considered both the structural and vibration properties of $S^p.a.p.$ functions. We will henceforth concentrate on their vibration properties, that is, their properties as sums of sinusoids. In this chapter, we will investigate the frequency content of the output of a nonlinearity that is driven by a bounded, real-valued $S^p.a.p.$ function.

3.1 Representation of a Nonlinear Function by a Power Series

For computational purposes, we do not intend to represent the memoryless nonlinearity using a power series. Such a representation for a grossly nonlinear function will have many terms, and hence will be difficult to handle in a numerical context. For theoretical purposes, however, a power series representation has advantages due to its inherent simplicity. Here, we will use it to determine the frequencies that occur in the response of the nonlinearity when its excitation is an $S^p.a.p.$ function.

First, we must make precise what we mean by "representing" the nonlinearity. We may use the following result:

Theorem 3.1.1: Let $\phi(x)$ be a continuous, bounded, real-valued function defined for $x \in [-1,1]$. Corresponding to $\phi(x)$ there is a Fourier–Legendre series

$$\sum_{n=0}^{\infty} c_n P_n(x), \quad x \in (-1,1), \tag{3.1.1}$$

where P_n is the n^{th}-order Legendre polynomial and

$$c_n = \frac{2n+1}{2} \int_{-1}^{1} \phi(x)\, P_n(x)\, dx. \tag{3.1.2}$$

Then the Fourier-Legendre series for $\phi(x)$ converges uniformly to $\phi(x)$ in any interval interior to [-1,1].

Proof: See Rees, Shah, and Stanojevic (1981), pp. 337-340. ∆

We may apply this theorem to any continuous, bounded function $\phi(x)$ defined for $\alpha \le x \le \beta$ by making the change of variable

$$\gamma = \frac{2x - \alpha - \beta}{\beta - \alpha}, \quad -1 \le \gamma \le 1 \tag{3.1.3}$$

so that

$$x = \frac{(\beta - \alpha)\gamma + \alpha + \beta}{2}. \tag{3.1.4}$$

We then expand the function $\phi[x(\gamma)]$, treating γ as the independent variable, and finally insert the definition of γ in terms of x to give a power series having the form

$$\phi(x) = \sum_{n=0}^{\infty} c_n\, P_n[\gamma(x)], \quad x \in (\alpha,\beta). \tag{3.1.5}$$

The uniform convergence of the series representation for $\phi(x)$ allows us to find the nonlinearity output $y(t) = \phi[x(t)]$ by finding the output of each term $c_n\, P_n\{\gamma[x(t)]\}$ and adding the results. We will use this observation in the next section.

3.2 Expansion of Power Series Terms

We now wish to expand a power series term $c_n\, P_n[x(t)]$, that is, to determine its output as a sum of sinusoids. (Expansion of the more general term $c_n\, P_n\{\gamma[x(t)]\}$ proceeds in a similar manner.) Since each Legendre polynomial $P_n(x)$

contains terms involving x^n, x^{n-2}, x^{n-4},..., it is sufficient to investigate the expansion of $[x(t)]^n$.

We will consider two forms for the excitation $x(t)$, the first a finite sum of sinusoids, the second a possibly infinite sum of sinusoids whose frequencies have a finite basis. In each case, $x(t)$ is restricted to be bounded and real–valued.

Finite Number of Input Frequencies

Let the excitation $x(t)$ be

$$x(t) = \sum_{q=-Q}^{Q} a_q e^{j\lambda_q t}, \tag{3.2.1}$$

where the frequencies λ_q are real numbers and the coefficients a_q are in general complex numbers. Because $x(t)$ is real–valued, we require that

$$\lambda_q = -\lambda_{-q}$$
$$a_q = (a_{-q})^*$$

for each q, where * denotes "complex conjugate." Otherwise, we assume no relationships among the components of $x(t)$; in particular, they need not be commensurable (that is, they need not have a common, finite period). Clearly, $x(t)$ is an $S^p.a.p.$ function.

We then have

$$[x(t)]^n = \left[\sum_{q=-Q}^{Q} a_q e^{j\lambda_q t} \right]^n$$
$$= \sum_{q_1=-Q}^{Q} \sum_{q_2=-Q}^{Q} \cdots \sum_{q_n=-Q}^{Q} a_{q_1} a_{q_2} \cdots a_{q_n} e^{j(\lambda_{q_1}+\lambda_{q_2}+\cdots+\lambda_{q_n})t}. \tag{3.2.2}$$

Thus, the frequencies contributed by the term x^n are of the form

$$\lambda = \lambda_{q_1} + \lambda_{q_2} + \cdots + \lambda_{q_n} \tag{3.2.3}$$

where q_1, q_2,..., q_n run independently through the integers $-Q$, $-Q$+1,..., Q-1, Q. Expressed another way, these frequencies are of the form

$$\lambda = k_{-Q}\,\lambda_{-Q} + k_{-Q+1}\,\lambda_{-Q+1} + \cdots + k_{Q-1}\,\lambda_{Q-1} + k_Q\,\lambda_Q, \tag{3.2.4}$$

where, for a given permutation of the indices of summation in equation (3.2.2), k_i is the number of indices taking on the value i. It is obvious that

$$\sum_{i=-Q}^{Q} k_i = n. \tag{3.2.5}$$

These results give us a useful concept: that of the "order" of a nonlinearity. A nonlinearity is said to be of N^{th} order if its Fourier–Legendre series expansion terminates for $n = N$, or equivalently, if it contains no powers of x greater than N. As we have seen, a term x^N contributes frequencies that are all N–fold sums and differences of the input frequencies, so we may also speak of "N^{th}–order frequencies" as those appearing for the first time in a term x^N. We note that a term x^N also contributes frequencies of order N-2, N-4,..., because of permutations of the indices of summation in which one pair of indices has opposite values, two pairs have opposite values,... . Further, if the input $x(t)$ contains a d.c. component, that is, if a_0 is non–zero, then x^N contributes frequencies of all orders from 0 to N.

Input Frequencies with a Finite Basis

Let us now consider a finite set of frequencies $\{\lambda_m,\ m=1,2,...,M\}$ that are independent over the integers; that is, the relation

$$i_1\,\lambda_1 + i_2\,\lambda_2 + \cdots + i_M\,\lambda_M = 0 \tag{3.2.6}$$

cannot be satisfied for any integer coefficients i_m except $i_1 = i_2 = \cdots = i_M = 0$. Clearly, the frequencies $\{\lambda_m\}$ are incommensurable, although not every set of incommensurable frequencies is independent over the integers. (For example, the frequencies 1, $\sqrt{2}$, and $1 + \sqrt{2}$ are incommensurable but not independent.) We wish to consider an input $x(t)$ that involves only frequencies that are expressible as linear combinations of the $\{\lambda_m\}$ using integer coefficients. Thus, we will let

$$x(t) = \sum_{q_1=-Q_1}^{Q_1} \sum_{q_2=-Q_2}^{Q_2} \cdots \sum_{q_M=-Q_M}^{Q_M} a_{q_1,q_2,\cdots,q_M} e^{j(q_1\lambda_1+q_2\lambda_2+\cdots+q_M\lambda_M)t}, \tag{3.2.7}$$

and again, $x(t)$ is $S^p.a.p.$ Applying this input to a term x^n gives

$$[x(t)]^n = \left[\sum_{q_{11}=-Q_1}^{Q_1} \sum_{q_{12}=-Q_1}^{Q_1} \cdots \sum_{q_{1n}=-Q_1}^{Q_1}\right]\left[\sum_{q_{21}=-Q_2}^{Q_2} \sum_{q_{22}=-Q_2}^{Q_2} \cdots \sum_{q_{2n}=-Q_2}^{Q_2}\right] \cdots$$
$$\left[\sum_{q_{M1}=-Q_M}^{Q_M} \sum_{q_{M2}=-Q_M}^{Q_M} \cdots \sum_{q_{Mn}=-Q_M}^{Q_M}\right]$$
$$(a_{q_{11},q_{21},\cdots,q_{M1}})(a_{q_{12},q_{22},\cdots,q_{M2}}) \cdots (a_{q_{1n},q_{2n},\cdots,q_{Mn}})$$
$$exp\Big[\, jt[(q_{11}+q_{12}+\cdots+q_{1n})\lambda_1 + (q_{21}+q_{22}+\cdots+q_{2n})\lambda_2 + \cdots$$
$$+ (q_{M1}+q_{M2}+\cdots+q_{Mn})\lambda_M] \Big]. \tag{3.2.8}$$

We thus see that $[x(t)]^n$ has the same basis frequencies as does $x(t)$. This result is true for all n, even for $n \to \infty$, so by our remarks at the end of Section 3.1, we conclude that $x(t)$ and $y(t) = \phi[x(t)]$ have the same basis frequencies.

We must be more careful about drawing this same conclusion for $Q_1, Q_2, \ldots, Q_M \to \infty$, however. If the series of equation (3.2.7) contains infinitely many terms, it is not necessarily allowable to equate this series to the function $x(t)$, since to do so would assume uniform convergence of the series, a property that is not guaranteed.

It will turn out that our conclusion is still true, namely that $x(t)$ and $\phi[x(t)]$ have the same basis frequencies, but we must wait until the end of Chapter 4 to prove it.

3.3 Enumeration of Frequencies in the Nonlinearity Output

It is useful to place a bound on the number of frequencies that result when an N^{th}–order nonlinearity is driven by a particular input. We will consider the simplest type of input, namely, one having a finite number of sinusoids with independent frequencies.

Let the input be

$$x(t) = \sum_{m=-M}^{M} a_m e^{j\lambda_m t} \tag{3.3.1}$$

where the λ_m, $1 \le m \le M$, are independent over the integers, and $\lambda_m = -\lambda_{-m}$ for each m. With this restriction, there can be no d.c. component in $x(t)$, since the frequency $\lambda_0 = 0$ is not independent of the remaining frequencies.

Let us first find the number of n^{th}–order frequencies for any $n \ge M$. As we saw in the previous section, n^{th}–order frequencies are those that arise for the first time in the term x^n; they occur when no two indices of summation in equation (3.2.2) assume opposite values.

We can model our problem using a basic concept of combinatorial theory, that of balls and cells. Let us define M cells, labeled with the integers 1 through M. We will further label each cell with either a "+" or "–" sign, with the choice made independently for each cell. These cells then represent the possible values for the indices of summation in equation (3.2.2), and this labeling scheme prevents any two indices from taking on opposite values. Let us also define n "indistinguishable"

balls to represent the indices of summation. A particular combination of indices is thus modeled by a corresponding distribution of the n balls into the M cells along with a corresponding scheme for labeling the cells as + or −. Since the balls are indistinguishable, we will be considering combinations, instead of permutations, of the indices.

Some thought will reveal that the total number of combinations in this model may be expressed as the sum of M terms, with the terms representing mutually–exclusive distributions of the balls: (We will call a cell "full" if it contains at least one ball.)

combinations

= (# ways of placing n balls into M cells such that 1 cell is full)

·(# ways of labeling the 1 cell as + or −)

+ (# ways of placing n balls into M cells such that 2 cells are full)

·(# ways of labeling the 2 cells as + or −)

+ ···

+ (# ways of placing n balls into M cells such that M cells are full)

·(# ways of labeling the M cells as + or −). (3.3.2)

The second factor in each term of equation (3.3.2) is 2^1, 2^2,..., 2^M. The first factor may be computed using well known results of combinatorial theory [see Roberts (1984), for example]. We write this factor, for the case of k cells full, as

(# ways of placing n balls into M cells such that k cells are full)

= (# ways of placing n balls into k distinguishable cells, no cell empty)

·(# ways to choose k cells from a set of M distinguishable cells)

$$= C(n\text{-}1,k\text{-}1) \cdot C(M,k)$$

$$= \frac{(n\text{-}1)!}{(k\text{-}1)!\ (n\text{-}k)!} \cdot \frac{M!}{k!\ (M\text{-}k)!}. \qquad (3.3.3)$$

The expression $C(M,k)$ is called the binomial coefficient. Combining results, we have

$$\#\text{ combinations} = \sum_{k=1}^{M} \frac{(n-1)!}{(k-1)!\ (n-k)!} \cdot \frac{M!}{k!\ (M-k)!} \cdot 2^k\ ,\ M \le n. \qquad (3.3.4)$$

We now consider the case of $n < M$. In examining equation (3.3.2), we note that only the first n terms will be non–zero, since if the number of cells exceeds the number of balls, no more than n cells can be full. Consequently, we must let k range from 1 to n instead of 1 to M in equation (3.3.4). Otherwise, no changes are needed.

A result applicable for any M and n is thus

$$\#\text{ combinations} = \sum_{k=1}^{min(M,n)} \frac{(n-1)!}{(k-1)!\ (n-k)!} \cdot \frac{M!}{k!\ (M-k)!} \cdot 2^k. \qquad (3.3.5)$$

This result tells us the number of n^{th}–order frequencies (counting both positive and negative frequencies) arising from an input having M independent frequencies. As we found earlier, a power series term x^n additionally contributes frequencies of order n-2, n-4,... . Further, if the input contains a d.c. component, the power series term x^n contributes frequencies of all orders.

As a final result, we find the total number of frequencies at the output of an N^{th}–order nonlinearity to be

$$\#\text{ frequencies} = 1 + M! \sum_{n=1}^{N} (n-1)! \left[\sum_{k=1}^{min(M,n)} \frac{2^k}{(k-1)!\ k!\ (n-k)!\ (M-k)!} \right]. \qquad (3.3.6)$$

The unity term in this expression accounts for d.c.

3.4 Distribution of Frequencies

Let us return to the finite basis formulation of Section 3.2, in which any frequency at the output of the nonlinearity can be written as

$$\lambda = q_1 \lambda_1 + q_2 \lambda_2 + \cdots + q_M \lambda_M$$

for integer q_m. We wish to answer the following question: For M finite and fixed, what is the distribution of the frequencies λ on the real axis as q_1, q_2,..., q_M approach infinity independently? That is, does the spacing of frequencies on the axis approach some finite minimum, or do the frequencies become dense?

Let us begin by considering a fixed frequency

$$\lambda' = q_1' \lambda_1 + q_2' \lambda_2 + \cdots + q_M' \lambda_M \tag{3.4.1}$$

and a "nearby" frequency

$$\lambda'' = q_1'' \lambda_1 + q_2'' \lambda_2 + \cdots + q_M'' \lambda_M \tag{3.4.2}$$

We define their difference as

$$\begin{aligned} \Delta\lambda \equiv \lambda'' - \lambda' &= (q_1'' - q_1') \lambda_1 + (q_2'' - q_2') \lambda_2 + \cdots + (q_M'' - q_M') \lambda_M \\ &\equiv \Delta q_1 \lambda_1 + \Delta q_2 \lambda_2 + \cdots + \Delta q_M \lambda_M \end{aligned} \tag{3.4.3}$$

and group this expression as

$$\left[\Delta q_1 \lambda_1 + \Delta q_2 \lambda_2\right] - -\left[\Delta q_3 \lambda_3 + \Delta q_4 \lambda_4 + \cdots + \Delta q_M \lambda_M\right] = \Delta\lambda. \tag{3.4.4}$$

(We could have chosen any two of the $\Delta q_m \lambda_m$ for the left-hand bracketed quantity.) Our question now becomes "How close can $(\Delta q_1 \lambda_1 + \Delta q_2 \lambda_2)$ be made to $-(\Delta q_3 \lambda_3 + \Delta q_4 \lambda_4 + \cdots + \Delta q_M \lambda_M)$, or more generally, to an arbitrary real number?" Accordingly, we consider the behavior of

$$|\Delta q_1 \lambda_1 + \Delta q_2 \lambda_2 - \theta| < \epsilon \tag{3.4.5}$$

for real θ and integer Δq_1, Δq_2, and we ask whether this inequality can be satisfied for an arbitrarily small ϵ.

With the problem stated in this form, the result follows immediately from a theorem due to Minkowski:

Theorem 3.4.1: Let $ax + by + c$ be a linear form with real coefficients a, b, and c. Suppose that a/b is irrational and that $ax + by + c = 0$ has no solutions for integer x, y. Then for any $\epsilon > 0$ there are infinitely many pairs of integers x, y such that

$$|ax + by + c| < \epsilon.$$

Proof: See Cassels (1957), pp. 46–48. ∎

Thus, we can indeed make $(\Delta q_1 \lambda_1 + \Delta q_2 \lambda_2)$ arbitrarily close to $-(\Delta q_3 \lambda_3 + \Delta q_4 \lambda_4 + \cdots + \Delta q_M \lambda_M)$, and consequently make $\Delta\lambda$ in equation (3.4.3) arbitrarily small. Hence, we conclude that the real line becomes dense with frequencies λ as q_1, q_2,..., q_M approach infinity.

CHAPTER 4
Computation of Fourier Series Coefficients

In Chapter 2, we saw that $S^p.a.p.$ functions may be approximated uniformly, in the sense of the distance D_{S^p}, by finite trigonometric polynomials. Hence, any $S^p.a.p.$ function $f(t)$ possesses a Fourier series, at least formally, and we write

$$f(t) \sim \sum_{n=-\infty}^{\infty} c_n e^{j\lambda_n t}$$

to indicate the formal correspondence. In the previous chapter, we addressed the question of finding the λ_n for the $S^p.a.p.$ response of a memoryless nonlinearity. Here, we will address the question of actually constructing the Fourier series. We will begin with computation of the coefficients, and then consider summability properties of the series.

4.1 Construction of the Fourier Series

Mean Value Over an Infinite Interval

We begin by defining the mean value of an $S^p.a.p.$ function $f(t)$ as

$$M\{f(t)\} \equiv \lim_{T\to\infty} \frac{1}{T} \int_0^T f(t)\, dt. \tag{4.1.1}$$

Besicovitch and Bohr (1931) prove that this mean value actually exists, that is, that the integral approaches a finite limit for any $S^p.a.p.$ function. It is further possible to show that this single-sided mean value is the same as the two-sided mean value

$$M\{f(t)\} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(t)\, dt, \tag{4.1.2}$$

and indeed the same as the mean value taken over any arbitrary infinite interval.

Orthogonality Over an Infinite Interval

We may use this concept of mean value to construct a set of orthonormal functions over the infinite interval $-\infty < t < \infty$. These functions will be the non–enumerable set of complex exponentials $\{e^{j\lambda t}\}$, where λ is any real number. It is easy to see that this set is both orthogonal and normalized, because

$$\begin{aligned} M\{e^{j\lambda_v t}\, e^{-j\lambda_n t}\} &= \lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} e^{j\lambda_v t}\, e^{-j\lambda_n t}\, dt \\ &= \begin{cases} 0 \text{ for } \lambda_v \neq \lambda_n \\ 1 \text{ for } \lambda_v = \lambda_n, \end{cases} \end{aligned} \tag{4.1.3}$$

since

$$\begin{aligned} \int_{0}^{T} e^{j\lambda_v t}\, e^{-j\lambda_n t}\, dt &= \int_{0}^{T} e^{j(\lambda_v - \lambda_n)t}\, dt \\ &= \begin{cases} \dfrac{e^{j(\lambda_v - \lambda_n)T} - 1}{j(\lambda_v - \lambda_n)} & \text{for } \lambda_v \neq \lambda_n \\ T & \text{for } \lambda_v = \lambda_n. \end{cases} \end{aligned} \tag{4.1.4}$$

Parseval's Equation

If we take the product of any $S^p.a.p.$ function $f(t)$ with a complex exponential $e^{-j\lambda t}$, we obtain another $S^p.a.p.$ function. Consequently, the mean value of the

product exists; we will denote it by $a(\lambda)$, and in analogy with periodic functions we will call it a Fourier series coefficient of $f(t)$. Hence,

$$a(\lambda) \equiv M\{f(t)\, e^{-j\lambda t}\} = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(t)\, e^{-j\lambda t} dt. \tag{4.1.5}$$

With these concepts, we can proceed to construct the Fourier series for $S^p.a.p.$ functions in the same manner as for purely periodic functions. The starting point for both is the representation of a function $f(t)$ by a finite sum of sinusoids, namely,

$$f(t) \simeq \sum_{n=-N}^{N} c_n\, e^{j\lambda_n t}. \tag{4.1.6}$$

For the purely periodic case these sinusoids had harmonically–related frequencies and we considered orthogonality over a finite interval; for the case of almost periodic functions (both *a.p.* and $S^p.a.p.$) the frequencies are chosen from the non–enumerable set of all real numbers, and we must consider orthogonality over an infinite interval, as shown by equation (4.1.3). Fortunately, it will turn out that only an enumerable subset of these non–enumerable frequencies can have non–zero Fourier series coefficients.

We now form the mean–square error between $f(t)$ and its series representation, giving

$$M\left\{\left|f(t) - \sum_{n=-N}^{N} c_n\, e^{j\lambda_n t}\right|^2\right\} = M\left\{\left[f(t) - \sum_{n=-N}^{N} c_n\, e^{j\lambda_n t}\right]\left[f^*(t) - \sum_{n=-N}^{N} c_n^*\, e^{-j\lambda_n t}\right]\right\}$$

$$= M\left\{f(t)\, f^*(t)\right\} - \sum_{n=-N}^{N} c_n^*\, M\left\{f(t)\, e^{-j\lambda_n t}\right\} - \sum_{n=-N}^{N} c_n\, M\left\{f^*(t)\, e^{j\lambda_n t}\right\}$$

$$+ \sum_{n_1=-N}^{N} \sum_{n_2=-N}^{N} c_{n_1}\, c_{n_2}^*\, M\left\{e^{j\lambda_{n_1} t}\, e^{-j\lambda_{n_2} t}\right\}. \tag{4.1.7}$$

All of the mean values in equation (4.1.7) exist because both the expression

$$f(t) - \sum c_n e^{j\lambda_n t}$$

and its magnitude squared are $S^p.a.p.$

If we use the orthogonality relation of equation (4.1.3), and the definition for $a(\lambda)$ of equation (4.1.5), then equation (4.1.7) reduces to

$$M\left\{\left|f(t) - \sum_{n=-N}^{N} c_n e^{j\lambda_n t}\right|^2\right\} = M\{|f(t)|^2\} - \sum_{n=-N}^{N} c_n^* a(\lambda_n) - \sum_{n=-N}^{N} c_n a^*(\lambda_n) + \sum_{n=-N}^{N} c_n c_n^*$$

$$= M\{|f(t)|^2\} + \sum_{n=-N}^{N} [c_n - a(\lambda_n)][c_n^* - a^*(\lambda_n)] - \sum_{n=-N}^{N} a(\lambda_n) a^*(\lambda_n)$$

$$= M\{|f(t)|^2\} - \sum_{n=-N}^{N} |a(\lambda_n)|^2 + \sum_{n=-N}^{N} |c_n - a(\lambda_n)|^2. \qquad (4.1.8)$$

If we now choose $c_n = a(\lambda_n)$ for each n, it is apparent that the mean–square error given by equation (4.1.8) is a minimum. In this case,

$$M\left\{\left|f(t) - \sum_{n=-N}^{N} a(\lambda_n) e^{j\lambda_n t}\right|^2\right\} = M\{|f(t)|^2\} - \sum_{n=-N}^{N} |a(\lambda_n)|^2, \qquad (4.1.9)$$

and because the right–hand side must be non–negative,

$$\sum_{n=-N}^{N} |a(\lambda_n)|^2 \le M\{|f(t)|^2\}. \qquad (4.1.10)$$

This last relation establishes that $a(\lambda)$ can be non–zero for, at most, an enumerably (countably) infinite set of frequencies λ_n. (The uncountably infinite sum of non–zero quantities $|a(\lambda_n)|^2$ cannot possibly be finite, as the inequality requires.) This equation also shows that to approximate $f(t)$ in the mean (that is, with minimum mean–square error) using a finite trigonometric sum, we should choose the terms of this sum from among the terms of the Fourier series for $f(t)$; any other choice of frequencies or of coefficients would cause a greater mean–square error.

The true Parseval equation

$$\sum_{n=-\infty}^{\infty} |a(\lambda_n)|^2 = M\{|f(t)|^2\} \tag{4.1.11}$$

does not immediately follow from the above development. Bohr (1951) proves the Parseval equation by first showing that it is equivalent to the uniqueness theorem [pp. 62–63], and then proving the latter [pp. 70–80]. Although his proof was carried out for *a.p.* functions, it is valid for all types of almost periodic function, including S^p*.a.p.*

The primary implication of Parseval's equation is that, for $N \to \infty$, the Fourier series of $f(t)$ converges in the mean to $f(t)$, that is,

$$\lim_{N\to\infty} \sum_{n=-N}^{N} |a(\lambda_n)|^2 = M\{|f(t)|^2\} \tag{4.1.12}$$

or

$$\lim_{N\to\infty} M\left\{\left|f(t) - \sum_{n=-N}^{N} a(\lambda_n)e^{j\lambda_n t}\right|^2\right\} = 0. \tag{4.1.13}$$

It says nothing about other types of convergence, such as a pointwise convergence of the series to the function $f(t)$. Consequently, the structural and vibration properties of an almost periodic function have not yet been connected. This connection will be established in the next section, when we consider summability of the Fourier series. Let us first state the important uniqueness theorem, which follows directly from Parseval's equation, (4.1.11).

Theorem 4.1.1 (Uniqueness Theorem): Let $f(t)$ and $g(t)$ be S^p*.a.p.* functions. The following statements, all forms of the uniqueness theorem, are equivalent:

1. If $f(t)$ and $g(t)$ have precisely the same Fourier series, then

   $D_{S^p}[f(t),g(t)] = 0.$

That is, the function $f(t) - g(t)$ is zero almost everywhere (everywhere except on a set of Lebesgue measure zero).

2. If the Fourier series of $f(t)$ contains no non–zero terms, then

$$D_{S^p}[f(t)] = 0.$$

That is, $f(t)$ is zero almost everywhere.

3. The system $\{e^{j\lambda t}\}$ is complete in the class of $S^p.a.p.$ functions.

Proof: See Bohr (1951), pp. 70–80. ∆

Statement 3 may be taken as a definition of completeness. In the special case of $f(t)$, $g(t) \subset \{a.p.\}$, the Stepanoff distance D_{S^p} may be replaced by the uniform distance D, since $D_{S^p} \to 0$ then implies $D \to 0$. Consequently, in statements 1 and 2, "almost everywhere" may be replaced by "everywhere."

4.2 Summability of the Fourier Series

Thus far, we have considered only convergence in the mean to relate $f(t)$ and its Fourier series. This form of convergence merely requires that $f(t)$ and its series representation be the same in an "average power" sense, viewed over all t. More restrictive types of convergence require that the Fourier series of $f(t)$ be a good approximation to the function when viewed over finite intervals, or even at each point on the t axis.

The most desirable type of convergence is uniform convergence, which requires that

$$\underset{-\infty < t < \infty}{\text{upper bound}} \left| f(t) - \sum_{n=-N}^{N} a(\lambda_n)\, e^{j\lambda_n t} \right| < \epsilon \quad \text{for } N \geq N_0 = N_0(\epsilon) \tag{4.2.1}$$

for any $\epsilon > 0$. The result for purely periodic functions is well–known:

Theorem 4.2.1 (Jordan's Test): If $f(t)$ is periodic with period $2\pi/\omega$, and is of bounded variation[1], then its Fourier series converges at each point t_0 to the average of its left– and right–hand limits, that is,

$$\lim_{N\to\infty} \sum_{n=-N}^{N} a(n\omega)\, e^{jn\omega t_0} = [\lim_{t \uparrow t_0} f(t) + \lim_{t \downarrow t_0} f(t)]\ /2. \tag{4.2.2}$$

In particular, if $f(t)$ is continuous at t_0, the Fourier series converges to $f(t_0)$. Further, the convergence is uniform within any interval where $f(t)$ is continuous.

Proof: See Rees, Shah, and Stanojevic, pp. 140–141. ∆

Unfortunately, a result corresponding to Jordan's Test does not exist in the general case for either *a.p.* or S^p*.a.p.* functions. Only for an *a.p.* function having independent frequencies in its Fourier series do we obtain uniform convergence [see Corduneanu (1968)]. However, by employing the more general concept of the summability of a Fourier series, we may still construct what is needed: finite trigonometric sums that approximate an *a.p.* or S^p*.a.p.* function arbitrarily well, in the sense of the appropriate distance. We will see that the Fourier series for all types of almost periodic functions, while they may not be convergent, are always summable. We will begin with the results for purely periodic functions to illustrate the basic concepts.

Summability Concepts

Let $f(t)$ be a periodic function with period $2\pi/\omega$. We assume that $f(t)$ has a Fourier series

---

1 A function is said to be of bounded variation if, roughly speaking, it has a finite number of finite excursions on any finite partition of its domain.

$$f(t) \sim \sum_{v=-\infty}^{\infty} a(v\omega)\, e^{jv\omega t}. \tag{4.2.3}$$

We will call $s_n(t)$ the n^{th} partial sum of this series; thus,

$$s_n(t) \equiv \sum_{v=-n}^{n} a(v\omega)\, e^{jv\omega t}. \tag{4.2.4}$$

If we now form the average

$$\begin{aligned} S_n(t) &\equiv \frac{1}{n}\,(s_0 + s_1 + \cdots + s_{n-1}) \\ &= \sum_{v=-n}^{n} \left[\, 1 - \frac{|v|}{n} \,\right] a(v\omega)\, e^{jv\omega t}, \end{aligned} \tag{4.2.5}$$

we say that the series is summable with sum U if $S_n(t) \to U$ for $n \to \infty$. A well-known theorem due to Fejér gives the conditions under which the Fourier series of $f(t)$ is summable.

Theorem 4.2.2 (Fejér's Theorem): Let $f(t) \subset L^1$ be periodic. If the right- and left-hand limits of $f(t)$ at t_0 both exist, then the Fourier series of $f(t)$ is summable to the average of these limits, that is,

$$\lim_{n\to\infty} S_n(t) = [\lim_{t \downarrow t_0} f(t) + \lim_{t \uparrow t_0} f(t)]\ /2. \tag{4.2.6}$$

In particular, the series sums to $f(t_0)$ at points of continuity. Further, within regions of continuity, the convergence of S_n is uniform in t.

Proof: See Rees, Shah, and Stanojevic, p. 149. ∆

Summability is a more robust process than convergence, as we can see by comparing the requirements of Theorem 4.2.1 with those of Theorem 4.2.2. A series that is convergent to a certain limit will always be summable to the same limit; however, not all series that are summable are convergent. Summability works by introducing new terms gradually as n grows. That is, in going from S_n to S_{n+1} in equation (4.2.5), we acquire new terms $a(n\omega)\, e^{jn\omega t}/(n+1)$ and $a(-n\omega)\, e^{-jn\omega t}/(n+1)$.

In contrast, in going from s_{n-1} to s_n in equation (4.2.4), we acquire new terms $a(n\omega)\ e^{jn\omega t}$ and $a(-n\omega)\ e^{-jn\omega t}$.

The expression

$$K_n(\omega t) \equiv \sum_{v=-n}^{n} \left[1 - \frac{|v|}{n} \right] e^{-jv\omega t} \tag{4.2.7}$$

is known as the Fejér kernel. An alternate form for $S_n(t)$ in terms of $K_n(\omega t)$ is

$$S_n(t) = M_x\{f(x + t)\ K_n(\omega x)\} \tag{4.2.8}$$

where M_x denotes the mean value with respect to x. This result follows because

$$f(x + t) \sim \sum_{k=-\infty}^{\infty} a(k\omega)\ e^{jk\omega t}\ e^{jk\omega x}, \tag{4.2.9}$$

so that

$$\begin{aligned} M_x\{f(x + t)\ K_n(\omega x)\} &= M_x\left\{ \sum_{k=-\infty}^{\infty} a(k\omega)\ e^{jk\omega t}\ e^{jk\omega x} \sum_{v=-n}^{n} \left[1 - \frac{|v|}{n} \right] e^{-jv\omega x} \right\} \\ &= \sum_{v=-n}^{n} \left[1 - \frac{|v|}{n} \right] \sum_{k=-\infty}^{\infty} a(k\omega)\ e^{jk\omega t}\ M_x\{ e^{j(k-v)\omega x} \} \\ &= \sum_{v=-n}^{n} \left[1 - \frac{|v|}{n} \right] a(v\omega)\ e^{jv\omega t}. \end{aligned} \tag{4.2.10}$$

Summability of Fourier Series for a.p. Functions

Bohr (1951), using results from Bochner, extended the concept of summability to the Fourier series for an *a.p.* function $f(t)$. His first task was to induce an ordering in the terms of the Fourier series, since there is no natural order corresponding to the result for periodic functions. He accomplished this ordering by means of a basis for the frequencies in $f(t)$.

The concept of a basis was discussed in Chapter 3, where we assumed that a set of frequencies had a finite basis with integer coefficients. Bohr generalized this concept in two ways to allow the representation of an arbitrary, enumerably infinite set of frequencies. First, he allowed the basis itself to contain an enumerably infinite number of frequencies, and second, he allowed rational and not merely integer coefficients.

Thus, Bohr proposed to represent each member of the set $\{\lambda_i\}$, that is, the set of frequencies in $f(t)$, as

$$\lambda_i = r_1 \beta_1 + r_2 \beta_2 + \cdots + r_m \beta_m + \cdots, \tag{4.2.11}$$

where r_1, r_2,... are rational numbers and the basis frequencies $\{\beta_m\}$ are independent over the rational numbers. It is easy to see that any particular λ_i has such a representation for sufficiently large m.

We will now choose a particular form for the numbers r_m, following Besicovitch and Bohr (1931). Let N_1, N_2,..., N_m,... be positive integers, let n_1, n_2,..., n_m,... be positive integers such that $n_m = N_m N_m!$ for each m, and let v_1, v_2,..., v_m,... be integers such that $|v_m| \leq n_m$ for each m. We will then write equation (4.2.11) as

$$\lambda_i = \frac{v_1}{N_1!} \beta_1 + \frac{v_2}{N_2!} \beta_2 + \cdots + \frac{v_m}{N_m!} \beta_m + \cdots. \tag{4.2.12}$$

Again, each λ_i has such a representation for sufficiently large m and sufficiently large N_1, N_2,..., N_m. We have chosen this particular form for the following reason: If a particular λ_i has the representation of equation (4.2.12) for certain values of m, N_1, N_2,..., N_m, then it will also have a representation for any larger values of m, N_1, N_2,..., N_m.

In analogy with the one–dimensional Fejér kernel of equation (4.2.7), we now construct the m–dimensional Bochner–Fejér kernel

$$K_{\begin{bmatrix} n_1, n_2, \ldots, n_m \\ \beta_1, \beta_2, \ldots, \beta_m \end{bmatrix}}(t) \equiv K_{n_1}(\beta_1 t/N_1!)\, K_{n_2}(\beta_2 t/N_2!) \cdots K_{n_m}(\beta_m t/N_m!)$$

$$= \sum_{v_1=-n_1}^{n_1} \sum_{v_2=-n_2}^{n_2} \cdots \sum_{v_m=-n_m}^{n_m} \left[1 - \frac{|v_1|}{n_1}\right] \left[1 - \frac{|v_2|}{n_2}\right] \cdots \left[1 - \frac{|v_m|}{n_m}\right]$$

$$exp\left\{-jt\left[\frac{v_1\beta_1}{N_1!} + \frac{v_2\beta_2}{N_2!} + \cdots + \frac{v_m\beta_m}{N_m!}\right]\right\} \qquad (4.2.13)$$

and use this composite kernel to form the sum

$$\sigma_{\begin{bmatrix} n_1, n_2, \ldots, n_m \\ \beta_1, \beta_2, \ldots, \beta_m \end{bmatrix}}(t) = M_x\{f(x+t)\, K_{\begin{bmatrix} n_1, n_2, \ldots, n_m \\ \beta_1, \beta_2, \ldots, \beta_m \end{bmatrix}}(x)\}$$

$$= \sum_{v_1=-n_1}^{n_1} \sum_{v_2=-n_2}^{n_2} \cdots \sum_{v_m=-n_m}^{n_m} a\left[\frac{v_1\beta_1}{N_1!} + \frac{v_2\beta_2}{N_2!} + \cdots + \frac{v_m\beta_m}{N_m!}\right]$$

$$\left[1 - \frac{|v_1|}{n_1}\right] \left[1 - \frac{|v_2|}{n_2}\right] \cdots \left[1 - \frac{|v_m|}{n_m}\right]$$

$$exp\left\{jt\left[\frac{v_1\beta_1}{N_1!} + \frac{v_2\beta_2}{N_2!} + \cdots + \frac{v_m\beta_m}{N_m!}\right]\right\}. \qquad (4.2.14)$$

We may use such sums, called Bochner–Fejér polynomials, to approximate the *a.p.* function $f(t)$, in the following sense:

Theorem 4.2.3: The sums $\sigma_{\begin{bmatrix} n_1, n_2, \ldots, n_m \\ \beta_1, \beta_2, \ldots, \beta_m \end{bmatrix}}(t)$ converge uniformly to $f(t)$ as $m \to \infty$, $N_1 \to \infty$, $N_2 \to \infty, \ldots$. That is, for any $\epsilon > 0$ there exist integers m', N_1', $N_2', \ldots$, such that

$$D[f(t), \sigma_{\begin{bmatrix} n_1, n_2, \ldots, n_m \\ \beta_1, \beta_2, \ldots, \beta_m \end{bmatrix}}(t)] < \epsilon \qquad (4.2.15)$$

whenever $m > m'$, $N_1 > N_1'$, $N_2 > N_2', \ldots$.

Proof: See Bohr (1951), pp. 80–88. ∎

This theorem gives the connection between the structural and vibration properties of *a.p.* functions, for it tells us that we may approximate any *a.p.* function arbitrarily well (in the sense of the uniform distance D) using Bochner–Fejér polynomials.

Summability of Fourier Series for S^p.*a.p.* Functions

The foregoing results carry over to the case of S^p.*a.p.* functions, with an appropriate change in the convergence measure. To simplify the notation, let $\sigma_k(t)$ be a particular $\sigma_{\begin{bmatrix} n_1,n_2,\ldots,n_m \\ \beta_1,\beta_2,\ldots,\beta_m \end{bmatrix}}(t)$, and let $\sigma_{k+1}(t)$ be another polynomial having larger m, N_1, N_2,... . Then we have the following result:

Theorem 4.2.4: For $l > 0$ and any $\epsilon > 0$, there exists a k_0 such that

$$D_{S_l^p}[f(t),\sigma_k(t)] < \epsilon \text{ for } k \geq k_0. \tag{4.2.16}$$

Consequently, taking $l = 1$ for convenience, we conclude that

$$D_{S^p}[f(t),\sigma_k(t)] \to 0 \text{ for } k \to \infty. \tag{4.2.17}$$

Proof: See Besicovitch and Bohr (1931), pp. 263–264. ∆

The choice of $l = 1$ is, of course, arbitrary; if the sequence of polynomials converges to $f(t)$ for $l = 1$, then it must converge for any $l > 0$. Let us choose a particular $l = l_0$ and any corresponding interval $[t,t+l_0]$ where $f(t)$ is uniformly continuous. Then on any interior subinterval the sequence of polynomials converges uniformly to $f(t)$, that is,

$$\lim_{k\to\infty} D[f(t),\sigma_k(t)] = 0. \tag{4.2.18}$$

This result follows because the Bochner–Fejér polynomials $\sigma_k(t)$ are all uniformly continuous and because, in regions of continuity, the Stepanoff distance D_{S^p} may be

replaced by the uniform distance D.

We may now use concepts developed in this chapter to verify a statement that we made at the end of Section 3.2, namely, that the nonlinearity input $x(t)$ and output $y(t) = \phi[x(t)]$ have the same basis frequencies. We previously proved this statement for an $x(t)$ having the finite Fourier series of equation (3.2.7). What remains is to prove it for an $x(t)$ having an infinite Fourier series.

Let $x(t)$ have the Fourier series representation

$$x(t) \sim \sum_{q_1=-Q_1}^{Q_1} \sum_{q_2=-Q_2}^{Q_2} \cdots \sum_{q_M=-Q_M}^{Q_M} a(q_1\lambda_1+q_2\lambda_2+\cdots+q_M\lambda_M)\, e^{j(q_1\lambda_1+q_2\lambda_2+\cdots+q_M\lambda_M)t}, \tag{4.2.19}$$

where now $Q_1, Q_2, \ldots, Q_M \to \infty$. We know that we can approximate $x(t)$ uniformly in the sense of the distance $D_{\mathcal{S}^p}$, using Bochner-Fejér polynomials $\sigma_k(t)$ that contain integer combinations of the basis frequencies $\lambda_1, \lambda_2, \ldots, \lambda_M$. From the definition of $D_{\mathcal{S}^p}$, these polynomials have the property that

$$\lim_{k\to\infty} \sigma_k(t) \underset{a.e.}{=} x(t) \tag{4.2.20}$$

where $\underset{a.e.}{=}$ means "equal almost everywhere." Because $\phi(x)$ is a uniformly continuous function, it is also true that

$$\phi\left[\lim_{k\to\infty} \sigma_k(t)\right] = \lim_{k\to\infty} \phi[\sigma_k(t)] \underset{a.e.}{=} \phi[x(t)] = y(t). \tag{4.2.21}$$

As we showed in equation (3.2.8), $\phi[\sigma_k(t)]$ for finite k contains only integer combinations of the basis frequencies $\lambda_1, \lambda_2, \ldots, \lambda_M$; taking $k \to \infty$ simply generates all possible integer combinations of these basis frequencies.

Now, we know from the uniqueness theorem that $\lim_{k\to\infty} \phi[\sigma_k(t)]$ and $y(t)$ must have the same Fourier series, because they are equal almost everywhere. Specifically, these two functions must contain the same frequencies in their Fourier

series, namely, all integer combinations of the basis frequencies λ_1, λ_2,..., λ_M. Thus, $x(t)$ and $y(t)$ have the same basis frequencies and the statement is proven.

CHAPTER 5
Uniform Distribution of Phase

We have now arrived at the most important part of our work. In this chapter, we will prove the following theorem: Let $\mu_1, \mu_2, \ldots, \mu_M$ be real numbers that are independent over the integers, and let $\langle \cdot \rangle$ denote the "fractional part." Then the vector values $(\langle \mu_1 t \rangle, \langle \mu_2 t \rangle, \ldots, \langle \mu_M t \rangle)$ are uniformly distributed in an M–dimensional unit cube as t ranges from 0 to ∞.

We will employ this theorem in the next chapter to convert the expression for the Fourier series coefficients of an almost periodic function from a single integral over an infinite interval into a multiple integral over finite limits. The numbers $\mu_m t$ of this theorem will, in the next chapter, be the "normalized" (Hertzian) phases corresponding to independent frequencies of complex exponentials. Consequently, we will call our work here a proof of the uniform distribution of phase for these independent frequencies.

5.1 One–Dimensional Case

We will first prove our result for the special case of $M = 1$. This approach will establish basic notation and concepts that carry over easily to the general case in the next section.

Uniform Distribution for a Discrete Variable

We begin by defining the concept of uniform distribution for a set of points:

Definition 5.1.1: Let $\psi_1, \psi_2, \ldots, \psi_N$ be a sequence of real numbers satisfying $0 \le \psi_n < 1$ for $1 \le n \le N$. Let $[\alpha,\beta)$ be any subinterval of $[0,1)$, and denote by $F_N(\psi_n,\alpha,\beta)$ the number of ψ_n that lie in $[\alpha,\beta)$. Then the sequence $\{\psi_n\}$ is said to be uniformly distributed in $[0,1)$ if and only if

$$\lim_{N\to\infty} \frac{F_N(\psi_n, \alpha,\beta)}{N} = \beta - \alpha \tag{5.1.1}$$

for every possible subinterval $[\alpha,\beta)$. ∆

This definition clearly coincides with our intuitive understanding of uniform distribution, for it says that as N becomes large, each subinterval of $[0,1)$ will contain the correct fraction of the total number of points.

For any real number γ, we will use the notation

$$\langle\gamma\rangle \equiv \text{(fractional part of } \gamma\text{)}$$

$$= \gamma - \text{(greatest integer not exceeding } \gamma\text{)}.$$

We may now state the following theorem from a branch of mathematics called Diophantine Approximation:

Theorem 5.1.1: Let μ be an arbitrary irrational number. Then the sequence $\{\psi_n = \langle\mu n\rangle,\ n=1,2,\ldots,N\}$ becomes dense and uniformly distributed in the unit interval $[0,1)$ as $N \to \infty$. Further, this uniform distribution holds uniformly throughout the interval $[0,1)$ in the following sense: Let $[\alpha,\alpha+\eta)$ be a subinterval of $[0,1)$ having length η. Then given an arbitrary $\Delta > 0$ and a minimum subinterval length η_0, we can choose an integer N_0 such that

$$\left| \frac{F_N(\psi_n,\alpha,\alpha+\eta)}{N} - \eta \right| < \Delta \tag{5.1.2}$$

holds for all $N \geq N_0$ and for all possible subintervals of length $\eta \geq \eta_0$. Specifically, the inequality can be made independent of α.

Proof: See Niven (1963), pp. 24–27 and Cassels (1957), pp. 60–65. ∆

By rearranging equation (5.1.2), we deduce that the maximum number of points in the subinterval $[\alpha,\alpha+\eta)$ is equal to $(\eta + \Delta)N$, and the minimum number of points is equal to $(\eta - \Delta)N$, for any N.

It is clear that augmenting the infinite sequence $\{\psi_n\}$ with the number zero will not affect its properties of denseness or uniform distribution, so for convenience we may begin the sequence with $n = 0$ rather than $n = 1$.

Following the approach of Cassels (1957), we may also measure uniformity in another, more general way. To do so, we must define a new counting function.

Definition 5.1.2: Let $\{\gamma_n,\ n=0,1,...,N\}$ be a sequence of real numbers satisfying $-\infty < \gamma_n < \infty$ for $0 \leq n \leq N$. Let A be any real number, and let B be a real number satisfying $0 < B - A < 1$. As before, $F_N(\gamma_n,A,B)$ will denote the number of values γ_n that lie in the interval $[A,B)$. Then

$$F_N^*(\gamma_n,A,B) \equiv \sum_{k=-\infty}^{\infty} F_N(\gamma_n,A+k,B+k). \ \Delta \tag{5.1.3}$$

The function F_N^* thus counts values of γ_n on an infinite number of intervals $[A+k,\ B+k)$ as k ranges over all integers. To understand the usefulness of this function, let us consider the sequence of Theorem 5.1.1, namely $\{\psi_n = \langle \mu n \rangle,\ n=0,1,...,N\}$, and a modified sequence $\{\psi_n^* = \mu n,\ n=0,1,...,N\}$, in which we have deleted the "fractional part" operator. If we again take $[\alpha,\alpha+\eta)$ to be any subinterval of $[0,1)$, then clearly

$$F_N^*(\psi_n^*,\alpha,\alpha+\eta) = F_N(\psi_n,\alpha,\alpha+\eta), \tag{5.1.4}$$

for we have simply transferred the "fractional part" operation from the values μn

onto the measure of these values F_N^*. In other words, in the case of $F_N(\psi_n,\alpha,\alpha+\eta)$, we translate the numbers μn into the unit interval [0,1) to count them, while in the case of $F_N^*(\psi_n^*,\alpha,\alpha+\eta)$, we translate the counting function into all unit intervals $[0+k, 1+k)$ containing numbers μn. Translating the counting function rather than the numbers will simplify the notation considerably in our upcoming work.

This new approach also allows us to state a more general form of Theorem 5.1.1, the form actually proven by Cassels, using the modified sequence $\{\psi_n^*\}$. In keeping with our earlier remark, we will begin the sequence with $n = 0$.

Theorem 5.1.2: Let μ be an arbitrary irrational number, and consider the sequence $\{\psi_n^* = \mu n,\ n=0,1,\ldots,N\}$. Let $[\alpha,\alpha+\eta)$ be any interval on the real axis having length η, $0 < \eta < 1$. Then given an arbitrary $\Delta > 0$ and a minimum interval length η_0, we can choose an integer N_0 such that

$$\left| \frac{F_N^*(\psi_n^*,\alpha,\alpha+\eta)}{N+1} - \eta \right| < \Delta \tag{5.1.5}$$

holds for all $N \geq N_0$ and for all intervals of length $\eta \geq \eta_0$, independent of α. ∆

This result is clearly more powerful than Theorem 5.1.1. It tells us that the numbers $\{\psi_n^*\}$, translated by integer values to lie in any given unit interval, are uniformly distributed. Theorem 5.1.1 is a special case in which the numbers have been translated into the interval [0,1). Thus, measuring uniformity of the sequence $\{\psi_n\}$ on a subinterval $[\alpha,\alpha+\eta)$ of [0,1) is equivalent to measuring uniformity of the modified sequence $\{\psi_n^*\}$ on any interval of length η on the real axis.

An immediate consequence of this general result is that an interval $[\alpha,\beta)$ satisfying Theorem 5.1.1 may include 0 and 1 as interior points. In other words, when graphed on a linear ψ axis, the interval may have two components $[\alpha,1)$, $[0,\beta)$ with $\eta = (1-\alpha) + (\beta-0)$. When graphed on a linear ψ^* axis, however, the interval properly appears as continuous and of length η.

The use of ψ^* instead of ψ often provides better insight into a problem because of the property just illustrated, namely, that it represents mathematical continuity by geometrical continuity. Another representation that is sometimes used for the same reason is one in which ψ represents the circumferential displacement around a circle with unit perimeter. In this case, the points 0 and 1 are geometrically the same, and again, the interval mentioned above would appear to be continuous.

With this ψ^* formulation and Theorem 5.1.2, we may also prove the following theorem:

Theorem 5.1.3: Let τ be any real number. Then the sequence $\{\psi_n = \langle\mu(n+\tau)\rangle, n=0,1,...,N\}$ becomes uniformly distributed in the sense of Theorem 5.1.1 as $N \to \infty$.

Proof: We define $\phi = \mu\tau$ and write the sequence as $\{\psi_n = \langle\mu n + \phi\rangle\}$. By our earlier remarks, it is sufficient to prove uniformity of the modified sequence $\{\psi_n^* = \mu n + \phi\}$ over any unit interval of the real axis; hence, we consider $F_N^*(\psi_n^*, \alpha, \alpha+\eta)$ for arbitrary α, and some η satisfying $0 < \eta < 1$. By a simple change of variable,

$$F_N^*(\psi_n^*, \alpha, \alpha+\eta) = F_N^*(\psi_n^* - \phi,\ \alpha-\phi,\ \alpha+\eta-\phi). \tag{5.1.6}$$

However, the right–hand side of equation (5.1.6) obeys Theorem 5.1.2, since

$$\psi_n^* - \phi = \mu n \tag{5.1.7a}$$

and

$$(\alpha + \eta - \phi) - (\alpha - \phi) = \eta, \tag{5.1.7b}$$

so the left–hand side must as well. Thus, the sequence $\{\psi_n^*\}$ is uniformly distributed in the sense of Theorem 5.1.2, and the sequence $\{\psi_n\}$ is uniformly distributed in the sense of Theorem 5.1.1. ∆

Uniform Distribution for a Continuous Variable

We have just considered a sequence $\{\langle\mu(n+\tau)\rangle,\ n=0,1,...\}$ and found that the numbers in this sequence are uniformly distributed in the interval [0,1). If we had set $t = n + \tau$, then t would have been a discrete variable (that is, one that takes on discrete values). We now generalize to the case in which t is a continuous variable; we will examine the expression $\langle\mu t\rangle$, $0 \le t < \infty$, with the goal of proving that the numbers resulting from it are also uniformly distributed in the interval [0,1).

We must first generalize the concept of uniform distribution. The measure F_N for a discrete variable now becomes an integral over the density function[1] of a continuous variable:

Definition 5.1.3: Let ψ be a variable that may take on either discrete or continuous values in the interval [0,1), and let $p(\psi)$ be the density function of ψ. Then ψ is said to be uniformly distributed over [0,1) if and only if

$$\int_{\alpha}^{\beta} p(\psi)\, d\psi = \beta - \alpha \tag{5.1.8}$$

for all possible subintervals $[\alpha,\beta) \subset [0,1)$. ∆

We may apply this definition to the infinite sequence $\{\psi_n=\langle\mu(n+\tau)\rangle,\ n=0,1,...\}$ studied earlier by taking

$$\begin{aligned} p(\psi) &= \lim_{N\to\infty} p_N(\psi) \\ &= \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} \delta[\ \psi - \langle\mu(n+\tau)\rangle\ ], \end{aligned} \tag{5.1.9}$$

---

1 By "density function" we mean simply a function having the characteristics of a probability density function from random variable theory. We avoid the use of the word "probability" since the variables here are not considered random quantities. The function $p(\gamma)$ is a valid density function of the variable γ if $p(\gamma) \ge 0$ for all γ and if the integral of $p(\gamma)$ over all γ is unity.

where $\delta(\cdot)$ is the Dirac delta (impulse) function. It is easy to verify equation (5.1.9) using Definition 5.1.3 and Theorem 5.1.1, since for any $[\alpha,\beta) \subset [0,1)$,

$$\begin{aligned}\int_{\alpha}^{\beta} p(\psi)\, d\psi &= \int_{\alpha}^{\beta} \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} \delta[\, \psi - \langle \mu(n+\tau) \rangle\,]\, d\psi \\ &= \lim_{N\to\infty} \frac{1}{N+1} \int_{\alpha}^{\beta} \sum_{n=0}^{N} \delta[\, \psi - \langle \mu(n+\tau) \rangle\,]\, d\psi \\ &= \lim_{N\to\infty} \frac{1}{N+1} F_N(\psi_n, \alpha, \beta) \\ &= \beta - \alpha. \end{aligned} \tag{5.1.10}$$

When we interchanged integration with limiting in the second step of equation (5.1.10), we assumed uniform convergence of the infinite series. We do indeed have this uniform convergence in the sense of generalized functions or distributions, since the δ functions are uniformly distributed on the ψ axis in accordance with Theorem 5.1.1.

We can also extend the concept of the measure function F_N^* in Definition 5.1.2 to the case of a continuous variable:

Definition 5.1.4: Let γ be a real variable with $-\infty < \gamma < \infty$, let A be any real number, and let B be a real number satisfying $0 < B - A < 1$. Then

$$\int_{A}^{B*} p(\gamma)\, d\gamma \equiv \sum_{k=-\infty}^{\infty} \int_{A+k}^{B+k} p(\gamma)\, d\gamma. \quad \Delta \tag{5.1.11}$$

If we set $\psi^* = \mu t$, and take $[\alpha,\beta)$ to be any subinterval of $[0,1)$, it follows from Definition 5.1.4 that

$$\int_{\alpha}^{\beta} p(\psi)\, d\psi = \int_{\alpha}^{\beta*} p(\psi^*)\, d\psi^*. \tag{5.1.12}$$

Again, we have transferred the "fractional part" operation from ψ onto the measure of uniformity $\int^*$. If we replace the right-hand side of equation (5.1.12) with the

definition of $\int^*$ in equation (5.1.11), we deduce that

$$\int_{\alpha}^{\beta} p(\psi)\, d\psi = \sum_{k=-\infty}^{\infty} \int_{\alpha+k}^{\beta+k} p(\psi^*)\, d\psi^*$$

$$= \int_{\alpha}^{\beta} \sum_{k=-\infty}^{\infty} p(\psi^* + k)\, d\psi^*. \qquad (5.1.13)$$

Equation (5.1.13) shows that the density function of ψ on the unit interval [0,1) is found by superimposing the components of the density function of ψ^*, these components coming from the infinite sequence of unit intervals $[0+k,\ 1+k)$ over which ψ^* may exist.

We are now able to begin our proof of the uniform distribution of $\langle \mu t \rangle$ for continuous t. We might observe that this proof can be accomplished by very elementary calculations; we choose a less direct approach here with the aim of generalizing it to higher dimensions later. This approach requires that we first work an intermediate problem, after which the desired result will follow easily.

Instead of considering a sequence of points $\{t = n+\tau,\ n=0,1,...,N\}$ on the t axis, as we did in the discrete variable case, let us consider a sequence of intervals

$$\{\ (n+\tau) \le t < (n+\tau+\ \epsilon/\mu),\ \ n=0,1,...,N\ \},$$

where μ is any positive, irrational number, and τ is any positive real number. We restrict ϵ to be a real number satisfying $0 < \epsilon << 1$ and $\epsilon/\mu << 1$, since ultimately we will let ϵ approach zero. We take $\psi = \langle \mu t \rangle$ and consider the distribution of ψ on the unit interval [0,1). We will show that this distribution becomes arbitrarily close to uniform for sufficiently large N and sufficiently small ϵ.

We will find it convenient to have explicit expressions for the density functions of t and ψ. Unfortunately, the latter would be rather cumbersome due to the wrap–around effect induced by the fractional part operation. For this reason, we

will resort to the ψ^* formulation and find $p(\psi^*)$ instead. To derive the density functions, we define a function

$$\Pi(\gamma,\epsilon) \equiv \begin{cases} 1/\epsilon, & 0 \le \gamma < \epsilon \\ 0, & \text{elsewhere.} \end{cases} \tag{5.1.14}$$

We assume that t is uniformly distributed within and among the intervals $[n+\tau, n+\tau+\epsilon/\mu)$, so that the density function of the finite sequence of intervals may be written as

$$p_N(t) = \frac{1}{N+1} \sum_{n=0}^{N} \Pi[t - (n+\tau), \epsilon/\mu], \quad 0 \le t < \infty. \tag{5.1.15}$$

This function is graphed in Figure 5.1.1a. Using a well-known property of density functions [see Papoulis (1965), pp. 126-127], we may write the density of $\psi^* = \mu t$ as

$$p_N(\psi^*) = \frac{1}{N+1} \sum_{n=0}^{N} \Pi[\psi^* - \mu(n+\tau), \epsilon], \quad 0 \le \psi^* < \infty, \tag{5.1.16}$$

which is graphed in Figure 5.1.1b.

Let us now measure the uniformity of ψ^* on an interval $[\alpha,\alpha+\eta)$ of the real axis, by computing upper and lower bounds on the expression

$$\begin{aligned} \int_{\alpha}^{\alpha+\eta}{}^{*} p_N(\psi^*)\, d\psi^* &= \int_{\alpha}^{\alpha+\eta}{}^{*} \frac{1}{N+1} \sum_{n=0}^{N} \Pi[\psi^* - \mu(n+\tau), \epsilon]\, d\psi^* \\ &= \int_{\alpha}^{\alpha+\eta} \sum_{k=-\infty}^{\infty} \frac{1}{N+1} \sum_{n=0}^{N} \Pi[\psi^* - k - \mu(n+\tau), \epsilon]\, d\psi^* \end{aligned} \tag{5.1.17}$$

for arbitrary N and α, and some η satisfying $0 < \eta < 1$.

The second form of equation (5.1.17) expresses our alternate view of the $\int^*$ operator [as in equation (5.1.13)], in which the integration is taken over a single interval $[\alpha,\alpha+\eta)$, and the integrand is an infinite sum of density functions $p_N(\psi^*)$ shifted by integer amounts.

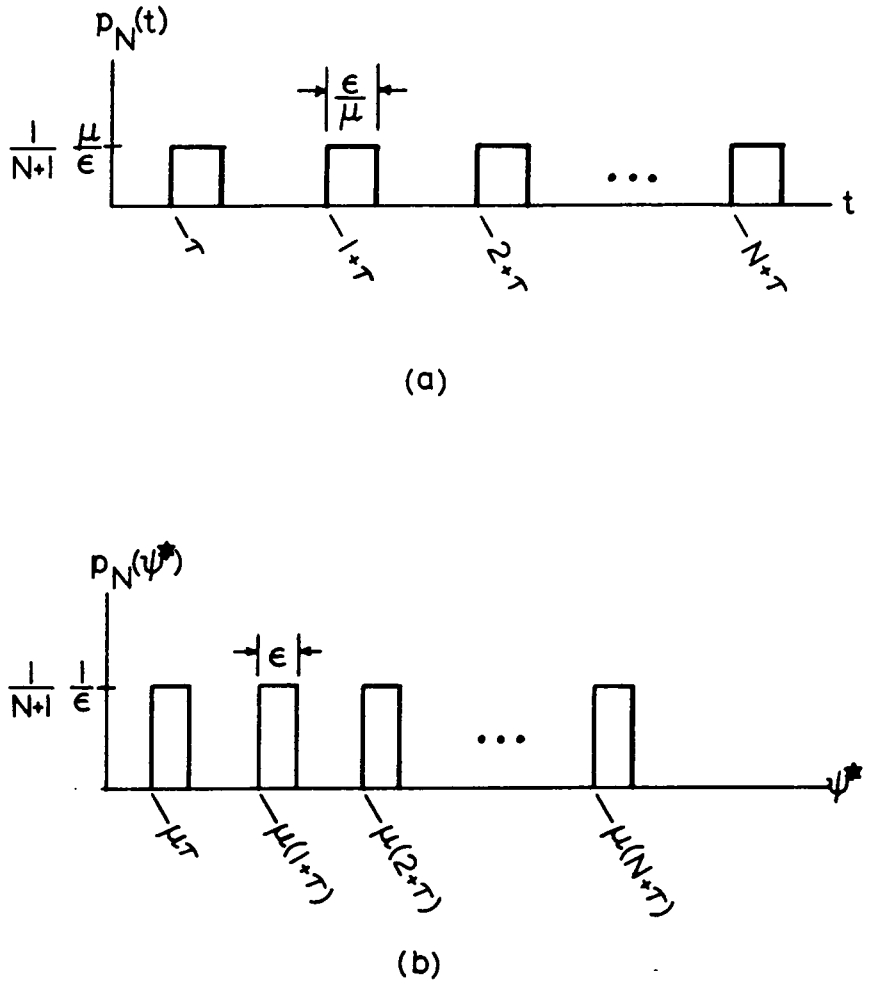

Figure 5.1.1. (a) Density function of t. (b) Density function of $\psi^* = \mu t$.

Pursuing this viewpoint, we have sketched in Figure 5.1.2 a pictorial representation of the integration process. This figure shows the region of integration $[\alpha,\alpha+\eta)$, which we will call region A, containing Π functions or portions of Π functions that have been translated from all intervals $[\alpha+k,\ \alpha+\eta+k)$ on the ψ^* axis. The initial points (rising edges) of these Π functions originally lay at values $\psi^* = \mu(n+\tau)$; consequently, the initial points are distributed in accordance with Theorem 5.2.1. The figure also shows a region $[\alpha-\epsilon,\alpha)$, called region B, that is external to A, and a region $[\alpha+\eta-\epsilon,\ \alpha+\eta)$, called region C, that lies at the upper end of A.

Figure 5.1.2 further shows two worst–case distributions of Π functions, corresponding to the upper and lower bounds on the integral of equation (5.1.17). In case 1, the maximum number of initial points allowed by Theorem 5.1.2 fall within region A. Further, the count of Π functions will be even higher because all of the Π functions with initial points in region B have intruded into A, and the integration in equation (5.1.17) will encompass these intruding Π functions. In case 2, the minimum number of initial points allowed by Theorem 5.1.2 fall within region A. Further, the count of Π functions will be even lower because all of the Π functions with initial points in region C have spilled out of the right end of A, and the integration of equation (5.1.17) will not encompass these Π functions. We now quantify these observations.

Let Δ be the maximum error associated with uniformity of the initial points, on any region of the ψ^* axis having a length of $\eta - \epsilon$. That is, from Theorem 5.1.2, Δ will be a number satisfying

$$\left| \left[\int_{\alpha}^{\alpha+(\eta-\epsilon)*} \frac{1}{N+1} \sum_{n=0}^{N} \delta[\ \psi^* - \mu(n+\tau)\ ]\ d\psi^* \right] - (\eta - \epsilon) \right| < \Delta, \qquad (5.1.18)$$

independent of α. We know from this theorem that equation (5.1.18) will be valid if

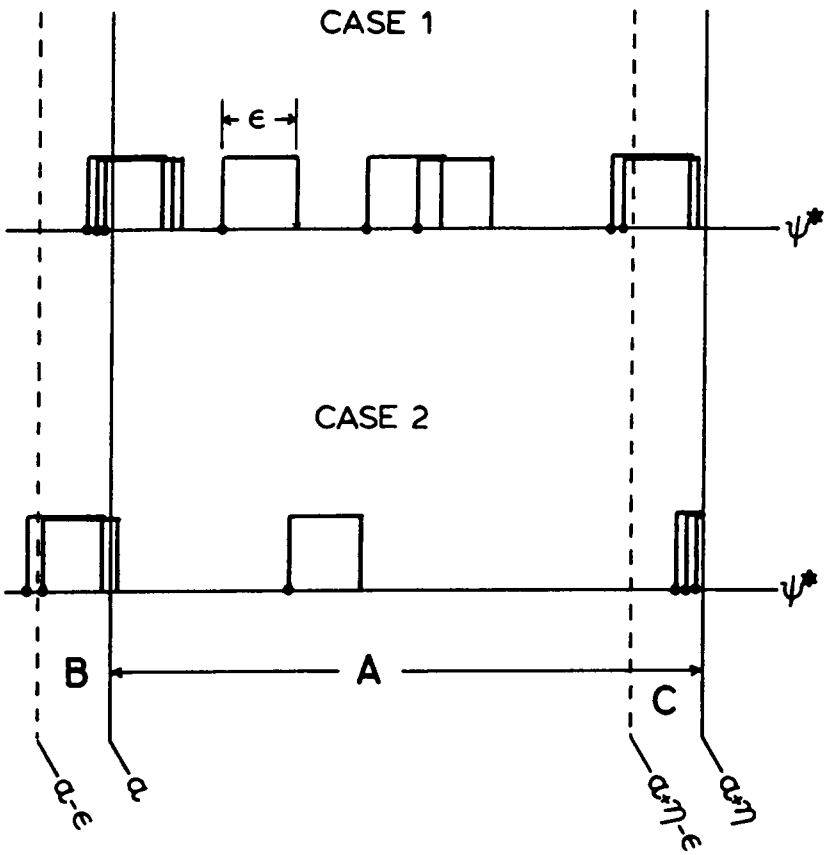

Figure 5.1.2. Representation of the integration process in equation (5.1.17). Case 1 is the upper bound on the integral; case 2 is the lower bound.

we replace $\eta - \epsilon$ with any length L satisfying $\eta - \epsilon \leq L < 1$. Since Δ here depends on both η and ϵ, we will choose it based on an initial value of ϵ; afterwards, we will let $\epsilon \to 0$, so this error bound will still hold.

If we now denote the maximum number of initial points in region A by $M(A)$ and the minimum number by $m(A)$, we find from equation (5.1.18) and the remarks following Theorem 5.1.1 that

$$M(A) = (\eta + \Delta)(N + 1) \tag{5.1.19a}$$

$$m(A) = (\eta - \Delta)(N + 1). \tag{5.1.19b}$$

Using similar notation, we also find that

$$M(A+B) = (\eta + \epsilon + \Delta)(N + 1) \tag{5.1.20}$$

$$m(A-C) = (\eta - \epsilon - \Delta)(N + 1). \tag{5.1.21}$$

Strictly speaking, we must require $\eta + \epsilon < 1$ to do the computation of equation (5.1.20); however, we may substitute our weaker conditions $\eta < 1$ and $\epsilon << 1$ with the recognition that the error bound $M(A+B)$ will simply be larger than the true least upper bound $(1 + \Delta)(N + 1)$ whenever $1 \leq \eta + \epsilon < 1 + \epsilon$.

Using these results, we now have

$$\begin{aligned} M(B) &= M(A+B) - m(A) \\ &= [\eta + \epsilon + \Delta - (\eta - \Delta)](N + 1) \\ &= (\epsilon + 2\Delta)(N + 1) \end{aligned} \tag{5.1.22}$$

and

$$\begin{aligned} M(C) &= M(A) - m(A-C) \\ &= [\eta + \Delta - (\eta - \epsilon - \Delta)](N + 1) \\ &= (\epsilon + 2\Delta)(N + 1). \end{aligned} \tag{5.1.23}$$

We note that the two calculations of equations (5.1.22) and (5.1.23) could not have been done directly since we do not have an error bound associated with a

region of length ϵ.

We may now compute

$$\text{maximum } \#\,\Pi \text{ functions counted in } A = M(A) + M(B)$$
$$= (\eta + \Delta + \epsilon + 2\Delta)(N+1)$$
$$= (\eta + \epsilon + 3\Delta)(N+1) \tag{5.1.24}$$

and

$$\text{minimum } \#\,\Pi \text{ functions counted in } A = m(A) - M(C)$$
$$= [\eta - \Delta - (\epsilon + 2\Delta)](N+1)$$
$$= (\eta - \epsilon - 3\Delta)(N+1). \tag{5.1.25}$$

Consequently, the desired error bound on equation (5.1.17) is

$$\left| \left[\int_{\alpha}^{\alpha+\eta *} \frac{1}{N+1} \sum_{n=0}^{N} \Pi[\psi^* - \mu(n+\tau),\, \epsilon]\, d\psi^* \right] - \eta \right| < \epsilon + 3\Delta, \tag{5.1.26}$$

so that

$$\left| \int_{\alpha}^{\alpha+\eta *} p_N(\psi^*)\, d\psi^* - \eta \right| < \epsilon + 3\Delta, \tag{5.1.27}$$

where Δ depends on η and N, but not on α. This error bound is valid if we replace η with any length L satisfying $\eta \le L < 1$.

Let us stop for a moment to review the foregoing work. We began by constructing a sequence of intervals on the t axis,

$$\{\, (n+\tau) \le t < (n+\tau+ \epsilon/\mu),\ n=0,1,\ldots,N \,\},$$

whose density function $p_N(t)$ is given by equation (5.1.15). After taking $\psi^* = \mu t$, we obtained a sequence of intervals on the ψ^* axis,

$$\{\, \mu(n+\tau) \le \psi^* < \mu(n+\tau)+\epsilon,\ n=0,1,\ldots,N \,\},$$

whose density function $p_N(\psi^*)$ is given by equation (5.1.16). By applying the $\int^*$ measure of uniformity from Definition 5.1.4, we concluded that the sequence of ψ^*

intervals is uniformly distributed over any region having length η, within an error given by equation (5.1.27). This error depends both on the width of the original intervals on the t axis, namely ϵ/μ, and on the number of such intervals in the sequence, namely $N + 1$.

Continuing with our proof, we now wish to take a union of interval sequences on the t axis, in such a way that the entire axis is covered from 0 to $N + 1$. From equation (5.1.15), we see that we may accomplish this goal in the following way: We choose ϵ such that $\epsilon/\mu = 1/K$ for some integer K, then construct K sequences of intervals with corresponding density functions $p_N^{(0)}(t)$, $p_N^{(1)}(t)$,..., $p_N^{(K-1)}(t)$; sequence 0 will have $\tau = 0$, sequence 1 will have $\tau = 1/K$,..., and sequence K-1 will have $\tau = (K\text{-}1)/K$.

Because each sequence contains the same number of intervals, namely $N + 1$, the density function of the union of sequences is just the average of the individual densities, that is,

$$p_N^{(union)}(t) = \frac{1}{K}\sum_{k=0}^{K-1} p_N^{(k)}(t) = \begin{cases} \frac{1}{N+1}, & 0 \le t < N+1 \\ 0, & \text{elsewhere.} \end{cases} \tag{5.1.28}$$

(This last equality may be verified using Figure 5.1.1a along with the fact that $1/K = \epsilon/\mu$.) Thus, this union of sequences causes t to be uniformly distributed in the interval $[0,N+1)$.

Each sequence of intervals on the t axis, having density function $p_N^{(k)}(t)$, generates a sequence of intervals on the ψ^* axis, having density function $p_N^{(k)}(\psi^*)$, for $0 \le k \le K\text{-}1$. In turn, the union of sequences on the t axis, having density function $p_N^{(union)}(t)$, generates a union of sequences on the ψ^* axis, having density function

$$p_N^{(union)}(\psi^*) = \frac{1}{K} \sum_{k=0}^{K-1} p_N^{(k)}(\psi^*). \tag{5.1.29}$$

Since each $p_N^{(k)}(\psi^*)$ obeys equation (5.1.27), the average $p_N^{(union)}(\psi^*)$ must as well. That is,

$$\left| \int_{\alpha}^{\alpha+\eta}{}^* p_N^{(union)}(\psi^*)\, d\psi^* - \eta \right| < \epsilon + 3\Delta. \tag{5.1.30}$$

Let us choose an arbitrary $\eta > 0$, and examine equation (5.1.30). We can make ϵ arbitrarily small by making K sufficiently large, that is, by using enough sequences in equation (5.1.28), because by definition $\epsilon = \mu/K$. In accordance with Theorem 5.1.2, we can also make Δ arbitrarily small by making N sufficiently large, that is, by using long enough sequences in equation (5.1.28). Further, we may make ϵ and Δ arbitrarily small for any $\eta > 0$. Consequently, the density function $p_N^{(union)}(\psi^*)$ becomes uniform as $K \to \infty$, $N \to \infty$.

Thus, the continuous variable $\psi^* = \mu t$, $0 \leq t < \infty$, has uniform density in the sense of $\int^*$ measure on any unit interval of the ψ^* axis, and $\psi = \langle \mu t \rangle$ has uniform density on [0,1).

This proof has been done under the assumption that μ is positive and irrational, but the result is also true for any real number μ. The restriction that μ be positive was made strictly for convenience, and it is easy to show that the foregoing proof holds for μ negative. To prove that μ need not be irrational, we suppose that μ is, in fact, rational. Then we can choose a number b such that $\mu/b = \mu'$ is irrational. We define $t' = bt$ and apply our results to the problem $\psi = \langle \mu' t' \rangle$. If $0 \leq t < \infty$, then $0 \leq t' < \infty$, and $\langle \mu' t' \rangle$ is uniformly distributed. Consequently, so is $\langle \mu t \rangle$.

5.2 Multi–Dimensional Case

We now turn to the case of $M > 1$; that is, we study the properties of values in M–space. Most of our results from the one–dimensional case generalize in an obvious way. We begin, as before, by considering discrete variables.

Uniform Distribution for Discrete Variables

We will let $\overline{\psi} = (\psi_1,\psi_2,\ldots,\psi_M)$ denote either a point in M–space, a vector–valued variable, or a vector in M–space. (The specific meaning will be clear from context, or will be stated explicitly.) The usual notation for vector arithmetic will also apply, so that, for example, $\overline{\psi} > \overline{0}$ means $\psi_m > 0$ for each $1 \le m \le M$.

In analogy with Definition 5.1.1, we may define the concept of uniform distribution for a set of points in M–space:

Definition 5.2.1: Let $\overline{\psi}^{(1)}, \overline{\psi}^{(2)},\ldots, \overline{\psi}^{(n)}$ be a sequence of points in M–space satisfying $\overline{0} \le \overline{\psi}^{(n)} < \overline{1}$ for $1 \le n \le N$. Let $[\overline{\alpha},\overline{\beta})$ be any partition of the unit cube $[\overline{0},\overline{1})$, that is, any region satisfying $\overline{0} \le \overline{\alpha} < \overline{\beta} < \overline{1}$. Denote by $F_N(\overline{\psi}^{(n)},\overline{\alpha},\overline{\beta})$ the number of $\overline{\psi}^{(n)}$ that lie in the partition $[\overline{\alpha},\overline{\beta})$. Then the sequence $\{\overline{\psi}^{(n)}\}$ is said to be uniformly distributed in the unit cube $[\overline{0},\overline{1})$ if and only if

$$\lim_{N\to\infty} \frac{F_N(\overline{\psi}^{(n)},\overline{\alpha},\overline{\beta})}{N} = (\beta_1-\alpha_1)(\beta_2-\alpha_2)\cdots(\beta_M-\alpha_M) \qquad (5.2.1)$$

for every possible partition $[\overline{\alpha},\overline{\beta})$. ∆

We will apply the "fractional part" operator to any point $\overline{\psi}$ on a component basis, so that

$$\langle\overline{\psi}\rangle \equiv (\langle\psi_1\rangle,\langle\psi_2\rangle,\ldots,\langle\psi_M\rangle).$$

Using it, we state the extension to Theorem 5.1.1 as follows:

Theorem 5.2.1: Let $\overline{\mu} = (\mu_1, \mu_2, \ldots, \mu_M)$ be a vector of arbitrary irrational numbers that are independent over the integers. Then the sequence of points $\{\overline{\psi}^{(n)} = \langle \overline{\mu} n \rangle, n=1,2,\ldots,N\}$ becomes dense and uniformly distributed in the unit cube $[\overline{0},\overline{1})$ as $N \to \infty$. Further, this uniform distribution holds uniformly throughout the unit cube in the following sense: Let $[\overline{\alpha},\overline{\alpha}+\overline{\eta})$ be a partition of the unit cube. Then given an arbitrary $\Delta > 0$, and a minimum partition dimension $\overline{\eta}^0$ satisfying $\overline{0} < \overline{\eta}^0 < \overline{1}$, we can choose an integer N_0 such that

$$\left| \frac{F_N(\overline{\psi}^{(n)}, \overline{\alpha},\overline{\alpha}+\overline{\eta})}{N} - \eta_1 \eta_2 \cdots \eta_M \right| < \Delta \qquad (5.2.2)$$

holds for all $N \geq N_0$ and for all possible volumes having $\overline{\eta} \geq \overline{\eta}^0$. Specifically, the inequality can be made independent of $\overline{\alpha}$. ∆

We may rearrange equation (5.2.2) to deduce that the maximum number of points in $[\overline{\alpha},\overline{\alpha}+\overline{\eta})$ is equal to $(\eta_1\eta_2\cdots\eta_M + \Delta)N$, and the minimum number is equal to $(\eta_1\eta_2\cdots\eta_M - \Delta)N$.

Again, we may augment the infinite sequence of points $\{\overline{\psi}^{(n)}\}$ with the point $\overline{0}$, without affecting the properties of denseness or uniform distribution.

As in the one–dimensional case, there is an alternate approach that incorporates the fractional part operation into the measure of the sequence, rather than the sequence itself. For multiple dimensions, this approach will prove quite useful in avoiding a great deal of tedious notation.

Definition 5.2.2: Let $\{\overline{\gamma}^{(n)}, n=0,1,\ldots,N\}$ be a sequence of points in M–space having $-\overline{\infty} < \overline{\gamma}^{(n)} < \overline{\infty}$, let $\overline{A}$ be any point in M–space, and let point $\overline{B}$ satisfy $\overline{0} < \overline{B} - \overline{A} < \overline{1}$. As before, we denote by $F_N(\overline{\gamma}^{(n)},\overline{A},\overline{B})$ the number of points of the sequence $\{\overline{\gamma}^{(n)}\}$ that lie in the partition of M–space $[\overline{A},\overline{B})$. Then we define the function $F^*_N(\overline{\gamma}^{(n)},\overline{A},\overline{B})$ as

$$F_N^*(\bar{\gamma}^{(n)},\bar{A},\bar{B}) \equiv \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_M=-\infty}^{\infty} F_N(\bar{\gamma}^{(n)},\bar{A}+\bar{k},\bar{B}+\bar{k}). \; \Delta \tag{5.2.3}$$

Thus, F_N^* counts points that lie in an M–fold sequence of volumes in M–space. If we now consider the modified sequence $\{\bar{\psi}^{*(n)} = \bar{\mu}n,\ n=0,1,...,N\}$, with $[\bar{\alpha},\bar{\beta})$ again any partition of the unit cube $[\bar{0},\bar{1})$, then

$$F_N^*(\bar{\psi}^{*(n)},\bar{\alpha},\bar{\beta}) = F_N(\bar{\psi}^{(n)},\bar{\alpha},\bar{\beta}). \tag{5.2.4}$$

For convenience, we take the summation in equation (5.2.3) over all unit cubes; however, for the case of F_N^* applied to $\bar{\psi}^{*(n)}$, the summands will be non–zero only in the sequence of unit cubes that contains the ray $\bar{\mu}t$, $0 \le t < \infty$.

Using Definition 5.2.2, we may now state the multi–dimensional extension of Theorem 5.1.2, which gives us a more general result for measuring uniformity than does Theorem 5.2.1.

Theorem 5.2.2: Let $\bar{\mu}$ be a vector of arbitrary irrational numbers that are independent over the integers, and consider the sequence of points $\{\bar{\psi}^{*(n)} = \bar{\mu}n,\ n=0,1,...,N\}$. Let $[\bar{\alpha},\bar{\alpha}+\bar{\eta})$ be any partition of M–space having dimension $\bar{\eta}$, with $\bar{0} < \bar{\eta} < \bar{1}$. Then given an arbitrary $\Delta > 0$, and a minimum partition dimension $\bar{\eta}^0$ satisfying $\bar{0} < \bar{\eta}^0 < \bar{1}$, we can choose an integer N_0 such that

$$\left| \frac{F_N(\bar{\psi}^{*(n)},\bar{\alpha},\bar{\alpha}+\bar{\eta})}{N+1} - \eta_1 \eta_2 \cdots \eta_M \right| < \Delta \tag{5.2.5}$$

holds for all $N \ge N_0$ and for all possible partitions having $\bar{\eta} \ge \bar{\eta}^0$. Specifically, the inequality can be made independent of $\bar{\alpha}$. Δ

Theorem 5.2.2 thus allows us to measure uniformity of the modified sequence $\{\psi^{*(n)} = \bar{\mu}n\}$ over any region of dimension $\bar{\eta}$ in M–space, as an alternative to measuring uniformity of the sequence $\{\psi^{(n)} = \langle\bar{\mu}n\rangle\}$ over a region of dimension $\bar{\eta}$ in $[\bar{0},\bar{1})$.

In one dimension, we were able to visualize the "fractional part" operation by wrapping the ψ axis around a circle of unit perimeter. In multiple dimensions, the circle becomes an M–dimensional torus, which is quite abstract except for the case $M = 2$. Consequently, we will have to content ourselves with the understanding that opposite faces of our M–dimensional unit cube are mathematically, if not geometrically, adjacent.

With these results, the M–dimensional extension to Theorem 5.1.3 may be stated. Its proof is an obvious generalization of the one–dimensional case.

Theorem 5.2.3: Let τ be any real number. Then the sequence $\{\overline{\psi}^{(n)} = \langle\overline{\mu}(n+\tau)\rangle, n=0,1,...\}$ is uniformly distributed in the sense of Theorem 5.2.1. ∆

Uniform Distribution for Continuous Variables

We will now examine the expression $\langle\overline{\mu}t\rangle$, $0 \le t < \infty$, with the goal of proving that the numbers resulting from it are also uniformly distributed in the unit cube. Our measure of uniformity will be an M–dimensional integral over the joint density function of the M variables $\psi_1, \psi_2, ..., \psi_M$ We formalize this statement in the following definition:

Definition 5.2.3: Let $\overline{\psi} = (\psi_1,\psi_2,...,\psi_M)$ be a vector variable that may take on either discrete or continuous values in the M–dimensional unit cube $[\overline{0},\overline{1})$. Let $p(\overline{\psi})$ be the joint density function of $\psi_1, \psi_2, ..., \psi_M$ Then $\overline{\psi}$ is said to be uniformly distributed over the unit cube if and only if

$$\int_{\alpha_M}^{\beta_M} \cdots \int_{\alpha_2}^{\beta_2} \int_{\alpha_1}^{\beta_1} p(\overline{\psi})\, d\overline{\psi} = (\beta_1-\alpha_1)(\beta_2-\alpha_2) \cdots (\beta_M-\alpha_M) \qquad (5.2.6)$$

for all possible partitions $[\overline{\alpha},\overline{\beta})$ of $[\overline{0},\overline{1})$. Here, $d\overline{\psi} \equiv d\psi_1\, d\psi_2 \cdots d\psi_M$ ∆

As an example, the density function of $\overline{\psi} = \langle \overline{\mu} t \rangle$ for the discrete variable case $t = n + \tau$ is

$$p(\overline{\psi}) = \lim_{N\to\infty} p_N(\overline{\psi})$$

$$= \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} \delta[\, \psi_1 - \langle \mu_1(n+\tau) \rangle\,]\ \delta[\, \psi_2 - \langle \mu_2(n+\tau) \rangle\,] \cdots$$

$$\delta[\, \psi_M - \langle \mu_M(n+\tau) \rangle\,]. \tag{5.2.7}$$

We can also extend the alternate uniformity measure of Definition 5.2.2 to the case of a continuous variable in M–space:

Definition 5.2.4: Let $\overline{\gamma}$ be a vector variable having $-\overline{\infty} < \overline{\gamma} < \overline{\infty}$, let $\overline{A}$ be any point in M–space, and let $\overline{B}$ be a point in M–space satisfying $\overline{0} < \overline{B} - \overline{A} < \overline{1}$. Then

$$\int_{\overline{A}}^{\overline{B}\,*} p(\overline{\gamma})\, d\overline{\gamma} \equiv \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_M=-\infty}^{\infty} \int_{A_M+k_M}^{B_M+k_M} \cdots$$

$$\int_{A_2+k_2}^{B_2+k_2} \int_{A_1+k_1}^{B_1+k_1} p(\gamma_1,\gamma_2,\ldots,\gamma_M)\, d\gamma_1\, d\gamma_2 \cdots d\gamma_M\, \Delta \tag{5.2.8}$$

If we set $\overline{\psi}^* = \overline{\mu} t$, and take $[\overline{\alpha},\overline{\beta})$ to be any partition of the unit cube $[\overline{0},\overline{1})$,

$$\int_{\alpha_M}^{\beta_M} \cdots \int_{\alpha_2}^{\beta_2} \int_{\alpha_1}^{\beta_1} p(\overline{\psi})\, d\overline{\psi} = \int_{\overline{\alpha}}^{\overline{\beta}\,*} p(\overline{\psi}^*)\, d\overline{\psi}^*. \tag{5.2.9}$$

Again, we have transferred the "fractional part" operation from $\overline{\psi}$ onto the measure of uniformity $\int^*$. By substituting equation (5.2.8) into the right–hand side of equation (5.2.9), interchanging integration with summation, and making a change of variables, we can obtain the multi–dimensional extension of equation (5.1.13). From it, we conclude that the density function of $\overline{\psi}$ on the unit cube $[\overline{0},\overline{1})$ is found by superimposing the components of the density function of $\overline{\psi}^*$, these components coming from the M–fold infinity of unit cubes over which $\overline{\psi}^*$ may exist.

We now begin our proof of the uniform distribution of $\langle \overline{\mu} t \rangle$ for continuous t. In contrast with the one–dimensional case, the result is not trivial, and we must resort to the sort of proof used there. The difference is the following: In the one–dimensional case, the unit interval $\psi \subset [0,1)$ becomes continuous with values $\psi = \langle \mu t \rangle$ for continuous t; in the multi–dimensional case, the unit volume $\overline{\psi} \subset [\overline{0},\overline{1})$ becomes dense with values $\overline{\psi} = \langle \overline{\mu} t \rangle$, but not continuous.

To begin, let $\overline{\mu}$ be a vector of positive, irrational numbers that are independent over the integers, and let us define the magnitude of $\overline{\mu}$ as

$$\mu \equiv |\overline{\mu}| = [\, \mu_1^2 + \mu_2^2 + \cdots + \mu_M^2 \,]^{1/2}. \tag{5.2.10}$$

For later use, we will define the vector constant $\overline{\epsilon}$ as

$$\overline{\epsilon} = (\epsilon_1, \epsilon_2, \ldots, \epsilon_M) \tag{5.2.11a}$$

with magnitude

$$\epsilon \equiv |\overline{\epsilon}|. \tag{5.2.11b}$$

We will restrict the vector $\overline{\epsilon}$ to lie in the same direction as $\overline{\mu}$, so that

$$\overline{\epsilon} = \epsilon \cdot \overline{\mu}/\mu. \tag{5.2.12}$$

As before, we consider a sequence of intervals on the t axis,

$$\{\, (n+\tau) \le t < (n+\tau+ \epsilon/\mu),\ n=0,1,\ldots,N \,\},$$

in which ϵ, defined in equation (5.2.11), satisfies $0 < \epsilon << 1$ and $\epsilon/\mu << 1$; ultimately, we will let ϵ approach zero.

We may express the density function of t exactly as we did for the one–dimensional case. As before, we define the function

$$\Pi(t,\epsilon) = \begin{cases} 1/\epsilon, & 0 \le t < \epsilon \\ 0, & \text{elsewhere} \end{cases} \tag{5.2.13}$$

so that

$$p_N(t) = \frac{1}{N+1} \sum_{n=0}^{N} \Pi[t - (n+\tau),\ \epsilon/\mu],\ \ 0 \le t < \infty. \tag{5.2.14}$$

We now take $\overline{\psi}^* = \overline{\mu}t$. Corresponding to an interval of length ϵ/μ on the t axis is a line segment of length ϵ in M–dimensional $\overline{\psi}^*$ space; this line segment lies in the direction given by the unit vectors $\overline{\mu}/\mu = \overline{\epsilon}/\epsilon$. Further, a point t_0 on the t axis corresponds to a point $\overline{\mu}t_0$ in $\overline{\psi}^*$ space. From these observations, we may express $p_N(\overline{\psi}^*)$ by making an extension of the Π function to higher dimensions. Our multi–dimensional function $\Pi(\overline{\psi}^*, \overline{\epsilon})$ is to represent a line segment extending from $\overline{0}$ to $\overline{\epsilon}$ in M–space. The line $\overline{\epsilon}t$ containing this segment may be defined by the intersection of $M - 1$ planes, each containing the line. Clearly, the planes

$$\frac{\psi_1^*}{\epsilon_1} - \frac{\psi_M^*}{\epsilon_M} = 0; \quad \frac{\psi_2^*}{\epsilon_2} - \frac{\psi_M^*}{\epsilon_M} = 0; \quad \cdots \quad ; \frac{\psi_{M-1}^*}{\epsilon_{M-1}} - \frac{\psi_M^*}{\epsilon_M} = 0 \tag{5.2.15}$$

fulfill the requirements. Their density functions have the form

$$\delta[\psi_1^* - \epsilon_1\psi_M^*/\epsilon_M]; \quad \delta[\psi_2^* - \epsilon_2\psi_M^*/\epsilon_M]; \quad \cdots \quad ; \delta[\psi_{M-1}^* - \epsilon_{M-1}\psi_M^*/\epsilon_M] \tag{5.2.16}$$

and the density function of the line $\overline{\epsilon}t$ contains a product of these individual densities. Enforcing the requirement that the M–fold integral over $\Pi(\overline{\psi}, \overline{\epsilon})$ must be unity, we deduce that

$$\Pi(\overline{\psi}^*, \overline{\epsilon}) = \begin{cases} \delta[\psi_1^* - \epsilon_1\psi_M^*/\epsilon_M]\,\delta[\psi_2^* - \epsilon_2\psi_M^*/\epsilon_M] \cdots \delta[\psi_{M-1}^* - \epsilon_{M-1}\psi_M^*/\epsilon_M]\,/\epsilon_M, \\ \quad \overline{0} \le \overline{\psi}^* < \overline{\epsilon} \\ 0, \quad \text{elsewhere}. \end{cases} \tag{5.2.17}$$

We note that this particular choice of planes was arbitrary, and that any permutation of the ψ_m^* and ϵ_m in equation (5.2.17) yields an equivalent result.

To construct $p_N(\overline{\psi}^*)$, we must now locate Π functions in space with their initial points at $\overline{\mu}(n+\tau)$, $n = 0, 1, \ldots, N$. Thus,

$$p_N(\overline{\psi}^*) = \frac{1}{N+1} \sum_{n=0}^{N} \Pi[\overline{\psi}^* - \overline{\mu}(n+\tau), \overline{\epsilon}], \quad 0 \le \psi^* < \infty. \tag{5.2.18}$$

Let us now measure the uniformity of $\overline{\psi}^*$ on a partition of M–space $[\overline{\alpha}, \overline{\alpha}+\overline{\eta})$ by computing upper and lower bounds on

$$\int_{\overline{\alpha}}^{\overline{\alpha}+\overline{\eta}} {}^{*}\, p_N(\overline{\psi}^*)\, d\overline{\psi}^* = \int_{\overline{\alpha}}^{\overline{\alpha}+\overline{\eta}} {}^{*}\, \frac{1}{N+1} \sum_{n=0}^{N} \Pi[\overline{\psi}^* - \overline{\mu}(n+\tau), \overline{\epsilon}]\, d\overline{\psi}^* \tag{5.2.19}$$

for arbitrary N and $\overline{\alpha}$, and some $\overline{\eta}$ satisfying $\overline{0} < \overline{\eta} < \overline{1}$. To help visualize the procedure, we have sketched in Figure 5.2.1 a typical situation for $M = 2$. The region of interest, which we will again call region A, has dimensions $\eta_1 \times \eta_2$ and an arbitrary "initial corner" $\overline{\alpha}$. In accordance with the definition of the $\int^*$ operator in equation (5.2.8), this region contains Π functions (representing line segments in 2–space) that have been translated from all regions $[\overline{\alpha}+\overline{k}, \overline{\alpha}+\overline{\eta}+\overline{k})$, where $\overline{k} = (k_1, k_2)$ ranges over all pairs of integers. The initial points of these Π functions originally lay at values $\overline{\psi}^* = \overline{\mu}(n+\tau)$; consequently, the initial points are distributed in accordance with Theorem 5.2.2. The figure also shows a region B that is external to A, and a region C that lies inside of A.

In Figure 5.2.1, we have further shown two worst–case distributions of Π functions, namely case 1 and case 2, corresponding to upper and lower bounds on the integral of equation (5.2.19). In case 1, the maximum number of initial points allowed by Theorem 5.2.2 fall within region A. Further, the count of Π functions will be even higher because all of the Π functions with initial points in region B have intruded into A, and will enter into the integration of equation (5.2.19). In case 2, the minimum number of initial points allowed by Theorem 5.2.2 fall within region A. Further, the count of Π functions will be even lower because all of the Π functions with initial points in region C have spilled out of the boundary of A, and will not enter into the integral of equation (5.2.19). We notice that a Π function, having length ϵ in two–space, has projection ϵ_1 in the ψ_1^* direction and projection ϵ_2

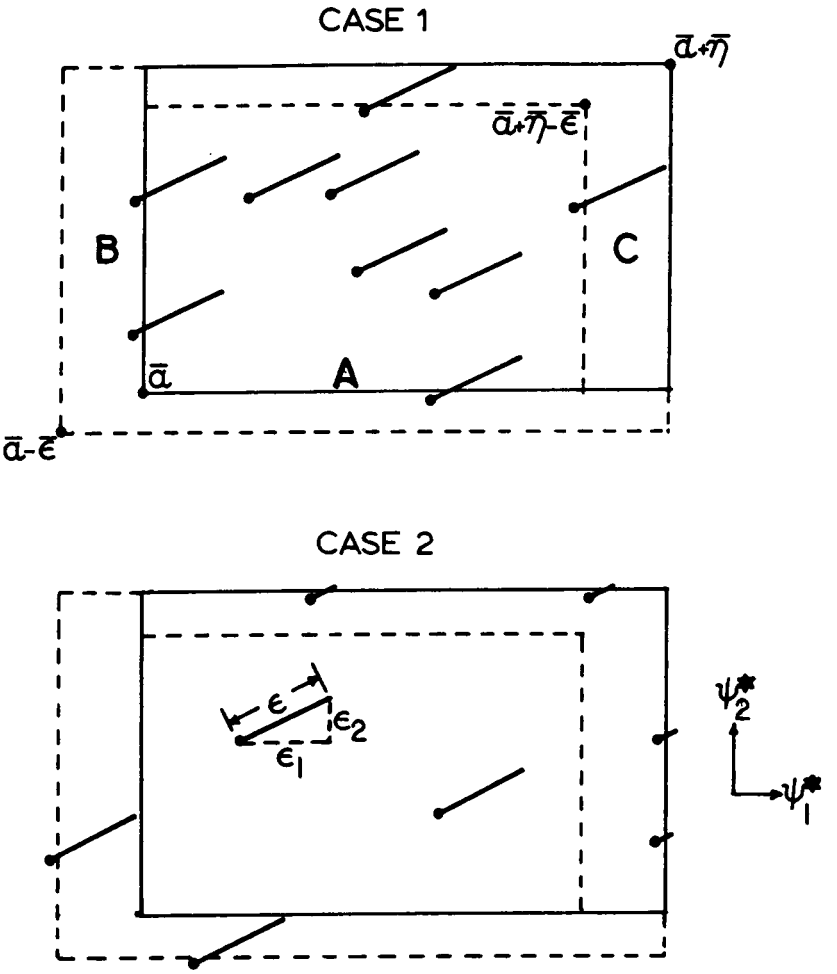

Figure 5.2.1. Representation of the integration process in equation (5.2.19), for $M = 2$. Case 1 is the upper bound on the integral; case 2 is the lower bound.

in the ψ_2^* direction.

Let us return now to the general case of $M > 1$, and take Δ to be the maximum error associated with uniformity of the initial points in a region of dimension $\bar{\eta} - \bar{\epsilon}$. From Theorem 5.2.2, Δ will be a number satisfying

$$\left| \left[\int_{\bar{\alpha}}^{\bar{\alpha}+(\bar{\eta}-\bar{\epsilon})*} \frac{1}{N+1} \sum_{n=0}^{N} \delta[\, \psi_1^* - \langle \mu_1(n+\tau) \rangle\,]\, \delta[\, \psi_2^* - \langle \mu_2(n+\tau) \rangle\,] \cdots \right.\right.$$

$$\left.\left. \delta[\, \psi_M^* - \langle \mu_M(n+\tau) \rangle\,]\, d\bar{\psi}^* \right]$$

$$- (\eta_1 - \epsilon_1)(\eta_2 - \epsilon_2) \cdots (\eta_M - \epsilon_M) \Bigg| < \Delta. \qquad (5.2.20)$$

We know from Theorem 5.2.2 that equation (5.2.20) will remain valid if, for any m, we replace $\eta_m - \epsilon_m$ with a length L_m satisfying $\eta_m - \epsilon_m \leq L_m < 1$. Since Δ depends on both $\bar{\eta}$ and $\bar{\epsilon}$, we will choose it based on an initial value of $\bar{\epsilon}$. Later, when $\epsilon = |\bar{\epsilon}|$ approaches zero, this error bound will still hold.

If we denote the maximum number of initial points in region A by $M(A)$ and the minimum number by $m(A)$, we find that

$$M(A) = (\eta_1 \eta_2 \cdots \eta_M + \Delta)\,(N+1) \qquad (5.2.21a)$$

$$m(A) = (\eta_1 \eta_2 \cdots \eta_M - \Delta)\,(N+1), \qquad (5.2.21b)$$

and that

$$M(A+B) = [(\eta_1+\epsilon_1)(\eta_2+\epsilon_2)\cdots(\eta_M+\epsilon_M) + \Delta]\,(N+1) \qquad (5.2.22)$$

$$m(A-C) = [(\eta_1-\epsilon_1)(\eta_2-\epsilon_2)\cdots(\eta_M-\epsilon_M) - \Delta]\,(N+1). \qquad (5.2.23)$$

We may simplify these expressions somewhat, under the assumption that ϵ_m/η_m is small for each m. Taking equation (5.2.22), for example, we may write

$$
\begin{aligned}
M(A+B) &= [(\eta_1+\epsilon_1)(\eta_2+\epsilon_2)\cdots(\eta_M+\epsilon_M) + \Delta]\,(N+1) \\
&= \eta_1\eta_2\cdots\eta_M\left[\left[1+\frac{\epsilon_1}{\eta_1}\right]\left[1+\frac{\epsilon_2}{\eta_2}\right]\cdots\left[1+\frac{\epsilon_M}{\eta_M}\right]+\frac{\Delta}{\eta_1\eta_2\cdots\eta_M}\right](N+1) \\
&= \eta_1\eta_2\cdots\eta_M\left[1+\frac{\epsilon_1}{\eta_1}+\frac{\epsilon_2}{\eta_2}+\cdots\right. \\
&\qquad\left.+\frac{\epsilon_M}{\eta_M}+O(\epsilon_m/\eta_m)^2+\frac{\Delta}{\eta_1\eta_2\cdots\eta_M}\right](N+1) \\
&\simeq \eta_1\eta_2\cdots\eta_M\left[1+\frac{\epsilon_1}{\eta_1}+\frac{\epsilon_2}{\eta_2}+\cdots+\frac{\epsilon_M}{\eta_M}+\frac{\Delta}{\eta_1\eta_2\cdots\eta_M}\right](N+1).
\end{aligned}
\tag{5.2.24}
$$

Throughout our remaining work, we will retain only first-order error terms such as these, since they dominate the error behavior.

Using the results leading to equation (5.2.24), we may also express equations (5.2.21a), (5.2.21b), and (5.2.23) as

$$
M(A) = \eta_1\eta_2\cdots\eta_M\left[1+\frac{\Delta}{\eta_1\eta_2\cdots\eta_M}\right](N+1), \tag{5.2.25}
$$

$$
m(A) = \eta_1\eta_2\cdots\eta_M\left[1-\frac{\Delta}{\eta_1\eta_2\cdots\eta_M}\right](N+1), \tag{5.2.26}
$$

and

$$
m(A-C) = \eta_1\eta_2\cdots\eta_M\left[1-\frac{\epsilon_1}{\eta_1}-\frac{\epsilon_2}{\eta_2}-\cdots-\frac{\epsilon_M}{\eta_M}-\frac{\Delta}{\eta_1\eta_2\cdots\eta_M}\right](N+1). \tag{5.2.27}
$$

Consequently, we have

$$
\begin{aligned}
M(B) &= M(A+B) - m(A) \\
&= \eta_1\eta_2\cdots\eta_M\left[\frac{\epsilon_1}{\eta_1}+\frac{\epsilon_2}{\eta_2}+\cdots+\frac{\epsilon_M}{\eta_M}+\frac{2\Delta}{\eta_1\eta_2\cdots\eta_M}\right](N+1)
\end{aligned}
\tag{5.2.28}
$$

and

$$
\begin{aligned}
M(C) &= M(A) - m(A-C) \\
&= \eta_1\eta_2\cdots\eta_M\left[\frac{\epsilon_1}{\eta_1}+\frac{\epsilon_2}{\eta_2}+\cdots+\frac{\epsilon_M}{\eta_M}+\frac{2\Delta}{\eta_1\eta_2\cdots\eta_M}\right](N+1).
\end{aligned}
\tag{5.2.29}
$$

The two calculations in equations (5.2.28) and (5.2.29) could not have been done directly since we do not have an error bound associated with a region of dimension less than $\bar{\eta} - \bar{\epsilon}$.

We may now compute

maximum # $\sqcap$ functions counted in $A = M(A) + M(B)$

$$= \eta_1\eta_2\cdots\eta_M\left[\, 1 + \frac{\epsilon_1}{\eta_1} + \frac{\epsilon_2}{\eta_2} + \cdots + \frac{\epsilon_M}{\eta_M} + \frac{3\Delta}{\eta_1\eta_2\cdots\eta_M} \,\right](N+1), \quad (5.2.30)$$

and

minimum # $\sqcap$ functions counted in $A = m(A) - M(C)$

$$= \eta_1\eta_2\cdots\eta_M\left[\, 1 - \frac{\epsilon_1}{\eta_1} - \frac{\epsilon_2}{\eta_2} - \cdots - \frac{\epsilon_M}{\eta_M} - \frac{3\Delta}{\eta_1\eta_2\cdots\eta_M} \,\right](N+1). \quad (5.2.31)$$

Finally, the desired error bound on equation (5.2.19) is

$$\left| \left[\int_{\underline{\alpha}}^{\bar{\alpha}+\bar{\eta}\,*} \frac{1}{N+1} \sum_{n=0}^{N} \sqcap \left[\, \bar{\psi}^* - \bar{\mu}(n+\tau),\, \bar{\epsilon} \,\right] d\bar{\psi}^* \right] - \eta_1\eta_2\cdots\eta_M \right|$$

$$< \eta_1\eta_2\cdots\eta_M\left[\frac{\epsilon_1}{\eta_1} + \frac{\epsilon_2}{\eta_2} + \cdots + \frac{\epsilon_M}{\eta_M} + \frac{3\Delta}{\eta_1\eta_2\cdots\eta_M} \right] \quad (5.2.32)$$

so that

$$\left| \int_{\underline{\alpha}}^{\bar{\alpha}+\bar{\eta}\,*} p_N(\bar{\psi}^*)\, d\bar{\psi}^* - \eta_1\eta_2\cdots\eta_M \right|$$

$$< \eta_1\eta_2\cdots\eta_M\left[\frac{\epsilon_1}{\eta_1} + \frac{\epsilon_2}{\eta_2} + \cdots + \frac{\epsilon_M}{\eta_M} + \frac{3\Delta}{\eta_1\eta_2\cdots\eta_M} \right]. \quad (5.2.33)$$

As before, we now take a finite union of $p_N(t)$ in such a way that the entire t axis is covered from 0 to $N + 1$. We do it by choosing an $\epsilon = |\bar{\epsilon}|$ such that $\epsilon/\mu = 1/K$ for some integer K, then constructing a sequence of intervals with $\tau = 0$, another with $\tau = 1/K$,..., and finally one with $\tau = (K\text{-}1)/K$. Each sequence of intervals on the t axis will have a corresponding sequence of line segments in M-space, according to the function $\bar{\psi}^* = \bar{\mu}t$. The density functions $p_N(\bar{\psi}^*)$ of these

latter sequences will be uniform, within the error given by equation (5.2.33), in any region of M-space having dimension $\bar{\eta}$ or larger. The density function of the union of sequences is simply the average of the individual densities, and the error of this union thus is also bounded by equation (5.2.33).

Now for any $\bar{\eta} > 0$, we can make ϵ arbitrarily small by making K sufficiently large, that is, by using enough sequences of intervals on the t axis. Hence, we can make $\max_m \epsilon_m$ arbitrarily small. For this same $\bar{\eta}$, we can also make Δ arbitrarily small by making N sufficiently large, that is, by using long enough sequences of intervals on the t axis. Consequently, the density function of the union of sequences in $\bar{\psi}^*$ space becomes uniform as $K \to \infty$, $N \to \infty$. Thus, $\bar{\psi}^* = \bar{\mu}t$, $0 \le t < \infty$, has uniform density in the sense of $\int^*$ measure on any unit cube, and $\bar{\psi} = \langle \bar{\mu}t \rangle$ has uniform density on $[\bar{0},\bar{1})$.

For our work in Chapter 6, we will also state a special form of equation (5.2.33). Suppose that we take N to be finite, so that each sequence of intervals on the t axis is of finite length, but that we let $K \to \infty$, so that an infinite number of sequences are used. Then $\epsilon \to 0$, and equation (5.2.33) reduces to

$$\left| \int_{\bar{\alpha}}^{\bar{\alpha}+\bar{\eta}} {}^* p_N(\bar{\psi}^*)\, d\bar{\psi}^* - \eta_1\eta_2\cdots\eta_M \right| < \eta_1\eta_2\cdots\eta_M \left[\frac{3\Delta}{\eta_1\eta_2\cdots\eta_M} \right] = 3\Delta \quad (5.2.34)$$

or equivalently, for any $[\bar{\alpha},\bar{\alpha}+\bar{\eta})$ in $[\bar{0},\bar{1})$,

$$\left| \int_{\bar{\alpha}}^{\bar{\alpha}+\bar{\eta}} p_N(\bar{\psi})\, d\bar{\psi} - \eta_1\eta_2\cdots\eta_M \right| < 3\Delta. \quad (5.2.35)$$

The foregoing proof has been carried out under the restriction that the μ_m are all positive and irrational, as well as independent over the integers. The results remain true for any μ_m that are independent over the integers, which clearly allows

at most one of the μ_m to be rational. The proof follows that for the one–dimensional case.

CHAPTER 6
Multiple Integral Formulation of the Fourier Series Coefficients

We are now in a position to prove the main result of our work, namely the multiple integral formulation of the Fourier series coefficients. We will begin by stating the problem precisely, to establish notation and restrictions. Next, we will carry out the proof in two ways, one an intuitive, superficial approach, the other, a more rigorous approach. Finally, we will comment on a few theoretical implications of the multiple integral form.

6.1 Problem Statement

Let $x(t)$ be a real–valued *S.a.p.* (Stepanoff almost periodic) function, as defined in Chapter 2, satisfying $a < x(t) < b$ for all t. Let $\phi(x)$ be a bounded, continuous, real–valued function of x, defined over the domain $a \leq x \leq b$. We know that $y(t) = \phi[x(t)]$ is also *S.a.p.*, and thus has a Fourier series. As we saw in Chapter 4, we may compute the Fourier series coefficients for $y(t)$ from the formula $a(\lambda) = M\{y(t)\, e^{-j\lambda t}\}$. The mean value $M\{\cdot\}$ is taken over the infinite interval $0 \leq t < \infty$, and consequently this expression cannot be computed in practice. Our goal in this chapter is to show that, for certain types of $x(t)$, the mean value over infinite t is equal to a multiple integration over finite limits. Our first task is to specify the allowable forms for the input $x(t)$.

In Chapters 3 and 4 we used the concept of a basis for the frequencies in an *S.a.p.* function. In general, the basis can contain enumerably infinite frequencies

that are independent over the rational numbers. For practical computations, however, we must restrict our attention to *S.a.p.* functions having finite bases; the members of these bases are independent over the integers. As we proved in Chapter 3, restricting the nonlinearity input $x(t)$ to have a particular finite basis will ensure that the output $y(t)$ has the same basis. Let us denote the basis frequencies for $x(t)$ and $y(t)$ as $\lambda_1, \lambda_2, \ldots, \lambda_M$

We must further restrict $x(t)$ to allow practical computation. Let $P_q(t)$ denote a purely periodic function of t, with a period of 2π. $P_q(t)$ could be a single sinusoid (sine or cosine), a finite sum of harmonically related sinusoids, or even an infinite sum of such sinusoids. In this last case, $P_q(t)$ will be known through the sum of its Fourier series, that is, as a function of t. We then require that $x(t)$ be expressible in the form

$$x(t) = \sum_{q=1}^{Q} P_q(\lambda^{(q)}t), \tag{6.1.1}$$

where

$$\lambda^{(q)} = v_{1q}\lambda_1 + v_{2q}\lambda_2 + \cdots + v_{Mq}\lambda_M \tag{6.1.2}$$

and the v_{mq} are integers. Thus, each $P_q(\lambda^{(q)}t)$ is periodic, with a radian frequency $\lambda^{(q)}$ that is a linear combination of the basis frequencies.

Let us rewrite each of the functions $P_q(\lambda^{(q)}t)$ as

$$P_q(\lambda^{(q)}t) = P_q\left\{2\pi\left[v_{1q}\frac{\lambda_1 t}{2\pi} + v_{2q}\frac{\lambda_2 t}{2\pi} + \cdots + v_{Mq}\frac{\lambda_M t}{2\pi}\right]\right\} \tag{6.1.3}$$

and make the definition

$$\frac{\lambda_m t}{2\pi} \equiv \theta_m = \theta_m(t), \tag{6.1.4}$$

so that

$$P_q(\lambda^{(q)}t) = P_q\left\{2\pi\left[v_{1q}\theta_1 + v_{2q}\theta_2 + \cdots + v_{Mq}\theta_M\right]\right\}. \tag{6.1.5}$$

We will call the θ_m "normalized phases," since they involve the Hertzian frequencies

$\lambda_m/(2\pi)$ rather than the radian frequencies λ_m. Because $P_q(t)$ is assumed periodic with period 2π, it is clearly permissible to replace each normalized phase with its fractional part, so that

$$P_q(\lambda^{(q)}t) = P_q\left\{2\pi\left[v_{1q}\langle\theta_1\rangle + v_{2q}\langle\theta_2\rangle + \cdots + v_{Mq}\langle\theta_M\rangle\right]\right\}. \tag{6.1.6}$$

Equation (6.1.6) shows that $P_q(\lambda^{(q)}t)$ is periodic, with a period of unity, with respect to each of its normalized phases.

Applying these results to each term of equation (6.1.1), we may express $x(t)$ as a function of the M normalized phases θ_1, θ_2,..., θ_M. Denoting this composite function by $\bar{x}$, we may write

$$x(t) = \bar{x}(\theta_1,\theta_2,\ldots,\theta_M) \equiv \bar{x}(\theta). \tag{6.1.7}$$

Likewise, the output $y(t)$ of the memoryless nonlinearity can be written as a function of the normalized phases. Calling this composite function $\bar{y}$, we have

$$y(t) = \bar{y}(\theta_1,\theta_2,\ldots,\theta_M) \equiv \bar{y}(\theta). \tag{6.1.8}$$

Clearly, both $\bar{x}(\theta)$ and $\bar{y}(\theta)$ are also periodic in each of their normalized phases. From Chapter 5, we know that the fractional parts of these normalized phases become uniformly distributed in the unit cube as t takes on all values from zero to infinity. We will make use of these observations in the next section.

Let us now consider the structural properties of $P_q(t)$. If $P_q(t)$ is a single sinusoid or a finite sum of sinusoids, it is a uniformly continuous function of t. If, on the other hand, it is an infinite sum of sinusoids, it need not be continuous. We will restrict a non–continuous $P_q(t)$ to be bounded and of bounded variation, so that in one period it has at most a finite number of finite jump discontinuities. That is, $P_q(t)$ will be discontinuous for at most a finite number of values of t in the interval $[0,2\pi)$. Let us suppose that there are K of these values, say $t = 2\pi\alpha_1, 2\pi\alpha_2,\ldots, 2\pi\alpha_K$, where $0 \le \alpha_k < 1$ for $1 \le k \le K$. From the periodicity of $P_q(t)$, it follows that $P_q(t)$ is

discontinuous for all values of t satisfying

$$t = 2\pi(\alpha_1 + i_1),\ 2\pi(\alpha_2 + i_2),\ldots,\ 2\pi(\alpha_K + i_K), \tag{6.1.9}$$

where each i_k takes on all integer values. Consequently, $P_q(\lambda^{(q)}t)$ is discontinuous for values of its argument satisfying

$$\lambda^{(q)}t = 2\pi(\alpha_1 + i_1),\ 2\pi(\alpha_2 + i_2),\ldots,\ 2\pi(\alpha_K + i_K). \tag{6.1.10}$$

Using equation (6.1.5), we thus conclude that the discontinuities occur for values of $\theta_1, \theta_2,\ldots, \theta_M$ satisfying

$$\begin{aligned} &2\pi(v_{1q}\theta_1 + v_{2q}\theta_2 + \cdots + v_{Mq}\theta_M) \\ &\qquad = 2\pi(\alpha_1 + i_1),\ 2\pi(\alpha_2 + i_2),\ldots,\ 2\pi(\alpha_K + i_K) \end{aligned} \tag{6.1.11}$$

and

$$\begin{aligned} &v_{1q}\theta_1 + v_{2q}\theta_2 + \cdots + v_{Mq}\theta_M \\ &\qquad = (\alpha_1 + i_1),\ (\alpha_2 + i_2),\ldots,\ (\alpha_K + i_K). \end{aligned} \tag{6.1.12}$$

Each of the equalities in equation (6.1.12) represents a plane in the M-dimensional space whose coordinate axes are $\theta_1, \theta_2,\ldots, \theta_M$ (we will refer to this space as the "phase space" of the corresponding function of t). Any finite partition of phase space will intersect at most a finite number of these planes; hence, $P_q(\lambda^{(q)}t)$ will be continuous throughout any unit cube in its phase space except on a finite number of planes intersecting this cube. Since $x(t)$ may contain at most a finite sum of such $P_q(\lambda^{(q)}t)$, we conclude that $x(t)$ is also continuous throughout any unit cube in its phase space except on a finite number of planes. (We will say that $\bar{x}(\bar{\theta})$ is almost everywhere continuous.) Finally, because of the uniform continuity of the nonlinear function $\phi(x)$, these continuity properties carry over to $\bar{y}(\bar{\theta})$ as well.

6.2 Proof of the Multiple Integral Formulation

We are now prepared to prove the main result of our work, first in a somewhat heuristic manner that will illustrate the basic concepts, and second, in a rigorous manner.

Expected Value Approach

As we learned in Chapter 4, the Fourier series coefficients of the *S.a.p.* function $y(t)$ are found by computing

$$a(\lambda) = M\{y(t)\, e^{-j\lambda t}\}$$

$$= \lim_{T\to\infty} \frac{1}{T} \int_0^T y(t)\, e^{-j\lambda t}\, dt. \tag{6.2.1}$$

The coefficients $a(\lambda)$ are non–zero for a set of frequencies λ that are, at most, enumerably infinite. These frequencies have the same basis as do the frequencies in $x(t)$ and $y(t)$. If we define

$$g(t) = y(t)\, e^{-j\lambda t} \tag{6.2.2}$$

then we may write simply

$$a(\lambda) = M\{g(t)\}. \tag{6.2.3}$$

For another viewpoint of this expression, we borrow concepts from the theory of random variables. For any continuous random variable t, and some function $\gamma(t)$, we define the "expected value" of $\gamma(t)$ as [see Papoulis (1965), p. 142]

$$E\{\gamma(t)\} = \lim_{T\to\infty} \int_{-T}^{T} \gamma(t)\, p_T(t)\, dt, \tag{6.2.4}$$

where $p_T(t)$ is a probability density function satisfying

$$p_T(t) = 0, \quad |t| > T, \tag{6.2.5}$$

and

$$\lim_{T\to\infty} p_T(t) = p(t). \tag{6.2.6}$$

In the event that $p(t) = \lim_{T\to\infty} p_T(t)$ is non–zero, we may delete the limiting operation and write

$$E\{\gamma(t)\} = \int_{-\infty}^{\infty} \gamma(t)\, p(t)\, dt. \tag{6.2.7}$$

If we consider t to be a random variable that is uniformly distributed over $[0,\infty)$, then we may take

$$p_T(t) = \begin{cases} 1/T, & 0 \le t < T \\ 0, & \text{elsewhere,} \end{cases} \tag{6.2.8}$$

and express the Fourier series coefficients as

$$a(\lambda) = E\{g(t)\}. \tag{6.2.9}$$

From our work in Section 6.1, we know that $y(t)$ may be written in the form

$$y(t) = \bar{y}(\theta_1,\theta_2,\ldots,\theta_M) = \bar{y}(\bar{\theta}), \tag{6.2.10}$$

where, again, the $\theta_m = \lambda_m t/(2\pi)$ are normalized phases involving the basis frequencies $\{\lambda_m\}$. Since $a(\lambda)$ is non–zero only for λ having the same basis, clearly the product $y(t)\, e^{-j\lambda t} = g(t)$ can also be written in the form

$$g(t) = \bar{g}(\theta_1,\theta_2,\ldots,\theta_M) = \bar{g}(\bar{\theta}). \tag{6.2.11}$$

Because $\bar{g}$ is periodic in each of the normalized phases, we may also express equation (6.2.11) as

$$g(t) = \bar{g}[\langle\theta_1\rangle,\langle\theta_2\rangle,\ldots,\langle\theta_M\rangle] \equiv \bar{g}[\langle\bar{\theta}\rangle]. \tag{6.2.12}$$

If we treat $\bar{g}$ as a function of the M variables $\langle\theta_1\rangle$, $\langle\theta_2\rangle$,..., $\langle\theta_M\rangle$, then from an elementary result in probability theory [see Papoulis (1965), p. 239], we may write

$$E\{g(t)\} = E\{\, \bar{g}[\langle\bar{\theta}\rangle]\, \} \tag{6.2.13}$$

where, by definition,

$$E\{\, \overline{g}[\langle \overline{\theta} \rangle] \,\} \equiv \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \overline{g}[\langle \theta_1 \rangle, \langle \theta_2 \rangle, \ldots, \langle \theta_M \rangle] \, p[\langle \theta_1 \rangle, \langle \theta_2 \rangle, \ldots, \langle \theta_M \rangle] \, d\langle \theta_1 \rangle \, d\langle \theta_2 \rangle \cdots d\langle \theta_M \rangle; \tag{6.2.14}$$

the expression $p[\langle \theta_1 \rangle, \langle \theta_2 \rangle, \ldots, \langle \theta_M \rangle]$ is called the joint probability density function of the M (random) variables $\langle \theta_1 \rangle$, $\langle \theta_2 \rangle$,..., $\langle \theta_M \rangle$. From the results of Chapter 5, however, we know that

$$p[\langle \theta_1 \rangle, \langle \theta_2 \rangle, \ldots, \langle \theta_M \rangle] = \begin{cases} 1, & \overline{0} \le \langle \overline{\theta} \rangle < \overline{1} \\ 0, & \text{elsewhere;} \end{cases} \tag{6.2.15}$$

that is, the fractional parts of the normalized phases are uniformly distributed in the M–dimensional unit cube $[\overline{0},\overline{1})$. Thus,

$$E\{\, \overline{g}[\langle \overline{\theta} \rangle] \,\} = \int_{0}^{1} \cdots \int_{0}^{1} \int_{0}^{1} \overline{g}[\langle \theta_1 \rangle, \langle \theta_2 \rangle, \ldots, \langle \theta_M \rangle] \, d\langle \theta_1 \rangle \, d\langle \theta_2 \rangle \cdots d\langle \theta_M \rangle. \tag{6.2.16}$$

Replacing λ in equation (6.2.2) with its representation in terms of the basis frequencies, we find the desired form for the Fourier series coefficients as

$$\begin{aligned} a(n_1\lambda_1 + n_2\lambda_2 + \cdots + n_M\lambda_M) &= \int_{0}^{1} \cdots \int_{0}^{1} \int_{0}^{1} \phi\left[\, \overline{x}[\langle \theta_1 \rangle, \langle \theta_2 \rangle, \ldots, \langle \theta_M \rangle] \,\right] \\ &\qquad e^{-j2\pi[n_1\langle \theta_1 \rangle + n_2\langle \theta_2 \rangle + \cdots + n_M\langle \theta_M \rangle]} \, d\langle \theta_1 \rangle \, d\langle \theta_2 \rangle \cdots d\langle \theta_M \rangle \\ &= \int_{0}^{1} \cdots \int_{0}^{1} \int_{0}^{1} \phi[\, \overline{x}(\theta_1, \theta_2, \ldots, \theta_M) \,] \, e^{-j2\pi[n_1\theta_1 + n_2\theta_2 + \cdots + n_M\theta_M]} \, d\theta_1 \, d\theta_2 \cdots d\theta_M \end{aligned} \tag{6.2.17}$$

where n_1, n_2,..., n_M range over all integers independently; in the second equality, we have deleted the fractional part operators because the integration is confined to the unit cube. In our upcoming work, we will often use the simplified notation

$a(n_1,n_2,\ldots,n_M)$ to mean $a(n_1\lambda_1 + n_2\lambda_2 + \cdots + n_M\lambda_M)$ whenever the basis frequencies are understood.

Rigorous Approach

We will now prove our result using fundamental principles. Let us consider the expression

$$\frac{1}{T}\int_0^T g(t)\,dt \equiv \frac{1}{T}\int_0^T \bar{g}[\theta_1(t),\theta_2(t),\ldots,\theta_M(t)]\,dt \tag{6.2.18}$$

for finite T, where

$$\theta_m(t) = \lambda_m t/(2\pi). \tag{6.2.19}$$

This single integral is equal to the M–dimensional integral

$$W \equiv \int_0^{\lambda_M T/(2\pi)} \cdots \int_0^{\lambda_2 T/(2\pi)} \int_0^{\lambda_1 T/(2\pi)} \bar{g}(\theta_1,\theta_2,\ldots,\theta_M)$$
$$\delta[\theta_1 - \lambda_1\theta_M/\lambda_M]\,\delta[\theta_2 - \lambda_2\theta_M/\lambda_M]\cdots\delta[\theta_{M-1} - \lambda_{M-1}\theta_M/\lambda_M]$$
$$\frac{2\pi}{\lambda_M T}\,d\theta_1\,d\theta_2\cdots d\theta_M, \tag{6.2.20}$$

as we may verify by direct calculation: After performing the first integration in θ_1, we have

$$W = \int_0^{\lambda_M T/(2\pi)} \cdots \int_0^{\lambda_2 T/(2\pi)} \bar{g}(\lambda_1\theta_M/\lambda_M,\ \theta_2,\ldots,\ \theta_M)$$
$$\delta[\theta_2 - \lambda_2\theta_M/\lambda_M]\cdots\delta[\theta_{M-1} - \lambda_{M-1}\theta_M/\lambda_M]$$
$$\frac{2\pi}{\lambda_M T}\,d\theta_2\cdots d\theta_M, \tag{6.2.21}$$

and after completing the θ_2, θ_3,..., θ_{M-1} integrations, we have

$$W = \int_0^{\lambda_M T/(2\pi)} \bar{g}(\lambda_1\theta_M/\lambda_M,\ \lambda_2\theta_M/\lambda_M,\ldots,\ \theta_M)\,\frac{2\pi}{\lambda_M T}\,d\theta_M \tag{6.2.22}$$

With the change of variable

$$\gamma = \frac{2\pi\ \theta_M}{\lambda_M};\ d\gamma = \frac{2\pi\ d\theta_M}{\lambda_M}, \tag{6.2.23}$$

the integral of equation (6.2.22) becomes

$$W = \int_0^T \bar{g}[\lambda_1\gamma/(2\pi),\ \lambda_2\gamma/(2\pi),\ldots,\ \lambda_M\gamma/(2\pi)]\ \frac{1}{T}\ d\gamma, \tag{6.2.24}$$

which is equal to the integral of equation (6.2.18). The same result follows for any other order of integration. Clearly, any permutation in order among the variables $\theta_1, \theta_2,\ldots, \theta_{M-1}$ will lead to the single integral of equation (6.2.24). It is also possible to show that the integration in θ_M need not be performed last, although we will not do so here.

From our work in Chapter 5, specifically equation (5.2.17), we recognize that

$$\delta[\theta_1 - \lambda_1\theta_M/\lambda_M]\ \delta[\theta_2 - \lambda_2\theta_M/\lambda_M]\cdots\delta[\theta_{M-1} - \lambda_{M-1}\theta_M/\lambda_M]\ \frac{2\pi}{\lambda_M T}$$

in equation (6.2.20) is the density function of the ray $\bar{\theta} = \bar{\lambda}t/(2\pi)$, $0 \le t < T$, in M-dimensional $\bar{\theta}$ space. For convenience, then, let us define

$$p_T(\bar{\theta}) = \begin{cases} \delta[\theta_1 - \lambda_1\theta_M/\lambda_M]\ \delta[\theta_2 - \lambda_2\theta_M/\lambda_M]\cdots\delta[\theta_{M-1} - \lambda_{M-1}\theta_M/\lambda_M]\ \dfrac{2\pi}{\lambda_M T}, \\ \quad \bar{0} \le \bar{\theta} < \bar{\lambda}T/(2\pi) \\ 0, \quad \text{elsewhere}. \end{cases} \tag{6.2.25}$$

Using this notation, we may express equation (6.2.20) as

$$\begin{aligned} W &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{g}(\theta_1,\theta_2,\ldots,\theta_M)\ p_T(\bar{\theta})\ d\bar{\theta} \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{g}(\bar{\theta})\ p_T(\bar{\theta})\ d\bar{\theta} \end{aligned} \tag{6.2.26}$$

where, as usual, $d\bar{\theta} \equiv d\theta_1\ d\theta_2\cdots d\theta_M$

We may view the integral of equation (6.2.26) as an M–fold infinite sum of integrals over adjacent unit cubes; that is,

$$W = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_M=-\infty}^{\infty} \int_{k_M}^{k_M+1} \cdots \int_{k_2}^{k_2+1} \int_{k_1}^{k_1+1} g(\overline{\theta})\, p_T(\overline{\theta})\, d\overline{\theta}$$

$$= \int_0^1 \cdots \int_0^1 \int_0^1 \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_M=-\infty}^{\infty} g(\overline{\theta}+\overline{k})\, p_T(\overline{\theta}+\overline{k})\, d\overline{\theta}. \qquad (6.2.27)$$

Since $\overline{g}(\overline{\theta})$ is periodic in each of its arguments, with a period of unity,

$$\overline{g}(\overline{\theta}+\overline{k}) = \overline{g}(\overline{\theta}), \qquad (6.2.28)$$

so that

$$W = \int_0^1 \cdots \int_0^1 \int_0^1 \overline{g}(\overline{\theta}) \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_M=-\infty}^{\infty} p_T(\overline{\theta}+\overline{k})\, d\overline{\theta}. \qquad (6.2.29)$$

From the comments following equation (5.2.9), we recognize that

$$\sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_M=-\infty}^{\infty} p_T(\overline{\theta}+\overline{k}) = p_T[\langle\overline{\theta}\rangle] \text{ for } \overline{0} \le \overline{\theta} < \overline{1}, \qquad (6.2.30)$$

where $p_T[\langle\overline{\theta}\rangle]$ is the joint density function of the fractional parts of the normalized phases, that is, the density function of $\langle\overline{\theta}\rangle = \langle\overline{\lambda}t/(2\pi)\rangle$, $0 \le t < T$. Consequently, we may write

$$W = \int_0^1 \cdots \int_0^1 \int_0^1 \overline{g}(\overline{\theta})\, p_T[\langle\overline{\theta}\rangle]\, d\overline{\theta}. \qquad (6.2.31)$$

As we noted in Chapter 5, an explicit expression for $p_T[\langle\overline{\theta}\rangle]$ is quite cumbersome because of the fractional part operator. We do know, however, that $p_T[\langle\overline{\theta}\rangle]$ has the uniformity properties of equation (5.2.35), if we set $\langle\overline{\theta}\rangle = \overline{\psi}$ and $T = N + 1$.

We will now consider the behavior of equation (6.2.31) as T becomes infinite. Let us first divide the unit cube $[\overline{0},\overline{1})$ into a set of equal–sized partitions

(rectangular prisms) having dimension $\bar{\eta}$ and volume $\eta_1\ \eta_2 \cdots \eta_M$; let us denote the i^{th} such partition by $\bar{\eta}^{(i)}$. With each partition we associate a value $\bar{g}^{(i)}$ such that

$$\int_{\bar{\eta}^{(i)}} \bar{g}(\bar{\theta})\ p_T[\langle\bar{\theta}\rangle]\ d\bar{\theta} = \bar{g}^{(i)} \int_{\bar{\eta}^{(i)}} p_T[\langle\bar{\theta}\rangle]\ d\bar{\theta}. \tag{6.2.32}$$

If $\bar{g}(\bar{\theta})$ is continuous within the partition $\bar{\eta}^{(i)}$, then $\bar{g}(\bar{\theta}) = \bar{g}^{(i)}$ for some $\bar{\theta}$ in $\bar{\eta}^{(i)}$. If $\bar{g}(\bar{\theta})$ is discontinuous within $\bar{\eta}^{(i)}$, then $\bar{g}(\bar{\theta})$ may not take on the value $\bar{g}^{(i)}$ anywhere within $\bar{\eta}^{(i)}$, or for that matter anywhere within the unit cube. Since $\bar{g}(\bar{\theta})$ is bounded, however, $\bar{g}^{(i)}$ will always be bounded as well. We recall from earlier in this chapter that discontinuities in $\bar{g}(\bar{\theta})$ are restricted to values of $\bar{\theta}$ that lie on a finite number of planes passing through the unit cube. Hence, $\bar{g}(\bar{\theta})$ is almost everywhere continuous in the unit cube.

With this notation and the error bound of equation (5.2.35), we may now write

$$\begin{aligned}
\int_{\bar{0}}^{\bar{1}} \bar{g}(\bar{\theta})\ p_T[\langle\bar{\theta}\rangle]\ d\bar{\theta} &= \sum_i \int_{\bar{\eta}^{(i)}} \bar{g}(\bar{\theta})\ p_T[\langle\bar{\theta}\rangle]\ d\bar{\theta} \\
&= \sum_i \bar{g}^{(i)} \int_{\bar{\eta}^{(i)}} p_T[\langle\bar{\theta}\rangle]\ d\bar{\theta} \\
&= \sum_i \bar{g}^{(i)}\ \eta_1\ \eta_2 \cdots \eta_M \left[\, 1 + \frac{3D^{(i)}}{\eta_1\eta_2 \cdots \eta_M} \,\right] \\
&= \sum_i \bar{g}^{(i)}\ \eta_1\ \eta_2 \cdots \eta_M + \sum_i \bar{g}^{(i)}\, 3D^{(i)},
\end{aligned} \tag{6.2.33}$$

where the error $D^{(i)}$ satisfies $|D^{(i)}| \leq \Delta$ for some $\Delta > 0$. The notation $D^{(i)}$ indicates that the error is a function of i (that is, a function of the partition location), but our results from equation (5.2.35) assure us that we can find a Δ such that $|D^{(i)}| \leq \Delta$ for all i simultaneously. This Δ is a function only of the partition dimension $\bar{\eta}$ and of T. Rearranging equation (6.2.33), we have

$$\int_{\bar{0}}^{\bar{1}} \bar{g}(\bar{\theta})\, p_T[\langle \bar{\theta} \rangle]\, d\bar{\theta} - \sum_i \bar{g}^{(i)}\, \eta_1\, \eta_2 \cdots \eta_M = \sum_i \bar{g}^{(i)}\, 3D^{(i)}. \tag{6.2.34}$$

Then

$$\begin{aligned} &\left| \int_{\bar{0}}^{\bar{1}} \bar{g}(\bar{\theta})\, p_T[\langle \bar{\theta} \rangle]\, d\bar{\theta} - \sum_i \bar{g}^{(i)}\, \eta_1\, \eta_2 \cdots \eta_M \right| \\ &\quad = \left| \sum_i \bar{g}^{(i)}\, 3D^{(i)} \right| \\ &\quad \le \max_i \left| \bar{g}^{(i)} \right| 3\Delta \sum_i 1 \\ &\quad = \max_i \left| \bar{g}^{(i)} \right| \frac{3\Delta}{\eta_1\, \eta_2 \cdots \eta_M}. \end{aligned} \tag{6.2.35}$$

In the last step, we have used the fact that there are $1/(\eta_1\, \eta_2 \cdots \eta_M)$ partitions in the unit cube.

Let us focus attention on

$$\sum_i \bar{g}^{(i)}\, \eta_1\, \eta_2 \cdots \eta_M$$

in equation (6.2.34) for a moment. Let us denote by $\bar{\eta}^{(c)}$ a partition (rectangular prism) within which $\bar{g}(\bar{\theta})$ is continuous, and let us denote by $\bar{\eta}^{(n)}$ a partition within which $\bar{g}(\bar{\theta})$ is not continuous. We may write

$$\sum_i \bar{g}^{(i)}\, \eta_1 \eta_2 \cdots \eta_M = \sum_c \bar{g}^{(c)}\, \eta_1 \eta_2 \cdots \eta_M + \sum_n \bar{g}^{(n)}\, \eta_1 \eta_2 \cdots \eta_M \tag{6.2.36}$$

Let I be the total number of partitions $\bar{\eta}^{(i)}$ in the unit cube, let C be the total number of $\bar{\eta}^{(c)}$, and let N be the total number of $\bar{\eta}^{(n)}$. Since $\bar{g}(\bar{\theta})$ is almost everywhere continuous, we know that

$$\lim_{I \to \infty} \frac{C}{I} = 1 \tag{6.2.37}$$

and

$$\lim_{I \to \infty} \frac{N}{I} = 0. \tag{6.2.38}$$

The collection of partitions $\overline{\eta}^{(c)}$ is, however, nothing but a Riemann inner partition of the regions where $\overline{g}(\overline{\theta})$ is continuous. Consequently, by the definition of Riemann integration, there exists for every $\epsilon > 0$ an integer I_0 such that the relation

$$\left| \sum_{c} \overline{g}^{(c)} \eta_1 \eta_2 \cdots \eta_M - \int_{\overline{0}}^{\overline{1}} \overline{g}(\overline{\theta}) \, d\overline{\theta} \right| < \epsilon \tag{6.2.39}$$

is satisfied for all $I > I_0 = I_0(\epsilon)$. (Recall that $\overline{g}^{(c)}$ is a value of $\overline{g}$ within the partition $\overline{\eta}^{(c)}$.) Further, since

$$\sum_{n} \overline{g}^{(n)} \eta_1 \eta_2 \cdots \eta_M$$

is bounded, and approaches zero for $I \to \infty$, we may also write

$$\left| \sum_{i} \overline{g}^{(i)} \eta_1 \eta_2 \cdots \eta_M - \int_{\overline{0}}^{\overline{1}} \overline{g}(\overline{\theta}) \, d\overline{\theta} \right| < \epsilon' \tag{6.2.40}$$

for every $\epsilon' > 0$ and all $I > I_0' = I_0'(\epsilon')$.

We now combine our results to compute an upper bound on the quantity

$$\left| \int_{\overline{0}}^{\overline{1}} \overline{g}(\overline{\theta}) \, p_T[\langle \overline{\theta} \rangle] \, d\overline{\theta} - \int_{\overline{0}}^{\overline{1}} \overline{g}(\overline{\theta}) \, d\overline{\theta} \right|.$$

Using the Triangle Inequality, and then equations (6.2.35) and (6.2.40), we find that

$$\begin{aligned} &\left| \int_{\overline{0}}^{\overline{1}} \overline{g}(\overline{\theta}) \, p_T[\langle \overline{\theta} \rangle] \, d\overline{\theta} - \int_{\overline{0}}^{\overline{1}} \overline{g}(\overline{\theta}) \, d\overline{\theta} \right| \\ &\quad \leq \left| \int_{\overline{0}}^{\overline{1}} \overline{g}(\overline{\theta}) \, p_T[\langle \overline{\theta} \rangle] \, d\overline{\theta} - \sum_{i} \overline{g}^{(i)} \eta_1 \eta_2 \cdots \eta_M \right| \\ &\qquad + \left| \sum_{i} \overline{g}^{(i)} \eta_1 \eta_2 \cdots \eta_M - \int_{\overline{0}}^{\overline{1}} \overline{g}(\overline{\theta}) \, d\overline{\theta} \right| \\ &\quad \leq \max_i \left| \overline{g}^{(i)} \right| \frac{3\Delta}{\eta_1 \eta_2 \cdots \eta_M} + \epsilon'. \end{aligned} \tag{6.2.41}$$

Now, we may make ϵ' arbitrarily small by taking a sufficiently fine partition of the unit cube (that is, by taking I sufficiently large, so that we use enough rectangular prisms $\overline{\eta}^{(i)}$). For any arbitrarily fine, but fixed, partition, we may make Δ as small as we wish by taking T large enough. For a compact result, we may consider I to be an implied function of T, so that $\Delta \to 0$ and $\epsilon' \to 0$ as $T \to \infty$. Then we have that

$$\lim_{T\to\infty} \int_{\underline{0}}^{\overline{1}} \overline{g}(\overline{\theta})\, p_T[\langle\overline{\theta}\rangle]\, d\overline{\theta} = \int_{\underline{0}}^{\overline{1}} \overline{g}(\overline{\theta})\, d\overline{\theta}. \tag{6.2.42}$$

Finally, using equations (6.2.3), (6.2.18), (6.2.20), (6.2.26), (6.2.31), and (6.2.42), we conclude that

$$\begin{aligned}
a(\lambda) &= \lim_{T\to\infty} \frac{1}{T} \int_0^T g(t)\, dt \\
&= \lim_{T\to\infty} \frac{1}{T} \int_0^T \overline{g}[\overline{\theta}(t)]\, dt \\
&= \lim_{T\to\infty} \int_{\underline{0}}^{\overline{\lambda}T/(2\pi)} \overline{g}(\overline{\theta})\, \delta[\theta_1 - \lambda_1\theta_M/\lambda_M]\, \delta[\theta_2 - \lambda_2\theta_M/\lambda_M] \cdots \\
&\qquad\qquad \delta[\theta_{M-1} - \lambda_{M-1}\theta_M/\lambda_M]\, \frac{2\pi}{\lambda_M T}\, d\overline{\theta} \\
&= \lim_{T\to\infty} \int_{\underline{-\infty}}^{\overline{\infty}} \overline{g}(\overline{\theta})\, p_T(\overline{\theta})\, d\overline{\theta} \\
&= \lim_{T\to\infty} \int_{\underline{0}}^{\overline{1}} \overline{g}(\overline{\theta})\, p_T[\langle\overline{\theta}\rangle]\, d\overline{\theta} \\
&= \int_{\underline{0}}^{\overline{1}} g(\overline{\theta})\, d\overline{\theta} \\
&\equiv \int_0^1 \cdots \int_0^1 \int_0^1 \overline{g}(\theta_1,\theta_2,\ldots,\theta_M)\, d\theta_1\, d\theta_2 \cdots d\theta_M
\end{aligned} \tag{6.2.43}$$

which is what we set out to prove.

6.3 Comments

Let us consider for a moment the following problem: Let $\overline{y}(\theta_1,\theta_2,\ldots\theta_M) \equiv \overline{y}(\overline{\theta})$ be a function of M real variables, and let it be defined and integrable L^1 in the unit cube $[\overline{0},\overline{1})$. We assume that the variables θ_1, θ_2,..., θ_M are independent, that is, that they take on all values independently in $[\overline{0},\overline{1})$. Then we may associate with $\overline{y}(\overline{\theta})$ a multi-dimensional Fourier series,

$$\overline{y}(\overline{\theta}) \sim \sum_{v_1=-\infty}^{\infty} \sum_{v_2=-\infty}^{\infty} \cdots \sum_{v_M=-\infty}^{\infty} a(v_1,v_2,\ldots,v_M)\, e^{j2\pi(v_1\theta_1+v_2\theta_2+\cdots+v_M\theta_M)}, \qquad (6.3.1)$$

where the coefficients $a(v_1,v_2,\ldots,v_M)$ are given by

$$a(v_1,v_2,\ldots,v_M) \equiv \int_0^1 \cdots \int_0^1 \int_0^1 \overline{y}(\theta_1,\theta_2,\ldots,\theta_M) \, e^{-j2\pi(v_1\theta_1+v_2\theta_2+\cdots+v_M\theta_M)}\, d\theta_1\, d\theta_2 \cdots d\theta_M \qquad (6.3.2)$$

We see that this expression for the coefficient a is identical with that derived at the end of the "expected value" proof, namely equation (6.2.17). Thus, it would seem that the work of this chapter was unnecessary, and that we could have moved immediately from the independence of the $\theta_m = \lambda_m t/2\pi$, $1 \le m \le M$, to the multiple integral form of the Fourier series coefficients. This approach is essentially that taken in the heuristic "expected value" proof given in Section 6.2. It is also implicit in the work of other authors, cited in Chapter 1, who have attempted multiple Fourier series expansions to analyze a memoryless nonlinearity. These authors apparently do not understand the following principle (which justifies the need for the rigorous proof of Section 6.2): Although by assumption the basis frequencies λ_m

are independent over the integers, the fractional parts of the normalized phases, $\langle \theta_m \rangle = \langle \lambda_m t/2\pi \rangle$, are independent (uniformly distributed) in the unit cube *only in the limit* as t ranges over an infinite interval. The essence of our rigorous proof in Section 6.2 was to create a Riemann partition of the unit cube such that $\overline{g}(\overline{\theta})$ was nearly constant within each partition volume, while at the same time, the $\langle \theta_m \rangle$ had a nearly uniform distribution over each volume. Had we not placed some restrictions on $\overline{g}$, it might have varied so rapidly as a function of the θ_m that we would never have been able to find a partition satisfying both requirements. The boundedness and continuity almost everywhere of $\overline{g}$ are sufficient restrictions, and, by the work of Section 6.1, reasonable ones, for they describe practical cases of interest.

We might note in passing that some of the authors cited in Chapter 1 confuse incommensurability with independence (stating, for instance, that incommensurability of the frequencies λ_m is sufficient to obtain a multiple Fourier series expansion). Clearly, incommensurable frequencies (those having no common period) are not necessarily independent over the integers, except for the trivial case of only two frequencies. Without independence of the basis $\{\lambda_m\}$, we do not have uniform distribution of the fractional parts of the normalized phases, and we may not express the Fourier series coefficients in the multiple integral form.

A difficult problem in the theory of multiple Fourier series is the question of convergence [Ash (1976), Ash and Welland (1972)]. Suppose that we let $s(n_1,n_2,\ldots,n_M)(\overline{\theta})$ be a partial sum of the multiple Fourier series in equation (6.3.1),

so that

$$s(n_1,n_2,\ldots,n_M)(\bar{\theta}) = \sum_{v_1=-n_1}^{n_1} \sum_{v_2=-n_2}^{n_2} \cdots \sum_{v_M=-n_M}^{n_M} a(v_1,v_2,\ldots,v_M)\, e^{j2\pi(v_1\theta_1+v_2\theta_2+\cdots+v_M\theta_M)}. \tag{6.3.3}$$

Let D_G denote a distance (in the sense of Chapter 2) appropriate to $\bar{y}(\bar{\theta})$. Then if

$$D_G[\, s(n_1,n_2,\ldots n_M)(\bar{\theta}) - \bar{y}(\bar{\theta})\,] < \epsilon \tag{6.3.4}$$

for n_1, n_2,..., n_M sufficiently large, and for any $\epsilon > 0$, we say that the Fourier series converges to $\bar{y}(\bar{\theta})$. There are, of course, an infinite variety of ways in which the n_m may approach infinity, the least restrictive of which allows them to approach infinity independently. This form of convergence is called unrestrictedly rectangular.

Unfortunately, we do not have this desirable result for the Fourier series of the function $\bar{y}(\bar{\theta})$ defined in equation (6.1.8). From the results of Chapter 4, however, we know that the Bochner–Fejér polynomials for $y(t)$ converge to $y(t)$, and we see from equation (4.2.14) that this convergence is unrestrictedly rectangular with respect to the M–fold summation over rational combinations of the M basis frequencies. We also know how to construct Bochner–Fejér polynomials from the Fourier series coefficients of $y(t)$, using the procedure of equations (4.2.12) and (4.2.13). Hence, we may construct an M–fold sequence of terms $\sigma(n_1,n_2,\ldots,n_M)(\bar{\theta})$ that converges unrestrictedly rectangularly to $\bar{y}(\bar{\theta})$ in the M–dimensional unit cube. We are able to make this conclusion because we have established the equivalence of the single integral in t with the multiple integral in $\bar{\theta}$ for finding the Fourier series coefficients.

CHAPTER 7
Evaluation of the Multiple Integral

The multiple integral formulation of the Fourier series coefficients is an exact one, subject to the restrictions discussed in the previous chapter. In a few simple cases, it is possible to carry out the multiple integration analytically, to obtain exact values for the Fourier series coefficients. In most cases of practical interest, however, the integration cannot be done analytically, and we must resort to numerical methods.

From the standpoint of numerical analysis, there is nothing remarkable about the multiple integration that we must perform: the integrand is well-behaved, and the limits of integration are finite. Any of several standard techniques for multiple numerical integration could be used; these methods generally allow for iterative improvement to any desired degree of accuracy, so that we may compute the Fourier series coefficients within a very small error.

The method that we will choose here is the simplest scheme for numerical integration, namely, a multi-dimensional rectangular rule. In one dimension, a rectangular rule computation to find Fourier series coefficients is called a Discrete Fourier Transform (DFT). In multiple dimensions, therefore, we will call our rectangular rule computation a Multiple Discrete Fourier Transform (MDFT).

Besides computational simplicity, the DFT and MDFT possess a straightforward relationship between the structure of the integrand and the error in the computed result. This relationship makes error prediction and correction quite simple.

We will first review the one–dimensional case, then apply these results to the multi–dimensional problem.

7.1 Discrete Fourier Transform

Let us consider a function $y(t)$ that is periodic in t with a period of unity, and let $y(t)$ have a finite Fourier series containing harmonics from $-V$ to V. Since the Fourier series is finite, $y(t)$ is equal to its Fourier series, and we may write

$$y(t) = \sum_{v=-V}^{V} a(v)\, e^{j2\pi vt}. \tag{7.1.1}$$

As usual, the coefficients of the series are found from

$$\begin{aligned} a(n) &= \int_0^1 e^{-j2\pi nt}\, y(t)\, dt \\ &= \int_0^1 e^{-j2\pi nt} \sum_{v=-V}^{V} a(v)\, e^{j2\pi vt}\, dt. \end{aligned} \tag{7.1.2}$$

(We are again using the shorthand notation $a(n)$ for the coefficient of the term having frequency $2\pi n$, that is, the n^{th} harmonic of the fundamental frequency 2π.) We may approximate the last integral in equation (7.1.2) using a set of K rectangular partitions; calling this approximation I, we have

$$\begin{aligned} I &= \sum_{k=0}^{K-1} e^{-j2\pi nk/K} \sum_{v=-V}^{V} a(v)\, e^{j2\pi vk/K}\, \frac{1}{K} \\ &= \frac{1}{K} \sum_{v=-V}^{V} a(v) \sum_{k=0}^{K-1} e^{j2\pi(v-n)k/K} \\ &= \frac{1}{K}\, a(n) \sum_{k=0}^{K-1} 1 + \frac{1}{K} \sum_{\substack{v=-V \\ v\neq n}}^{V} a(v) \sum_{k=0}^{K-1} e^{j2\pi(v-n)k/K} \end{aligned}$$

$$= a(n) + \frac{1}{K} \sum_{\substack{v=-V \\ v \neq n}}^{V} a(v) \frac{1 - e^{j2\pi(v-n)}}{1 - e^{j2\pi(v-n)/K}}. \tag{7.1.3}$$

In this last step, we have used the result for the sum of a finite geometric series. Since we wish to have $I = a(n)$ in equation (7.1.3), we require that each term in the summation

$$\sum_{\substack{v=-V \\ v \neq n}}^{V} a(v) \frac{1 - e^{j2\pi(v-n)}}{1 - e^{j2\pi(v-n)/K}}$$

be zero, hence that

$$\frac{1 - e^{j2\pi(v-n)}}{1 - e^{j2\pi(v-n)/K}} = 0 \quad \text{for } v \neq n, \ |v| \leq V. \tag{7.1.4}$$

Because $v - n$ is always an integer, the numerator of equation (7.1.4) is always zero; therefore, we may enforce the equality simply by ensuring that the denominator be non–zero. Thus, we must make

$$e^{j2\pi(v-n)/K} \neq 1, \tag{7.1.5}$$

which in turn requires that

$$\frac{v - n}{K} \neq \text{integer} \quad \text{for } v \neq n, \ |v| \leq V. \tag{7.1.6}$$

The smallest K satisfying this requirement is

$$K = V + |n| + 1. \tag{7.1.7}$$

To ensure that all non–zero $a(n)$ can be calculated correctly, we take the largest value of $|n|$, namely V, so that

$$K \geq 2V + 1. \tag{7.1.8}$$

V is the largest Hertzian frequency contained in $y(t)$; $1/K$ is the sample interval used in our rectangular rule approximation to the integral, so we may call K the Hertzian sample frequency. Thus, the requirement on K states that the sample frequency must be greater than twice the highest frequency in the signal y.

This requirement is known as the Sampling Theorem, and is a well–known result in communication theory. We may also justify this requirement in another way, by noting that y contains $2V + 1$ independent components (complex exponentials); certainly, to determine the amplitudes of these components, we need a minimum of $2V + 1$ samples.

Let us now extend the problem as follows: Let us choose $K = 2V + 1$, where V is the highest frequency of interest in $y(t)$; now, however, we will allow $a(n) \neq 0$ for $|n| > V$. In the extreme case, the Fourier series for y might have an infinite number of terms. In this case, we cannot necessarily write that y is equal to its Fourier series (to do so would require uniform convergence, which is not guaranteed); however, we may still replace y with its Fourier series in the above calculations, since clearly y and its Fourier series both have the same Fourier series representation.

Following the approach of equation (7.1.3), with a third term included to handle $|v| > V$, we find that

$$I = a(n) + \frac{1}{K} \sum_{\substack{v=-V \\ v \neq n}}^{V} a(v) \sum_{k=0}^{K-1} e^{j2\pi(v-n)k/K} + \frac{1}{K} \sum_{|v|>V} a(v) \sum_{k=0}^{K-1} e^{j2\pi(v-n)k/K}$$

$$= a(n) + 0 + \sum_{\substack{|v|>V \\ (v-n)/K \,=\, \text{integer} \neq 0}} a(v)$$

$$= a(n) + [a(n+K) + a(n-K)] + [a(n+2K) + a(n-2K)] + \cdots. \qquad (7.1.9)$$

Thus, the desired result $I = a(n)$ is corrupted by coefficients whose frequencies are distant from n by multiples of K. This effect is called aliasing, and again is well known in communication theory. Clearly, if the Fourier series for $y(t)$ is infinite, aliasing can never be fully eliminated, regardless of how large we make K.

The error due to aliasing can be made arbitrarily small, however, for most practical problems, because the coefficients $a(n)$ decrease in magnitude with increasing $|n|$.

It is also possible to use the DFT to find the Fourier series coefficients of an almost periodic function, although the results are not as favorable as for a periodic function. In the almost periodic case, reducing error in the computed coefficients to an acceptable level may require a very large number of very short sample intervals, and even if the Fourier series is of finite length, there is no way to eliminate the error completely. Not only is there an aliasing effect, in which a given coefficient is corrupted by coefficients corresponding to higher frequencies, but there is also a leakage effect, in which a given coefficient is corrupted by coefficients corresponding to nearby frequencies. The aliasing effect is similar to aliasing in the periodic case, but the leakage effect has no analog there because a periodic function possesses only harmonically–related frequencies.

The exact nature of these effects, and the strategy needed to control them, are probably best understood through Fourier Transform analysis [see Geckinli and Yavuz (1983), for example], and we will not pursue the matter any further.

7.2 Multiple Discrete Fourier Transform

We will now extend the DFT to multiple dimensions, using the results developed in Section 7.1. Let $\overline{y}(\overline{\theta}) = \overline{y}(\theta_1,\theta_2,\ldots,\theta_M)$ be a function periodic in each of the variables θ_1, θ_2,..., θ_M, with a period of unity. The structure of $\overline{y}$ will be that discussed in Chapter 6; specifically, $\overline{y}$ will be bounded, real–valued, and almost everywhere continuous in the unit cube $\overline{0} \leq \overline{\theta} < \overline{1}$. As we found in Chapter 6, $\overline{y}$ has the Fourier series representation

$$\bar{y}(\bar{\theta}) \sim \sum_{v_1=-\infty}^{\infty} \sum_{v_2=-\infty}^{\infty} \cdots \sum_{v_M=-\infty}^{\infty} a(v_1,v_2,\ldots,v_M)\, e^{j2\pi(v_1\theta_1+v_2\theta_2+\cdots+v_M\theta_M)} \tag{7.2.1}$$

with coefficients determined from

$$\begin{aligned} a(n_1,n_2,\ldots,n_M) &= \int_0^1 \cdots \int_0^1 \int_0^1 e^{-j2\pi(n_1\theta_1+n_2\theta_2+\cdots+n_M\theta_M)} \bar{y}(\theta_1,\theta_2,\ldots,\theta_M)\, d\theta_1\, d\theta_2 \cdots d\theta_M \\ &= \int_0^1 \cdots \int_0^1 \int_0^1 e^{-j2\pi(n_1\theta_1+n_2\theta_2+\cdots+n_M\theta_M)} \\ &\qquad \sum_{v_1=-\infty}^{\infty} \sum_{v_2=-\infty}^{\infty} \cdots \sum_{v_M=-\infty}^{\infty} a(v_1,v_2,\ldots,v_M)\, e^{j2\pi(v_1\theta_1+v_2\theta_2+\cdots+v_M\theta_M)} \\ &\qquad d\theta_1\, d\theta_2 \cdots d\theta_M \end{aligned} \tag{7.2.2}$$

We may regroup the factors in this last equation to emphasize the iterated nature of the multiple integral:

$$\begin{aligned} a(n_1,n_2,\ldots,n_M) = \int_0^1 e^{-j2\pi n_1\theta_1} \sum_{v_1} \Bigg[\int_0^1 e^{-j2\pi n_2\theta_2} \sum_{v_2} \Bigg[\int_0^1 e^{-j2\pi n_3\theta_3} \sum_{v_3} \cdots \\ \Bigg[\int_0^1 e^{-j2\pi n_M\theta_M} \sum_{v_M} a(v_1,v_2,\ldots,v_M)\, e^{j2\pi v_M\theta_M}\, d\theta_M \Bigg] \cdots \\ e^{j2\pi v_3\theta_3}\, d\theta_3 \Bigg] e^{j2\pi v_2\theta_2}\, d\theta_2 \Bigg] e^{j2\pi v_1\theta_1}\, d\theta_1. \end{aligned} \tag{7.2.3}$$

Let us consider the expression within the innermost pair of brackets in equation (7.2.3). With v_1, v_2,...,v_{M-1} fixed, this expression is simply a calculation to find the Fourier series coefficients of a periodic function of θ_M. Consequently, we can use the DFT discussed earlier to approximate the integral. If we use K_M

samples in this DFT, the result of the computation is

$$\sum_{i_M=-\infty}^{\infty} a(v_1,v_2,\ldots,n_M+i_M K_M).$$

If we repeat this procedure for the integration in θ_{M-1}, using K_{M-1} samples, the quantity within the second–innermost brackets becomes

$$\sum_{i_{M-1}=-\infty}^{\infty} \sum_{i_M=-\infty}^{\infty} a(v_1,v_2,\ldots, n_{M-1}+i_{M-1}K_{M-1}, n_M+i_M K_M).$$

Finally, after we apply the DFT M times, the approximation to the M–dimensional integral becomes

$$\begin{aligned} a(n_1,n_2,\ldots,n_M) &\simeq \sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} \cdots \sum_{k_M=0}^{K_M-1} e^{-j2\pi(n_1k_1/K_1 + n_2k_2/K_2 + \cdots + n_Mk_M/K_M)} \\ &\qquad \overline{y}(k_1/K_1, k_2/K_2,\ldots, k_M/K_M)\,\frac{1}{K_1K_2\cdots K_M} \\ &= \sum_{i_1=-\infty}^{\infty} \sum_{i_2=-\infty}^{\infty} \cdots \sum_{i_M=-\infty}^{\infty} a(n_1+i_1K_1, n_2+i_2K_2,\ldots, n_M+i_MK_M). \end{aligned} \tag{7.2.4}$$

We will call this approximation a Multiple Discrete Fourier Transform (MDFT). We see from equation (7.2.4) that aliasing in the MDFT is an M–dimensional effect, with the error in each dimension depending on the number of samples taken in that dimension. Moreover, there is no leakage among nearby Fourier series coefficients, because of the regular spacing of frequencies within each dimension.

The MDFT allows for iterative improvement during computation, just as the single–dimensional DFT does. The most efficient schemes for iterative improvement make use of all previously computed sample points, typically by repeatedly halving

any or all of the interval sizes $1/K_1$, $1/K_2$,..., $1/K_M$. As we have just found, aliasing occurs independently in each dimension. Thus, if dimension m contains only a few significant harmonics (that is, if $a(n_1,...,n_m,...,n_M)$ is non–zero only for small values of $|n_m|$), then only a relatively large partition size $1/K_m$ will be needed. Consequently, we can expend most of our computational effort on those dimensions having many harmonics. This result is important in practical problems, for we often wish to analyze a nonlinearity driven by one or two signals of rather large amplitude, and several others of quite small amplitude. (This situation would occur, for example, when a nonlinear element is used as a mixer). If these signals have frequencies independent over the integers, we might expect, at least intuitively, that the disparity of amplitudes at the input of the nonlinearity will reflect itself as a disparity of harmonic content at the output.

Matrix Representation of the MDFT

As defined in equation (7.2.4), the MDFT employs $K_1K_2\cdots K_M$ sample points distributed throughout the M–dimensional unit cube. Instead of indexing these points using an M–fold summation in k_1, k_2,..., k_M, let us number the points (in any order) from 1 to $K_1K_2\cdots K_M$ and index them using a single summation in q. Taking k_1, k_2,..., k_M to be functions of q, and defining $K_1K_2\cdots K_M \equiv Q$, we may write

$$a(n_1,n_2,...,n_M) \simeq \sum_{q=1}^{Q} e^{-j2\pi[n_1k_1(q)/K_1 + n_2k_2(q)/K_2 + \cdots + n_Mk_M(q)/K_M]} \overline{y}[k_1(q)/K_1, k_2(q)/K_2,..., k_M(q)/K_M] \frac{1}{K_1K_2\cdots K_M}. \qquad (7.2.5)$$

Let us assume that we wish to calculate P different Fourier series coefficients, for P specified sets of values of $n_1, n_2,..., n_M$. Following the approach just taken for the sample points, let us index these values of $n_1, n_2,..., n_M$ using the variable p, so that equation (7.2.5) becomes

$$a[n_1(p), n_2(p),..., n_M(p)] \simeq \sum_{q=1}^{Q} e^{-j2\pi[n_1(p)k_1(q)/K_1 + n_2(p)k_2(q)/K_2 + \cdots + n_M(p)k_M(q)/K_M]} \overline{y}[k_1(q)/K_1, k_2(q)/K_2,..., k_M(q)/K_M] \frac{1}{K_1K_2\cdots K_M},$$
$$1 \le p \le P. \tag{7.2.6}$$

We may now represent the computation of equation (7.2.6), for all values of p simultaneously, using matrix notation. Let us define a matrix E of dimension $P \times Q$, having elements

$$e_{p,q} = e^{-j2\pi[n_1(p)k_1(q)/K_1 + n_2(p)k_2(q)/K_2 + \cdots + n_M(p)k_M(q)/K_M]} \cdot \frac{1}{K_1K_2\cdots K_M},$$
$$1 \le p \le P,\ 1 \le q \le Q. \tag{7.2.7}$$

Let us also define a column vector Y of length Q, having

$$y_q = \overline{y}[k_1(q)/K_1, k_2(q)/K_2,..., k_M(q)/K_M],\ 1 \le q \le Q, \tag{7.2.8}$$

and a column vector A of length P, having

$$a_p = a[n_1(p), n_2(p),..., n_M(p)],\ 1 \le p \le P. \tag{7.2.9}$$

Then

$$A = E \cdot Y. \tag{7.2.10}$$

E depends only on the choice of Fourier series coefficients to be computed, and on the partition sizes used in the computation. Y is a vector of samples of the nonlinearity output, and depends on both the nonlinear function and the nonlinearity input. Clearly, once we have computed the elements of E, we may find

the Fourier series coefficients corresponding to any nonlinearity output by using a simple matrix multiplication (provided that the set of basis frequencies remains the same, and that the chosen partition sizes are small enough to give the required accuracy). Thus, implementation of the MDFT can be made very efficient in computer problems that require its repeated application. (An example is its use as part of the method of harmonic balance, which is iterative in nature.) Further, if we must perform iterative improvement on the MDFT by adding sample points, these additional samples simply increase the length of the vector Y and the column dimension of the matrix E.

Comparison of MDFT with Least Squares

Let us now compare the MDFT with the method of least squares. Although least squares is well known as a general technique in numerical analysis, it appears that Chua and Ushida (1981) were the first to apply it to the computation of Fourier series coefficients. Their work has been cited frequently in the current literature, and has formed the basis for at least one additional investigation [Sorkin, Kundert, and Sangiovanni-Vincentelli (1987)]. We will find that, at best, least squares is rather inefficient compared with the MDFT; at worst, it becomes computationally intractable.

Let $y(t)$ be an almost periodic function of t, and let us approximate y using a function $\tilde{y}(t)$ that has the finite-length trigonometric series

$$\tilde{y}(t) \equiv \sum_{v=-V}^{V} c(\lambda_v)\, e^{j\lambda_v t}. \tag{7.2.11}$$

We know from Chapter 4 that we should choose each coefficient $c(\lambda_v)$ to be the

Fourier series coefficient $a(\lambda_v)$ if we wish to achieve the smallest mean-square error between $y(t)$ and its series approximation $\tilde{y}(t)$.

To aid in writing a matrix representation of the problem, let us number the frequencies in any order from 1 to $2V + 1$, and index them with the variable p. Taking v to be a function of p, and setting $2V + 1 = P$, we may write

$$\tilde{y}(t) \equiv \sum_{p=1}^{P} c[\lambda_{v(p)}]\, e^{j\lambda_{v(p)}t}. \tag{7.2.12}$$

We now wish to take Q samples of the approximating function $\tilde{y}(t)$ for Q distinct values of t. Denoting these values of t by $t(q)$, we may express equation (7.2.12) as

$$\tilde{y}[t(q)] \equiv \sum_{p=1}^{P} c[\lambda_{v(p)}]\, e^{j\lambda_{v(p)}t(q)}, \quad 1 \le q \le Q. \tag{7.2.13}$$

There are a variety of possible choices for the sample instants $t(q)$, the simplest being a uniform spacing that makes $t(q) = (q\text{-}1)\Delta t$ for some Δt. Chua and Ushida (1981) advocate this approach, while Sorkin, Kundert, and Sangiovanni-Vincentelli (1987) suggest choosing the values randomly. The choice of $t(q)$ affects the numerical conditioning of the problem, and the wrong choice can result in an attempt to invert a nearly singular matrix. Unfortunately, making an informed choice of $t(q)$ requires knowing, in advance, the significant frequencies contained in $y(t)$.

We may now write the computation of equation (7.2.13), for all values of q simultaneously, in matrix form. Let us define a matrix Γ of dimension $Q \times P$, having elements

$$\gamma_{q,p} = e^{j\lambda_{v(p)}t(q)}, \quad 1 \le p \le P, \quad 1 \le q \le Q. \tag{7.2.14}$$

Let us also define a column vector C of length P, having elements

$$c_p = c[\lambda_{v(p)}], \quad 1 \le p \le P, \tag{7.2.15}$$

and a column vector $\tilde{Y}$ of length Q, having elements

$$\tilde{y}_q = \tilde{y}[t(q)], \quad 1 \le q \le Q. \tag{7.2.16}$$

Then

$$\tilde{Y} = \Gamma \cdot C. \tag{7.2.17}$$

Equation (7.2.17) is a single–dimensional inverse Discrete Fourier Transform (that is, a transform from the frequency to the time domain) involving C, the vector of unknown Fourier series coefficients. The vector $\tilde{Y}$, containing samples of the function $\tilde{y}(t)$, is also unknown, but it is intended to be an approximation to the vector Y, whose elements are

$$y_q = y[t(q)], \quad 1 \le q \le Q. \tag{7.2.18}$$

Consequently, we will choose the coefficients in C to minimize the mean–square error between the elements of $\tilde{Y}$ and their corresponding elements in Y. That is, we will choose the elements of C to minimize

$$\text{M.S.E.} = (Y - \Gamma \cdot C)^{\dagger} (Y - \Gamma \cdot C)/Q, \tag{7.2.19}$$

where $\dagger$ indicates the "complex conjugate – transpose" operation on a matrix. Equation (7.2.19) is just a least squares problem, and it is well known that

$$C = (\Gamma^{\dagger}\Gamma)^{-1} \Gamma^{\dagger} \cdot Y \tag{7.2.20}$$

minimizes the mean–square error. (Chua and Ushida (1981), for example, give a proof for the case of real–valued matrices that carries over immediately to complex–valued matrices.)

We should note again at this point that Γ is not necessarily a square matrix. If it is not square, its row dimension should exceed its column dimension, to ensure that there are more samples of $\tilde{y}$ than there are unknown coefficients in C. If Γ does

happen to be a square matrix, then $(\Gamma^\dagger\Gamma)^{-1}\,\Gamma^\dagger = \Gamma^{-1}(\Gamma^\dagger)^{-1}\,\Gamma^\dagger = \Gamma^{-1}$, so that

$$C = \Gamma^{-1}\,Y, \quad \Gamma \text{ square.} \tag{7.2.21}$$

This result tells us that when the number of coefficients equals the number of sample points, the least squares solution forces $y(t)$ and $\tilde{y}(t)$ to be equal at the Q points $t = t(q)$.

Let us compare the matrix formulation of the MDFT with that of the least squares method. For a given choice of Q samples of y, and a given choice of P coefficients to be calculated, both methods take the form of a $P \times Q$ matrix multiplying the same $Q \times 1$ vector of function samples [see equations (7.2.10) and (7.2.20)]. There is an essential difference between them, however: The MDFT (and, as a special case, the single–dimensional DFT for periodic functions) represents the function $y(t)$ by a Fourier series that may contain enumerably infinite frequencies; the MDFT computes series coefficients that minimize the mean–square error between this infinite series and the Q samples of $y(t)$. On the other hand, the least squares method approximates the function $y(t)$ by a finite Fourier series; least squares computes series coefficients that minimize the mean–square error between this finite series and the Q samples of $y(t)$.

Because least squares begins with an approximate series representation of $y(t)$, in general it returns values of coefficients different from those computed by the MDFT. Results of the two methods agree only in two cases: first, if the Fourier series for $y(t)$ is actually finite and all of its non–zero coefficients appear in C (in which case, both methods compute the exact values of Fourier series coefficients; second, if $y(t)$ is purely periodic and the sample points are uniformly spaced in one period of $y(t)$ (in which case, least squares is analytically the same as the periodic DFT, as Chua and Ushida (1981) proved).

Computationally, the MDFT and least squares methods exhibit substantial differences. Because the least squares problem is formulated as an inverse Discrete Fourier Transform [see equation (7.2.17)], its solution requires inversion of the $P \times P$ matrix $\Gamma^{\dagger}\Gamma$, while the MDFT, being a forward transform, requires no matrix inversion. Accurate results when using least squares require that we assume a sufficient number of coefficients in the vector C. Many practical functions $y(t)$ have tens or even hundreds of significant coefficients, all of which must be included in C and all of which contribute to the dimension of $\Gamma^{\dagger}\Gamma$. Further, iterative improvement is not possible using least squares. If we carry out a least squares solution and then discover that we did not include enough assumed coefficients in C, we must begin again, and carry out the inversion of an even larger matrix.

As we mentioned earlier, the least squares formulation also suffers from a potential numerical ill–conditioning if the sample points $t(q)$ are not properly chosen. Sorkin, Kundert, and Sangiovanni–Vincentelli (1987) noted this problem, and suggest oversampling $y(t)$, then performing an orthogonalization procedure to create a Γ matrix with linearly independent rows. The MDFT avoids this difficulty, first, by not requiring a matrix inversion, and second, by using a superposition of single–dimensional periodic DFT's, each of which is well–conditioned.

We can view the MDFT as providing an algorithm for selecting the optimum values of t at which to sample $y(t)$. We are able to do so because we have information regarding the structure of y. Recalling our work of Chapters 5 and 6, we can write the nonlinearity input $x(t)$ as

$$x(t) = \bar{x}(\theta_1, \theta_2, \ldots, \theta_M),$$

where each $\theta_m = \theta_m(t) = \lambda_m t/(2\pi)$, and the λ_m are the basis frequencies of x. From Chapter 3, we know that the nonlinearity output $y(t)$ has the same basis

frequencies, so that we can also write

$$y(t) = \bar{y}(\theta_1, \theta_2, \ldots, \theta_M).$$

In carrying out the MDFT, we sample $\bar{y}$ at regular intervals in M-space, rather than at regular intervals along the trajectory $\bar{\theta} = \bar{\lambda}t/(2\pi)$ in M-space. (The chosen sample points may not even lie on the trajectory for finite values of t.) The choice of these sample points is quite clear from the multiple integral form of the Fourier series coefficients; it is quite unclear from the original single integral form on which the least squares method and the DFT are based.

CHAPTER 8
Examples

We now consider five short examples that employ the multiple integral formulation of the Fourier series coefficients. For this purpose, the author wrote a Fortran program that implements the Multiple Discrete Fourier Transform (MDFT) algorithm discussed in Chapter 7. The program incorporates iterative improvement by repeatedly halving the interval size in a particular dimension until all coefficients exhibit a magnitude change smaller than the requested accuracy. Iteration stops when every dimension has been checked in this manner.

8.1 Example 1: Half–Wave Square Law I

Let us consider a nonlinear function $\phi(x)$ defined by

$$\phi(x) = \begin{cases} x^2, & x \geq 0 \\ 0, & x < 0. \end{cases} \tag{8.1.1}$$

Such a function represents a half–wave square–law detector. For $x(t)$, let us take

$$x(t) = 2 + 1 \cos \omega_1 t + 0.5 \cos \omega_2 t, \tag{8.1.2}$$

where ω_1 and ω_2 are independent over the integers. We wish to compute the Fourier series coefficients of $y(t) = \phi[x(t)]$ to an absolute accuracy of 10^{-6}.

Since ω_1 and ω_2 are independent, the frequency basis for $x(t)$ and $y(t)$ has two members, and a double integral will be needed to calculate the Fourier series coefficients. Hence, application of the MDFT will give rise to a double summation of sample points.

As in Chapter 3, we will use the notion of the "order" of a frequency in $y(t)$, saying for example that ω is a second–order frequency because it cannot be produced by a nonlinearity having a power of x less than the second. We will also refer to the "order" of a Fourier series coefficient, meaning the order of the frequency to which the coefficient corresponds.

We note in this example that $x(t) > 0$ for all t, so that effectively $\phi(x) = x^2$. Consequently, we know that the Fourier series for $y(t)$ is of finite length, and that it will contain only zero– through second–order frequencies. Using equations (3.2.7) and (3.2.8), with $M = 2$, $Q_1 = Q_2 = 1$, and $n = 2$, we deduce that these frequencies are of the form

$$\omega = v_1 \omega_1 + v_2 \omega_2, \quad |v_1|, |v_2| \leq 2. \tag{8.1.3}$$

A listing of the computer results is given in Table 8.1.1. The program was asked to compute all coefficients through fourth order as a check on its performance; just as we predicted, all coefficients of third and fourth order are zero, within the requested accuracy. It is necessary to give results in Table 8.1.1 only for positive values of v_1 and v_2, for the following reason: From equation (8.1.2), we see that $x(t)$ is an even function of its two arguments $\theta_1 = \omega_1 t/(2\pi)$ and $\theta_2 = \omega_2 t/(2\pi)$. Consequently, the composite function $\overline{y}(\theta_1,\theta_2) = \phi[\overline{x}(\theta_1,\theta_2)]$ is also an even function of both of its arguments θ_1 and θ_2, and further, it is real–valued. Examining the integral expression for $a(v_1,v_2)$ in equation (6.2.17) shows that under these conditions the imaginary parts of the coefficients should be zero (and they are, again within the requested accuracy) and that $a(v_1,v_2) = a(|v_1|,|v_2|)$.

It is an easy calculation to show that the results of Table 8.1.1 are indeed correct, recalling that these are exponential, not trigonometric, Fourier series coefficients.

Table 8.1.1

Results of Example 1.

Partitions in dimension 1: 36
Partitions in dimension 2: 18

v_1	v_2	Re[a(v_1,v_2)]	Im[a(v_1,v_2)]
0	0	4.62499902	0.00000000
1	0	1.99999935	−0.00000004
2	0	0.24999954	−0.00000008
3	0	−0.00000043	−0.00000011
4	0	−0.00000042	−0.00000015
0	1	0.99999956	−0.00000006
1	1	0.24999980	−0.00000003
2	1	−0.00000008	−0.00000001
3	1	−0.00000007	−0.00000002
4	1	−0.00000007	−0.00000003
0	2	0.06249968	−0.00000011
1	2	−0.00000012	−0.00000004
2	2	−0.00000001	0.00000000
3	2	0.00000000	0.00000000
4	2	0.00000000	0.00000000
0	3	−0.00000029	−0.00000017
1	3	−0.00000011	−0.00000006
2	3	−0.00000001	−0.00000001
3	3	0.00000000	0.00000000
4	3	0.00000000	0.00000000
0	4	−0.00000027	−0.00000023
1	4	−0.00000010	−0.00000008
2	4	−0.00000001	−0.00000001
3	4	0.00000000	0.00000000
4	4	0.00000000	0.00000000

8.2 Example 2: Half-Wave Square Law II

Let us use the same $\phi(x)$ as in Example 1, this time with

$$x(t) = 1 + 1\ cos\ \omega_1 t + 0.5\ cos\ \omega_2 t. \tag{8.2.1}$$

With this choice of excitation, $\phi(x)$ no longer has the finite power series representation $\phi(x) = x^2$. As a consequence, we would expect the Fourier series for $y(t)$ to be infinite; the computer results of Table 8.2.1 suggest that such is the case, as the third- and fourth-order coefficients are no longer zero. The much larger number of partitions now needed to achieve the requested accuracy of 10^{-6} also suggests the presence of many significant higher-order coefficients.

Table 8.2.1

Results of Example 2.

Partitions in dimension 1: 288
Partitions in dimension 2: 72

v_1	v_2	Re[a(v_1,v_2)]	Im[a(v_1,v_2)]
0	0	1.61512771	0.00000000
1	0	1.00924430	0.00000000
2	0	0.24244539	0.00000000
3	0	0.00529079	-0.00000001
4	0	-0.00304647	-0.00000001
0	1	0.50857704	-0.00000001
1	1	0.24198441	0.00000000
2	1	0.00650355	0.00000000
3	1	-0.00448948	0.00000000
4	1	0.00250957	0.00000000
0	2	0.05700217	-0.00000001
1	2	0.00509724	-0.00000001
2	2	-0.00402834	0.00000000
3	2	0.00262828	0.00000000
4	2	-0.00129476	0.00000000
0	3	0.00235592	-0.00000002
1	3	-0.00213805	-0.00000001
2	3	0.00156554	0.00000000
3	3	-0.00084468	0.00000000
4	3	0.00020887	0.00000000
0	4	-0.00043630	-0.00000002
1	4	0.00035430	-0.00000001
2	4	-0.00014848	0.00000000
3	4	-0.00008504	0.00000000
4	4	0.00024504	0.00000000

8.3 Example 3: Full-Wave Square Law

For this example, let us take

$$\phi(x) = x^2, \quad -\infty < x < \infty, \tag{8.3.1}$$

with

$$x(t) = 1 \cos \omega_1 t + 2 \cos \omega_2 t + 3 \cos \omega_3 t, \tag{8.3.2}$$

where ω_1, ω_2, and ω_3 are independent over the integers. There are now three basis frequencies, so that the MDFT will require a triple summation of sample points. We still expect the Fourier series for $y(t)$ to be finite, however, and to contain at most zero- through second-order frequencies.

In calculating the MDFT on the computer, the requested accuracy this time was 10^{-4}, although because the Fourier series is finite, somewhat better accuracy was achieved. As a check, all coefficients through third order were computed, and the third-order coefficients were indeed zero within computational error. To save space, Table 8.3.1 presents the results only for the zero- through second-order coefficients.

In Table 8.3.1, we have listed the coefficients for all octants of three-space, not just the first octant. We have done so for two reasons: First, it confirms the statement made in Example 1 that $a(v_1,v_2,\ldots) = a(|v_1|,|v_2|,\ldots)$ for cases where $\overline{y(\theta)}$ is real and an even function of its arguments θ_1, θ_2,... . Second, it allows us to check our work of Section 3.3, where we computed the number of frequencies at the output of an N^{th}-order nonlinearity driven by M sinusoids having independent frequencies. Here, $N = 2$ and $M = 3$; further, since $x(t)$ contains no d.c. component, we know that only second- and zero-order frequencies will appear in $y(t)$. Applying equation (3.3.5) for $n = 2$ gives us 18 second-order frequencies; by inspection, there

is 1 zero–order frequency (namely, d.c.), so that there are a total of 19 frequencies in $y(t)$. Table 8.3.1 confirms this result.

Table 8.3.1

Results of Example 3.

Partitions in dimension 1: 14
Partitions in dimension 2: 14
Partitions in dimension 3: 14

v_1	v_2	v_3	Re[a(v_1,v_2,v_3)]	Im[a(v_1,v_2,v_3)]
−2	−2	−2	0.00000000	0.00000000
−1	−2	−2	0.00000000	0.00000000
0	−2	−2	−0.00000015	0.00000007
1	−2	−2	0.00000000	0.00000000
2	−2	−2	0.00000000	0.00000000
−2	−1	−2	0.00000000	0.00000000
−1	−1	−2	−0.00000004	0.00000002
0	−1	−2	−0.00000037	0.00000015
1	−1	−2	−0.00000004	0.00000002
2	−1	−2	0.00000000	0.00000000
−2	0	−2	−0.00000011	0.00000005
−1	0	−2	−0.00000024	0.00000009
0	0	−2	2.24999915	0.00000025
1	0	−2	−0.00000024	0.00000004
2	0	−2	−0.00000011	−0.00000004
−2	1	−2	0.00000000	0.00000000
−1	1	−2	−0.00000004	0.00000002
0	1	−2	−0.00000037	0.00000011
1	1	−2	−0.00000004	0.00000002
2	1	−2	0.00000000	0.00000000
−2	2	−2	0.00000000	0.00000000
−1	2	−2	0.00000000	0.00000000
0	2	−2	−0.00000015	−0.00000003
1	2	−2	0.00000000	0.00000000
2	2	−2	0.00000000	0.00000000
−2	−2	−1	0.00000000	0.00000000
−1	−2	−1	−0.00000007	0.00000003
0	−2	−1	−0.00000031	0.00000014
1	−2	−1	−0.00000007	0.00000003
2	−2	−1	0.00000000	0.00000000

Table 8.3.1 (continued)

v_1	v_2	v_3	Re[a(v_1,v_2,v_3)]	Im[a(v_1,v_2,v_3)]
−2	−1	−1	−0.00000013	0.00000006
−1	−1	−1	−0.00000026	0.00000006
0	−1	−1	2.99999912	0.00000013
1	−1	−1	−0.00000026	−0.00000001
2	−1	−1	−0.00000013	−0.00000006
−2	0	−1	−0.00000014	0.00000007
−1	0	−1	1.49999956	0.00000006
0	0	−1	−0.00000095	0.00000012
1	0	−1	1.49999956	0.00000000
2	0	−1	−0.00000014	−0.00000006
−2	1	−1	−0.00000013	0.00000006
−1	1	−1	−0.00000026	0.00000003
0	1	−1	2.99999912	0.00000000
1	1	−1	−0.00000026	−0.00000004
2	1	−1	−0.00000013	−0.00000006
−2	2	−1	0.00000000	0.00000000
−1	2	−1	−0.00000007	−0.00000003
0	2	−1	−0.00000031	−0.00000012
1	2	−1	−0.00000007	−0.00000003
2	2	−1	0.00000000	0.00000000
−2	−2	0	−0.00000006	0.00000003
−1	−2	0	−0.00000014	0.00000005
0	−2	0	0.99999945	0.00000019
1	−2	0	−0.00000014	0.00000003
2	−2	0	−0.00000006	−0.00000002
−2	−1	0	−0.00000010	0.00000005
−1	−1	0	0.99999971	0.00000004
0	−1	0	−0.00000080	0.00000010
1	−1	0	0.99999971	0.00000000
2	−1	0	−0.00000010	−0.00000004
−2	0	0	0.24999963	0.00000016
−1	0	0	−0.00000059	0.00000008
0	0	0	6.99999827	0.00000000
1	0	0	−0.00000059	−0.00000008
2	0	0	0.24999963	−0.00000016

Table 8.3.1 (continued)

v_1	v_2	v_3	Re[$a(v_1,v_2,v_3)$]	Im[$a(v_1,v_2,v_3)$]
−2	1	0	−0.00000010	0.00000004
−1	1	0	0.99999971	0.00000000
0	1	0	−0.00000080	−0.00000010
1	1	0	0.99999971	−0.00000004
2	1	0	−0.00000010	−0.00000005
−2	2	0	−0.00000006	0.00000002
−1	2	0	−0.00000014	−0.00000003
0	2	0	0.99999945	−0.00000019
1	2	0	−0.00000014	−0.00000005
2	2	0	−0.00000006	−0.00000003
−2	−2	1	0.00000000	0.00000000
−1	−2	1	−0.00000007	0.00000003
0	−2	1	−0.00000031	0.00000012
1	−2	1	−0.00000007	0.00000003
2	−2	1	0.00000000	0.00000000
−2	−1	1	−0.00000013	0.00000006
−1	−1	1	−0.00000026	0.00000004
0	−1	1	2.99999912	0.00000000
1	−1	1	−0.00000026	−0.00000003
2	−1	1	−0.00000013	−0.00000006
−2	0	1	−0.00000014	0.00000006
−1	0	1	1.49999956	0.00000000
0	0	1	−0.00000095	−0.00000012
1	0	1	1.49999956	−0.00000006
2	0	1	−0.00000014	−0.00000007
−2	1	1	−0.00000013	0.00000006
−1	1	1	−0.00000026	0.00000001
0	1	1	2.99999912	−0.00000013
1	1	1	−0.00000026	−0.00000006
2	1	1	−0.00000013	−0.00000006
−2	2	1	0.00000000	0.00000000
−1	2	1	−0.00000007	−0.00000003
0	2	1	−0.00000031	−0.00000014
1	2	1	−0.00000007	−0.00000003
2	2	1	0.00000000	0.00000000

Table 8.3.1 (continued)

v_1	v_2	v_3	Re[$a(v_1,v_2,v_3)$]	Im[$a(v_1,v_2,v_3)$]
−2	−2	2	0.00000000	0.00000000
−1	−2	2	0.00000000	0.00000000
0	−2	2	−0.00000015	0.00000003
1	−2	2	0.00000000	0.00000000
2	−2	2	0.00000000	0.00000000
−2	−1	2	0.00000000	0.00000000
−1	−1	2	−0.00000004	−0.00000002
0	−1	2	−0.00000037	−0.00000011
1	−1	2	−0.00000004	−0.00000002
2	−1	2	0.00000000	0.00000000
−2	0	2	−0.00000011	0.00000004
−1	0	2	−0.00000024	−0.00000004
0	0	2	2.24999915	−0.00000025
1	0	2	−0.00000024	−0.00000009
2	0	2	−0.00000011	−0.00000005
−2	1	2	0.00000000	0.00000000
−1	1	2	−0.00000004	−0.00000002
0	1	2	−0.00000037	−0.00000015
1	1	2	−0.00000004	−0.00000002
2	1	2	0.00000000	0.00000000
−2	2	2	0.00000000	0.00000000
−1	2	2	0.00000000	0.00000000
0	2	2	−0.00000015	−0.00000007
1	2	2	0.00000000	0.00000000
2	2	2	0.00000000	0.00000000

8.4 Example 4: Exponential

We now consider a more abrupt nonlinearity, namely a real exponential function that is often used to model a junction diode or bipolar junction transistor (BJT). Let us take

$$y = \phi(x) = 5 \cdot 10^{-15}\, e^{40x}. \tag{8.4.1}$$

This function might represent the emitter current versus base voltage characteristic of a silicon BJT biased in the active region, for with $x = 0.65$ (volts), $\phi(x) \simeq 0.001$ (amps). Pursuing this representation, we will let

$$x(t) = 0.65 + 0.250\ cos\ \omega_1 t + 0.025\ cos\ \omega_2 t, \tag{8.4.2}$$

which is a typical situation when the BJT is used as a mixer (ω_1 would then be the local oscillator frequency and ω_2 the incoming signal frequency). We note that the amplitude of the second cosine is usually considered to be "small signal," while the amplitude of the first cosine most definitely is not. (When a BJT is used as a mixer, we desire the local oscillator signal to be large compared with the incoming signal, to provide good conversion gain.)

Computer calculations were carried out to an accuracy of 10^{-6}, which required a very large number of partitions due to the great number of significant Fourier series terms. In practical circuit analyses, such accuracies are far better than one usually needs; this accuracy was used here to illustrate the power of the MDFT. (Least squares, for example, would find such an accuracy impossible to achieve in practice because of the need to invert a matrix of enormous dimensions.) A particularly interesting feature of our results in Table 8.4.1 is that nearly eight times as many partitions were needed in the ω_1 dimension as in the ω_2 dimension. We predicted this very result in Section 7.2, when we studied the aliasing effect;

since the amplitude of the signal $cos\ \omega_2 t$ is much smaller than the amplitude of the signal $cos\ \omega_1 t$, we expect there to be fewer significant harmonics of ω_2 in $y(t)$ than there are harmonics of ω_1. Consequently, aliasing in the ω_2 dimension will be much less of a problem, and fewer partitions will be needed in that dimension to achieve a given accuracy. This decoupling of dimensions with regard to aliasing is one of the most powerful features of the multiple integral formulation. To illustrate these comments, we have plotted in Figure 8.4.1 the Fourier series coefficients $a(v_1,v_2)$ as functions of v_1 and v_2.

From Table 8.4.1, which again lists only first quadrant results, we see that the d.c. component of emitter current for our BJT is nearly 3.5 amps, although the device was biased for about 1 milliamp. This tremendous shift in the quiescent point occurs because of the large–signal input, which drives the peak base–emitter voltage to over 0.9 volts. In a well–designed practical circuit, the bias network for the transistor tends to stabilize the d.c. emitter current, rather than the d.c. base–emitter voltage, to counteract this effect. We can still extract some useful information from this simulation, however, and at the same time confirm our numerical results.

Following the approach of Clarke and Hess (1971), let us write

$$y(t) = 5 \cdot 10^{-15}\ e^{40(\ 0.65\ +\ 0.250\ cos\ \omega_1 t\ +\ 0.025\ cos\ \omega_2 t\ )}$$

$$= 5 \cdot 10^{-15}\ e^{40(0.65)}\ e^{40(0.250\ cos\ \omega_1 t)}\ e^{40(0.025\ cos\ \omega_2 t)}. \qquad (8.4.3)$$

Each of the last two exponentials in equation (8.4.3) is a periodic function of its exponent, and may be expanded in a Fourier series. The coefficients of these series are well known, and the result becomes

$$y(t) = 5 \cdot 10^{-15}\, e^{40(0.65)}$$
$$\cdot [I_0(\alpha) + I_1(\alpha)\,(e^{j\omega_1 t} + e^{-j\omega_1 t}) + I_2(\alpha)\,(e^{j2\omega_1 t} + e^{-j2\omega_1 t}) +$$
$$I_3(\alpha)\,(e^{j3\omega_1 t} + e^{-j3\omega_1 t}) + \cdots]$$
$$\cdot [I_0(\beta) + I_1(\beta)\,(e^{j\omega_2 t} + e^{-j\omega_2 t}) + I_2(\beta)\,(e^{j2\omega_2 t} + e^{-j2\omega_2 t}) +$$
$$I_3(\beta)\,(e^{j3\omega_2 t} + e^{-j3\omega_2 t}) + \cdots], \tag{8.4.4}$$

where $\alpha = 40 \cdot 0.250 = 10$, $\beta = 40 \cdot 0.025 = 1$, and $I_n(\alpha)$ is a modified Bessel function of the first kind, order n, having argument α. These functions are widely tabulated [see Clarke and Hess (1971), for example], and are also available in numerical analysis packages on large computers. If we use approximate values for the Bessel functions, our output becomes

$$y(t) = 0.000979$$
$$\cdot [2815.7 + 2671.0\,(e^{j\omega_1 t} + e^{-j\omega_1 t}) + 2281.5\,(e^{j2\omega_1 t} + e^{-j2\omega_1 t}) +$$
$$1758.4\,(e^{j3\omega_1 t} + e^{-j3\omega_1 t}) + \cdots]$$
$$\cdot [1.2661 + 0.5652\,(e^{j\omega_2 t} + e^{-j\omega_2 t}) + 0.1358\,(e^{j2\omega_2 t} + e^{-j2\omega_2 t}) +$$
$$0.0222\,(e^{j3\omega_2 t} + e^{-j3\omega_2 t}) + \cdots],$$
$$= 3.4888$$
$$\cdot [1 + 0.9486\,(e^{j\omega_1 t} + e^{-j\omega_1 t}) + 0.8103\,(e^{j2\omega_1 t} + e^{-j2\omega_1 t}) +$$
$$0.6245\,(e^{j3\omega_1 t} + e^{-j3\omega_1 t}) + \cdots]$$
$$\cdot [1 + 0.4464\,(e^{j\omega_2 t} + e^{-j\omega_2 t}) + 0.1072\,(e^{j2\omega_2 t} + e^{-j2\omega_2 t}) +$$
$$0.0175\,(e^{j3\omega_2 t} + e^{-j3\omega_2 t}) + \cdots]. \tag{8.4.5}$$

If we multiply out equation (8.4.5), no two terms will have the same frequency, due to the independence of ω_1 and ω_2. Thus, we may easily compare these results with

those of Table 8.4.1. For example, from equation (8.4.5), we immediately see that the d.c. component of $y(t)$ is 3.4888, the ω_1 component is 3.4888 · 0.9486 · 1 = 3.3094, and the ω_1-ω_2 component is 3.4888 · 0.9486 · 0.4464 = 1.4773; these numbers are in excellent agreement with the computer results.

From equation (8.4.4) we observe that the d.c. component of base–emitter voltage appears only in the leading factor

$$5 \cdot 10^{-15}\, e^{40(0.65)},$$

so that the relative amplitudes of the Fourier series terms depend only on the signal levels and not on the bias point of the transistor. From the remarks made above, the d.c. component of emitter current is

$$5 \cdot 10^{-15}\, e^{40(0.65)}\, I_0(\alpha)\, I_0(\beta),$$

which depends on both the d.c. and a.c. components of base–emitter voltage. Suppose that the bias circuit for our transistor maintained the d.c. component of emitter current perfectly constant, under all signal conditions, at its non–signal value of approximately 1 milliamp. Because $I_0(\alpha)$ and $I_0(\beta)$ are monotonically increasing functions of their arguments (that is, of signal level), the d.c. component of base–emitter voltage would have to be smaller than 0.65 under signal conditions. This decrease of average base–emitter voltage is readily observable in actual large–signal BJT circuits.

Table 8.4.1

Results of Example 4.

Partitions in dimension 1: 336
Partitions in dimension 2: 44

v_1	v_2	Re[a(v_1,v_2)]	Im[a(v_1,v_2)]
0	0	3.48876420	0.00000000
1	0	3.30944104	–0.00000001
2	0	2.82687573	–0.00000002
3	0	2.17869023	–0.00000004
4	0	1.51966080	–0.00000005
5	0	0.96296054	–0.00000006
6	0	0.55669896	–0.00000007
7	0	0.29492022	–0.00000009
8	0	0.14380882	–0.00000010
9	0	0.06482402	–0.00000011
10	0	0.02712324	–0.00000012
0	1	1.55734913	–0.00000003
1	1	1.47730109	–0.00000003
2	1	1.26188880	–0.00000003
3	1	0.97254533	–0.00000003
4	1	0.67836125	–0.00000003
5	1	0.42985587	–0.00000003
6	1	0.24850480	–0.00000004
7	1	0.13164941	–0.00000004
8	1	0.06419481	–0.00000004
9	1	0.02893679	–0.00000005
10	1	0.01210754	–0.00000005
0	2	0.37406522	–0.00000005
1	2	0.35483819	–0.00000005
2	2	0.30309755	–0.00000004
3	2	0.23359912	–0.00000004
4	2	0.16293800	–0.00000003
5	2	0.10324861	–0.00000002
6	2	0.05968925	–0.00000002
7	2	0.03162134	–0.00000001
8	2	0.01541918	–0.00000001
9	2	0.00695043	–0.00000001
10	2	0.00290815	–0.00000001

Table 8.4.1 (continued)

v_1	v_2	Re[a(v_1,v_2)]	Im[a(v_1,v_2)]
0	3	0.06108684	−0.00000008
1	3	0.05794697	−0.00000007
2	3	0.04949744	−0.00000006
3	3	0.03814798	−0.00000005
4	3	0.02660864	−0.00000003
5	3	0.01686105	−0.00000002
6	3	0.00974757	−0.00000001
7	3	0.00516393	−0.00000001
8	3	0.00251803	0.00000000
9	3	0.00113504	0.00000000
10	3	0.00047492	0.00000000
0	4	0.00754204	−0.00000010
1	4	0.00715438	−0.00000010
2	4	0.00611117	−0.00000008
3	4	0.00470991	−0.00000006
4	4	0.00328522	−0.00000004
5	4	0.00208174	−0.00000003
6	4	0.00120348	−0.00000002
7	4	0.00063756	−0.00000001
8	4	0.00031089	0.00000000
9	4	0.00014014	0.00000000
10	4	0.00005864	0.00000000
0	5	0.00074770	−0.00000013
1	5	0.00070927	−0.00000012
2	5	0.00060585	−0.00000010
3	5	0.00046693	−0.00000008
4	5	0.00032569	−0.00000006
5	5	0.00020638	−0.00000004
6	5	0.00011931	−0.00000002
7	5	0.00006321	−0.00000001
8	5	0.00003082	−0.00000001
9	5	0.00001389	0.00000000
10	5	0.00000581	0.00000000

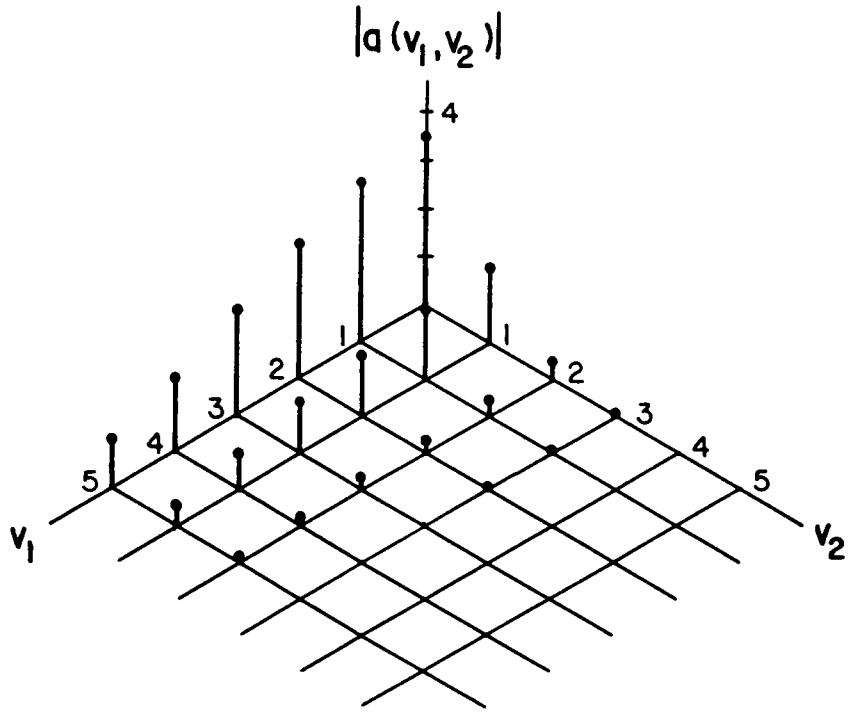

Figure 8.4.1. Plot of the Fourier series coefficients for Example 4.

8.5 Example 5: Exponential with Series Resistance

For our last example, we will modify the nonlinear function of Example 4. Suppose that our hypothetical transistor is no longer ideal, but instead has five ohms of resistance in series with the emitter. Such resistance could be due to the bulk resistance of the semiconductor. Now, we no longer have a simple functional relationship between the emitter current and applied voltage, although the function is still single–valued. We may, however, write the applied voltage x as a function of emitter current y, obtaining

$$x = \phi^{-1}(y) = \frac{1}{40} ln \left[\frac{y}{5 \cdot 10^{-15}} \right] + 5y, \tag{8.5.1}$$

and solve this equation numerically for y, given a value of x. Since the MDFT requires $y = \phi(x)$ only at a finite number of discrete points, we may still use this algorithm to compute the Fourier series coefficients of $y(t)$.

Table 8.5.1 lists our results, for a requested accuracy of 10^{-6}. We note that the d.c. emitter current is now only 9.4 milliamps, compared with nearly 3.5 amps in Example 4, and that the levels of all other components are greatly reduced as well. More importantly, however, the introduction of the emitter feedback resistor has resulted in a much more linear amplification in the transistor. For example, the ratio of emitter currents for the ω_1-ω_2 and d.c. terms in Example 4 is 1.4773/3.4888 = 0.4234; in Example 5, this ratio is only 0.0006325/0.009404 = 0.06727. In an amplifier, a component having frequency ω_1-ω_2 is undesirable, for it represents distortion. Conversely, it is usually the desired output term of a mixer. Thus we observe that emitter feedback is beneficial in building a linear amplifier, but detrimental in building a mixer. We have plotted the Fourier series coefficients

$a(v_1,v_2)$ in Figure 8.5.1. Comparison of this graph with that of Figure 8.4.1 shows quite clearly the linearizing effect of emitter resistance in a BJT.

Table 8.5.1

Results of Example 5.

Partitions in dimension 1: 84
Partitions in dimension 2: 22

v_1	v_2	Re[a(v_1,v_2)]	Im[a(v_1,v_2)]
0	0	0.00940387	0.00000000
1	0	0.00760275	0.00000000
2	0	0.00374217	0.00000000
3	0	0.00066449	0.00000000
4	0	-0.00037952	0.00000000
5	0	-0.00018698	0.00000000
6	0	0.00007698	0.00000000
7	0	0.00006324	0.00000000
8	0	-0.00001841	0.00000000
9	0	-0.00002298	0.00000000
10	0	0.00000441	0.00000000
0	1	0.00089640	0.00000000
1	1	0.00063254	0.00000000
2	1	0.00013676	0.00000000
3	1	-0.00011822	0.00000000
4	1	-0.00006478	0.00000000
5	1	0.00003604	0.00000000
6	1	0.00003158	0.00000000
7	1	-0.00001193	0.00000000
8	1	-0.00001542	0.00000000
9	1	0.00000383	0.00000000
10	1	0.00000746	0.00000000
0	2	0.00001998	0.00000000
1	2	0.00000638	0.00000000
2	2	-0.00001167	0.00000000
3	2	-0.00000732	0.00000000
4	2	0.00000614	0.00000000
5	2	0.00000573	0.00000000
6	2	-0.00000296	0.00000000
7	2	-0.00000389	0.00000000
8	2	0.00000129	0.00000000
9	2	0.00000244	0.00000000
10	2	-0.00000048	0.00000000

Table 8.5.1 (continued)

v_1	v_2	Re[a(v_1,v_2)]	Im[a(v_1,v_2)]
0	3	−0.00000006	0.00000000
1	3	−0.00000032	0.00000000
2	3	−0.00000027	0.00000000
3	3	0.00000046	0.00000000
4	3	0.00000047	0.00000000
5	3	−0.00000037	0.00000000
6	3	−0.00000049	0.00000000
7	3	0.00000023	0.00000000
8	3	0.00000042	0.00000000
9	3	−0.00000012	0.00000000
10	3	−0.00000032	0.00000000
0	4	0.00000000	0.00000000
1	4	0.00000000	0.00000000
2	4	0.00000001	0.00000000
3	4	0.00000001	0.00000000
4	4	−0.00000002	0.00000000
5	4	−0.00000003	0.00000000
6	4	0.00000002	0.00000000
7	4	0.00000004	0.00000000
8	4	−0.00000002	0.00000000
9	4	−0.00000004	0.00000000
10	4	0.00000001	0.00000000
0	5	0.00000000	0.00000000
1	5	0.00000000	0.00000000
2	5	0.00000000	0.00000000
3	5	0.00000000	0.00000000
4	5	0.00000000	0.00000000
5	5	0.00000000	0.00000000
6	5	0.00000000	0.00000000
7	5	0.00000000	0.00000000
8	5	0.00000000	0.00000000
9	5	0.00000000	0.00000000
10	5	0.00000000	0.00000000

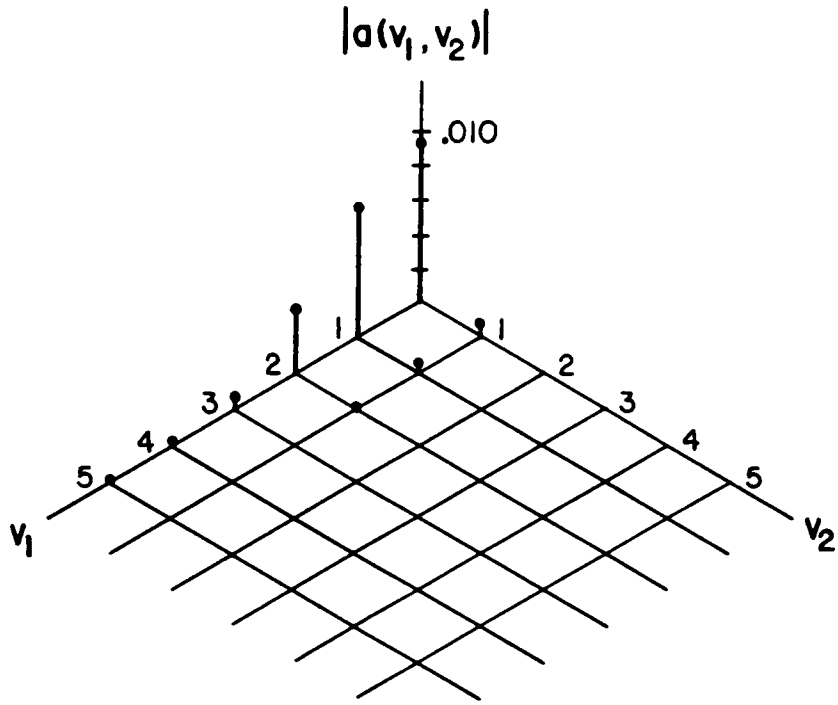

Figure 8.5.1. Plot of the Fourier series coefficients for Example 5.

CHAPTER 9
Conclusions

9.1 Summary

In this dissertation, we set out to develop an efficient means of analyzing memoryless nonlinear functions in the frequency domain. We first defined the problem and looked at existing solutions. Next, we examined the basic theory of almost periodic functions, and proved that the output of a memoryless nonlinearity is almost periodic whenever the input is. We continued by defining the concept of a set of basis frequencies, and we proved that the frequencies at the input and output of the nonlinearity have the same basis. We also considered the notion of a Fourier series for almost periodic functions, and the convergence and summability properties of such series.

We then proved the theorem upon which our method is based, namely, that the normalized phases corresponding to a set of basis frequencies have their fractional parts uniformly distributed in the unit cube. Using this result, we converted the single integral form of the Fourier series coefficients into a multiple integral over finite limits. Finally, we considered a numerical implementation of the multiple integral, which we called the Multiple Discrete Fourier Transform, and looked at five simple examples of its use.

9.2 Results

In Chapter 1, the author set forth the requirements on a good method for frequency domain analysis of a memoryless nonlinearity. It would now be appropriate to review how our new method meets these requirements.

As desired, the multiple integral formulation handles any type of memoryless nonlinearity that can be represented as a continuous, bounded, real–valued function. Because there is no need to expand the nonlinear function in a power series, gross nonlinearities (which have many terms in their power series approximations) pose no particular difficulty. A variety of mathematical descriptions of the nonlinear function are acceptable, including analytical, piecewise, or even pointwise, since computationally all that we must do is to generate an output value for each input value over the domain of the function.

The multiple integral formulation allows the input of the nonlinearity to be any finite sum of bounded, periodic functions. The frequencies of these functions need not be harmonically related, and as a result, the frequencies in the output of the nonlinearity need not be harmonically related either. Both the input and output frequencies can be expressed as linear combinations, using integer coefficients, from the same finite set of basis frequencies. Being independent over the integers, these basis frequencies do not appear explicitly in the multiple integral formulation, and their numerical values are arbitrary. Thus, the computational efficiency of the multiple integral formulation remains good even when the basis frequencies are very closely spaced or very widely spaced (situations in which the Discrete Fourier Transform encounters difficulty). Further, an input frequency that is independent of the remaining input frequencies may assume any numerical value without

increasing the computational effort. In the event that an input frequency is not independent, but is very close to or very far from the other input frequencies, it might be possible to approximate the former as independent to take advantage of the improved computational efficiency.

The multiple integral formulation of the Fourier series coefficients is exact. To carry out the calculations numerically, any of a variety of multi-dimensional integration schemes can be used. The simple scheme presented here, the Multiple Discrete Fourier Transform (MDFT), allows for iterative improvement and error checking. The MDFT can be written in the form of a matrix multiplied by a column vector of samples of the nonlinearity output. Thus, once the elements of the matrix are computed and stored, the MDFT may be applied repeatedly for different nonlinearity outputs by carrying out a single matrix multiplication, provided that the set of basis frequencies for the output does not change. In contrast to the least squares method, the MDFT allows for iterative improvement, while it does not require the inversion of a large matrix, it does not require any assumption regarding the form of the nonlinearity output, and it does not require the inclusion of all significant Fourier series terms just to find the coefficients of a few desired ones. The MDFT is also efficient relative to the Discrete Fourier Transform whenever the nonlinearity input and output contain frequencies that are not harmonically related; in the case of purely periodic inputs and outputs, it reduces to the DFT.

9.3 Further Research

The multiple integral formulation of the Fourier series coefficients, because of its elegance and theoretical soundness, provides many opportunities for further

research. Some suggested areas are the following:

The multiple integral formulation was derived under the restriction that the nonlinearity be described by a continuous, bounded function. It would be useful to remove this continuity restriction and replace it with a weaker requirement, for example by requiring only that the function be of bounded variation.

Stronger results are needed regarding the convergence of the Fourier series for almost periodic functions, if only for special cases. Whether or not such results even exist is apparently still an open question.

It would be beneficial to investigate error prediction and correction methods for use with the MDFT; for example, the single–dimensional DFT possesses a Richardson extrapolation algorithm, and there may be a multi–dimensional extension of this algorithm. It would also be beneficial to investigate methods other than the MDFT for evaluating the multiple integral formulation, to see whether they offer any computational advantages. Cizek (1986) discusses various numerical integration schemes for finding the Fourier series coefficients of a periodic function, and it may be possible to extend his work to multiple dimensions.

To be a useful electric circuit analysis tool, the multiple integral formulation must, of course, be included in a complete frequency–domain analysis algorithm. Probably the most versatile algorithm is harmonic balance, introduced into the literature by Nakhla and Vlach (1976). Since then, a number of articles have appeared [for example, Ushida and Chua (1984), Gilmore (1986), Gayral et al. (1987), and Gilmore (1988)] concerning the application of harmonic balance to practical problems. Although the basic scheme is rather straightforward, there is still controversy in the literature regarding the best analysis method for the nonlinear circuit elements. Our work here in developing the multiple integral

formulation addressed this very question, and a logical extension of this work is to incorporate the multiple integration formulation into the harmonic balance algorithm.

The multiple integral formulation for the Fourier series coefficients is not limited to applications with harmonic balance. It could also be used with some of the approximate circuit analysis methods of Clarke and Hess (1971), and it might well find applications in areas outside of electric circuit analysis and in disciplines other than electrical engineering.

REFERENCES

Antonov, O. E., and V. S. Ponkratov (1974), Further discussion of the action of the sum of harmonic oscillations on nonlinear elements, *Radio Engineering and Electronic Physics, 19*, 2, 44–53.

Ash, J. Marshall (1976), Multiple trigonometric series, in *Studies in Harmonic Analysis, MAA Studies in Mathematics Volume 13*, Ed., J. Marshall Ash, The Mathematical Association of America, U.S.A.

Ash, J. Marshall, and Grant V. Welland (1972), Convergence, uniqueness, and summability of multiple trigonometric series, *Trans. American Mathematical Society, 163*, 401–436.

Beckmann, Petr (1973), *Orthogonal Polynomials for Engineers and Physicists*, The Golem Press, Boulder, Colorado.

Bennett, W. R. (1933), New results in the calculation of modulation products, *Bell System Technical Journal, 12*, 228–243.

Besicovitch, A. S., and Harald A. Bohr (1931), Almost periodicity and general trigonometric series, *Acta Mathematica, 57*, 203–292.

Bhatkar, V. P., and S. R. Atre (1970), Intermodulation response of exponential-law devices, *Proc. IEEE* (Lett.), *58*, 6, 951–952.

Bohr, Harald A. (1951), *Almost Periodic Functions*, Chelsea Publishing Company, New York [originally published in 1933 by Julius Springer, Berlin].

Camacho-Peñalosa, Carlos (1983), Numerical steady-state analysis of nonlinear microwave circuits with periodic excitation, *IEEE Trans. Microwave Theory and Techniques, MTT-31*, 9, 724–730.

Cassels, John William Scott (1957), *An Introduction to Diophantine Approximation, Cambridge Tracts in Mathematics and Mathematical Physics No. 45*, Cambridge University Press, Cambridge, England.

Chua, Leon O., and Akio Ushida (1981), Algorithms for computing almost periodic steady-state response of nonlinear systems to multiple input frequencies, *IEEE Trans. Circuits and Systems, CAS-28*, 10, 953–971.

Cizek, Vaclav (1986), *Discrete Fourier Transforms and Their Applications*, Adam Hilger Ltd., Bristol, England.

Clarke, Kenneth K., and Donald T. Hess (1971), *Communication Circuits: Analysis and Design*, Addison-Wesley Publishing Company, Reading, Massachusetts.

Corduneanu, C. (1968), *Almost Periodic Functions*, Interscience Publishers, New York.

Gardiner, J. G. (1968), The relationship between cross-modulation and intermodulation distortions in the double-balanced modulator, *Proc. IEEE* (Lett.), *56*, 11, 2069–2071.

Gayral, M., et al. (1987), The spectral balance: a general method for analysis of nonlinear microwave circuits driven by non-harmonically related generators, *IEEE MTT-S International Microwave Symp. Digest*, 119–121.

Geckinli, Nezih C., and Davras Yavuz (1983), *Discrete Fourier Transformation and Its Applications to Power Spectral Estimation*, Elsevier Scientific Publishing Company, Amsterdam, Netherlands.

Gilmore, Rowan J. (1986), Nonlinear circuit design using the modified harmonic balance algorithm, *IEEE Trans. Microwave Theory and Techniques*, *MTT-34*, 12, 1294–1307.

Gilmore, Rowan J. (1988), Design in the nonlinear world—the method of harmonic balance, *Proceedings RF Technology Expo 88*, 447–456.

Gretsch, W. R. (1966), The spectrum of intermodulation generated in a semiconductor diode junction, *Proc. IEEE*, *54*, 11, 1528–1535.

Hicks, Ross G., and Peter J. Khan (1982), Numerical analysis of nonlinear solid-state device excitation in microwave circuits, *IEEE Trans. Microwave Theory and Techniques*, *MTT-30*, 3, 251–259.

Nakhla, Michael S., and Jiri Vlach (1976), A piecewise harmonic balance technique for determination of periodic response of nonlinear systems, *IEEE Trans. Circuits and Systems*, *CAS-23*, 2, 85–91.

Niven, Ivan Morton (1963), *Diophantine Approximations*, *Interscience Tracts in Pure and Applied Mathematics Number 14*, Interscience Publishers, New York.

Papoulis, Athanasios (1965), *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill Book Company, New York.

Rees, Charles Sparks, S. M. Shah, and C. V. Stanojevic (1981), *Theory and Applications of Fourier Analysis*, Marcel Dekker, Inc., New York.

Roberts, Fred S. (1984), *Applied Combinatorics*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Sorkin, Gregory B., Kenneth S. Kundert, and Alberto Sangiovanni-Vincentelli (1987), An almost-periodic Fourier Transform for use with harmonic balance, *IEEE MTT-S International Microwave Symp. Digest*, 717–720.

Stoodley, L. G. (1964), Intermodulation analysis of crystal mixer, *Proc. IEEE* (Lett.), *52*, 8, 784.

Surana, D. C., and J. G. Gardiner (1971), Multiple Fourier series analysis for mixers and modulators, *Proc. IEEE* (Lett.), *59*, 11, 1627–1628.

Ushida, Akio, and Leon O. Chua (1984), Frequency–domain analysis of nonlinear circuits driven by multi–tone signals, *IEEE Trans. Circuits and Systems*, *CAS–31*, 9, 766–779.

Weiner, Donald D., and John F. Spina (1980), *Sinusoidal Analysis and Modeling of Weakly Non–Linear Circuits*, Van Nostrand Reinhold Company, New York.

The two page vita has been removed from the scanned document. Page 1 of 2

The two page vita has been removed from the scanned document. Page 2 of 2