The FFT
Fundamentals and Concepts

Robert W. Ramirez

Tektronix, Inc.

PRENTICE-HALL, INC./ Englewood Cliffs, N.J. 07632

Library of Congress Cataloging in Publication Data

RAMIREZ, ROBERT W. (date)
The FFT, fundamentals and concepts.

Bibliography: p.
Includes index.
1. Fourier analysis. 2. Fourier analysis—Data processing. 3. Fourier transformations. 4. Fourier transformations—Data processing. I. Title.
QA403.5.R36 1985 515'.2433 84-8284
ISBN 0-13-314386-4

Editorial/production supervision and interior design: *Sylvia Schmokel*
Cover design: *20/20 Services, Inc.*
Manufacturing buyer: *Gordon Osbourne*

Printed in the United States of America

10 9 8 7 6 5 4 3 2

ISBN 0-13-314386-4 01

PRENTICE-HALL INTERNATIONAL, INC., *London*
PRENTICE-HALL OF AUSTRALIA PTY. LIMITED, *Sydney*
EDITORA PRENTICE-HALL DO BRASIL, LTDA., *Rio de Janeiro*
PRENTICE-HALL CANADA INC., *Toronto*
PRENTICE-HALL OF INDIA PRIVATE LIMITED, *New Delhi*
PRENTICE-HALL OF JAPAN, INC., *Tokyo*
PRENTICE-HALL OF SOUTHEAST ASIA PTE. LTD., *Singapore*
WHITEHALL BOOKS LIMITED, *Wellington, New Zealand*

Preface

Fourier analysis is not a new subject. It has been around since the early 1800s when J.B.J. Fourier developed the initial concepts and theory. Since then numerous papers and books dealing with Fourier theory have been published, and the Fourier series and integral have found their way into various college curricula.

So why another book on Fourier theory? Fourier analysis exists in a different context today. It used to be a pencil-and-paper issue, an interesting mathematical approach to getting frequency-domain information, but generally too difficult to apply in most practical cases. And even with the arrival of the digital computer, useful Fourier analyses were too time-consuming and computer-expensive for widespread use. Then, in the 1960s, J W. Cooley and J.W. Tukey published *An Algorithm for the Machine Calculation of Complex Fourier Series*. Their algorithm became known as the *fast Fourier transform*, or *FFT*, and has become the new context for Fourier analysis. This is not just a digital context either, but a new context that allows quick, economical application of Fourier techniques to a wide variety of measurement and analysis tasks.

Thus the FFT is becoming a general analysis tool. FFT routines are found in most comprehensive software libraries, and FFT analyzers are becoming a more frequently encountered item. But even more than that, the FFT has joined forces with general-purpose instrumentation.

Today, instrument manufacturers are offering a variety of waveform digitizers, the most common type being what is generally referred to as a *digitizing oscilloscope*. These waveform digitizers are usually operated in conjunction with a minicomputer or desk-top calculator and a software library that often includes an FFT algorithm. The result is that Fourier analysis, as well as convolution and correlation, has been taken out of the textbook and put on the engineering bench. Because of its usefulness and increasing availability, it is expected that the FFT will become a major and commonplace measurement tool.

There is still, however, a missing link in the chain of events leading to general-purpose use of the FFT. That missing link is general familiarity with Fourier theory. You don't have to know all the details of the equations and their derivations. But you do need to know the concepts that they embody. To successfully use the FFT as a measurement tool, you do need to know what to expect in the frequency domain and how digital techniques affect the frequency domain.

For the most part, these concepts can be demonstrated through simple diagrams and pictures and can be discussed in simple terms. That is the approach taken in the following pages. Part I introduces classical Fourier theory with a slant toward later discussions of digital implementations. Part II covers the digital approach to Fourier analysis and makes heavy use of a waveform digitizer and an FFT algorithm to provide specific examples. Every effort has been made throughout to illustrate fully each concept and to discuss each concept in easy-to-understand terms. Part III provides a brief look at an FFT algorithm.

Why another book on Fourier theory?—to bridge the gap between classroom theory and practical use, and to do it in a language that people of different backgrounds and technical levels can understand.

ROBERT W. RAMIREZ

Acknowledgments

Though many solitary hours are spent putting words to paper, one can never truly write a book alone. Accordingly, I extend my appreciation to Tektronix, Inc., for providing the creative atmosphere as well as access to the instruments and software necessary for developing much of the material in this book. Also, I would particularly like to thank Lyle Ochs of Tektronix for his many suggestions. And, finally, my deep appreciation to my wife, Barbara, for her patience and encouragement throughout this project.

Contents

Part I

INTRODUCTION TO FOURIER ANALYSIS

There are two chapters in Part I. The first is but a few pages—just enough to start you thinking about time and frequency as two related concepts. Though the relationship may not be intuitively obvious, an important relationship does exist. This relationship becomes more obvious in the second chapter, where the Fourier transform is explored. A good grasp of the concepts covered in these first two chapters is necessary to understand the digital analysis techniques covered in Part II.

Chapter 1

Time and Frequency: Two Bases of Description

Time. This is one of the fundamental concerns of people. How often during the day do we look at a clock or check our watches? Since birth, our lives have been geared to time. There is a time to wake up, a time to eat, a time to work, a time to play, and a time to go to sleep. We measure each day of our lives in time and use it to order the events that concern and affect us.

Time is universal. All people recognize its passage. All people live by it. Time, in itself, is central to many philosophical questions: Does time flow by us, or do we advance through time? And the measurement of time is an established science (horology) with a long history.

As far back as 3500 B.C., people were erecting poles and towers to cast shadows, the length of the shadow being an indication of the time of day. By the eighth century B.C., the Egyptians had refined this shadow concept to a fairly accurate sundial. They also developed water clocks to substitute during the night and on cloudy days. Later, the Romans and Greeks refined these devices further. But it wasn't until the fourteenth century A.D., that anything resembling a modern timepiece was developed. Then, in 1582, Galileo observed the constancy of a pendulum, and Christian Huygens, in 1665, incorporated Galileo's observation into the first pendulum clock. Until the advent of electrically driven clocks, the pendulum clock was the most accurate timepiece available. Now, by the 1967 agreement of the International Conference on Weights and Measures, the atomic clock is the ultimate standard.

TIME HISTORIES NEED TIME BASES

Today, one second is equal to 9,192,631,770 transitions between two specified, hyperfine levels of the cesium 133 atom.

But why so much precision in measuring the passage of time? The answer: Science demands it. A great deal of scientific theory is couched in terms of time histories. Furthermore, experimental proof of these theories requires time-domain measurements, and the precision of these measurements depends upon our ability to measure time.

As an example, Galileo reportedly used his own pulse as a timepiece in making his original pendulum observations. Each complete swing of a large pendulum took so many heart beats. With no greater precision than that, it was a natural experimental conclusion to say that a pendulum always shows the same simple harmonic motion. But theory tells us that this is not the case. In fact, for large displacements, the time for a complete swing of a pendulum is greater than for small displacements (Fig. 1–1). Proving this experimentally, however, requires a more precise timepiece than what Galileo had access to.

An electronic oscillator is in many ways analogous to a mechanical pendulum. The output of a sine-wave oscillator has a time history that closely resembles the time history of a pendulum's angular displacement. Galileo's concept of counting pulse beats can also be applied to measure the time for a complete voltage swing in an oscillator. The modern version of this concept is used in frequency counters. However, as shown in Fig. 1–2, an electronic pulse is used as a time base instead of a human pulse.

There is a time base involved in all time-domain measurements. In the case of Fig. 1–2b, the time base is used directly to measure the period of the signal. In other types of measurements, the time base is used to generate a time axis for an amplitude history. The precisely controlled speed of the paper drive for a chart recorder is an example of this latter case. Another example is the oscilloscope, which uses a ramp voltage to drive an electron beam at a constant rate across the face of a CRT (cathode ray tube). In both examples, the event or activity being captured for observation drives the pen or CRT trace in a direction normal to the time-base drive. The result is a time history of amplitude variation, as shown in the CRT photo of Fig. 1–3.

At this point, it might be well to pause and examine the CRT photo of Fig. 1–3 in a little more detail. This examination may seem trivial at first. But then it is surprising how much insight can be gained by starting with the seemingly trivial aspects of a subject. It is also surprising—no, embarrassing—to think of the number of measurements that get tripped up by trivia. So let's get on with the examination, which can justly be described as "attention to detail."

Referring to the CRT photo in Fig. 1–3, it is conventional to assign time zero to the left side of the display. Then, according to the time-base setting, time proceeds to the right. In Fig. 1–3, the CRT readout in the upper right portion indicates the time increments for each major division of the display. Vertical amplitude scaling

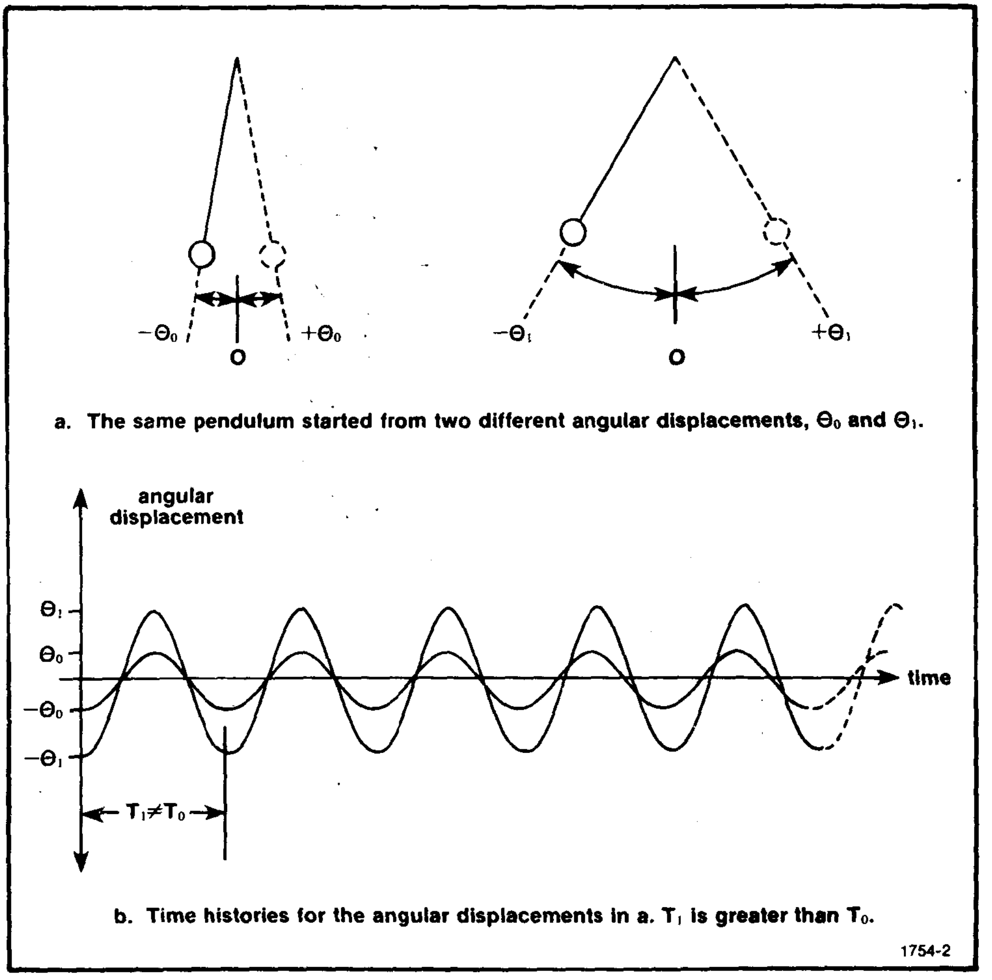


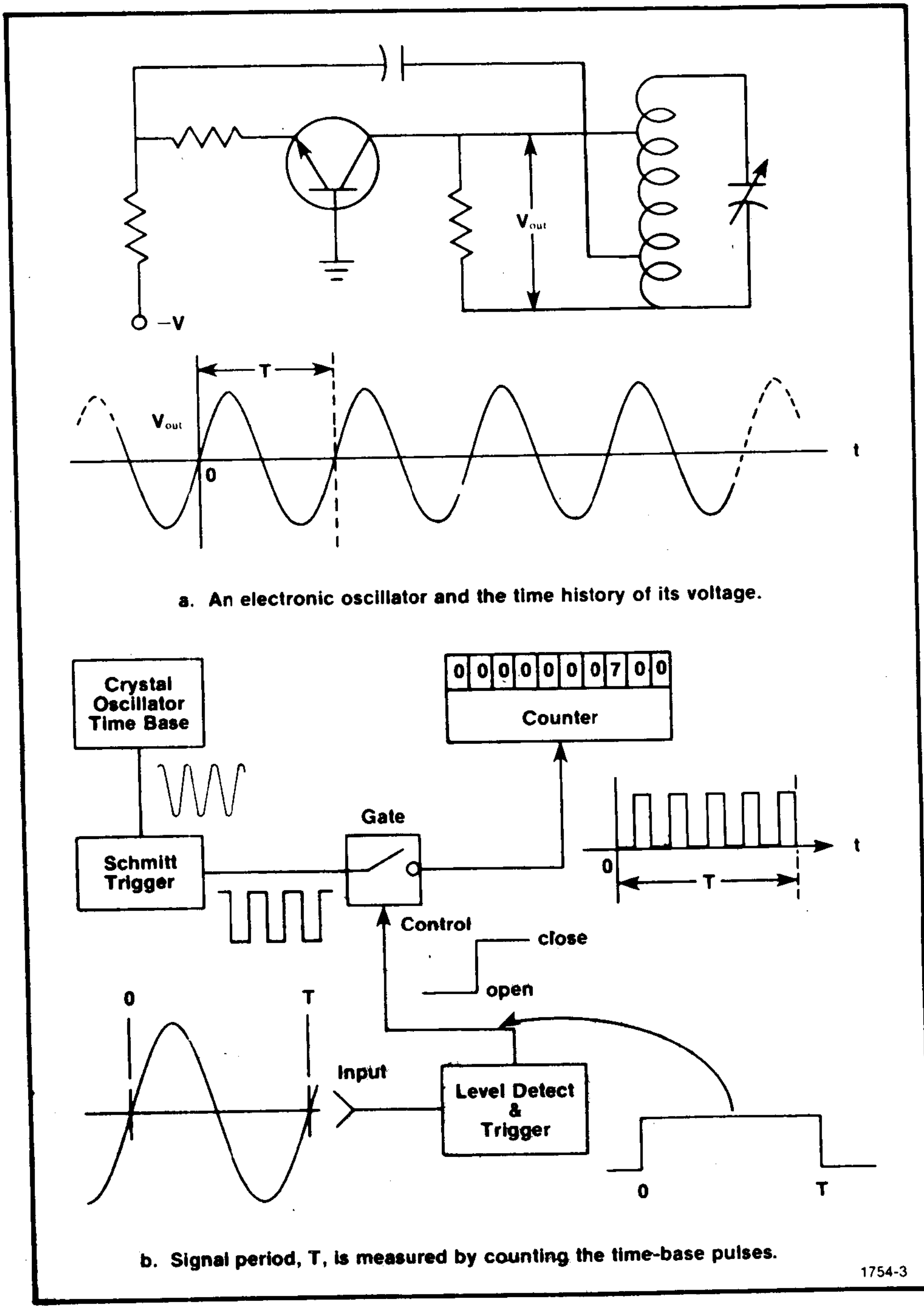
Figure 1–1 The period of a pendulum, T, varies according to its angular displacement, θ.

is indicated by the readout in the upper left corner. But what is the vertical amplitude referenced to?

In the case of Fig. 1–3, which is a photo of a waveform captured by a digitizing oscilloscope after a zero-referencing operation, the reference is indicated by the 0 DIV in the lower right corner. Here, the 0 DIV refers to the vertical center division as being the vertical zero reference. Other possibilities for zero reference might be above (for example, 3 DIV) or below (for example, −3 DIV) center.

With these three things defined—vertical and horizontal scale factors and a zero reference—the value of any point on the displayed waveform is defined. However, this is still not enough to fully define or describe a waveform.

To fully describe a waveform in time, it must not only be possible to pick off



a. An electronic oscillator and the time history of its voltage.

b. Signal period, T, is measured by counting the time-base pulses.

Figure 1–2 Electronic version of Galileo's pendulum observations.

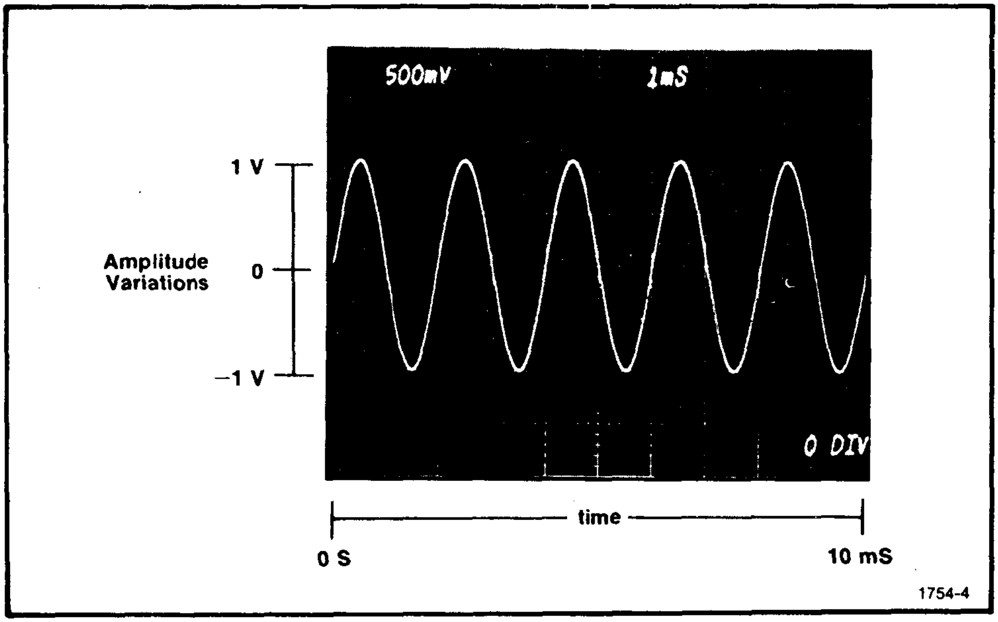


Figure 1–3 Time history of sinusoidal amplitude variations obtained with an oscilloscope.

its values at various points, but there must also be enough of the waveform displayed to discern its shape or type. For example, imagine the display if the sinusoid in Fig. 1–3 had been captured with a time-base setting of only 0.01 msec per division. That would stretch just the first tenth of a time division to cover the entire display. The waveform would appear to be a ramp instead of a sinusoid!

However, the CRT photo of Fig. 1–3 does give a complete time-domain description of a sinusoid. It is complete because it is fully scaled in time and at least one full repetition of the waveform is displayed. From this it can be assumed that the waveform is sinusoidal, at least within the bounds of the display area. What happens outside the display is not recorded and, therefore, is actually undefined. In the case of Fig. 1–3, however, experience and common sense lead us to assume continuation of the waveform in the same manner beyond the confines of the display. And so, we have a time history of a sinusoid.

SINUSOIDS LOOK DIFFERENT FROM A FREQUENCY VIEWPOINT

Once a sinusoid is completely described with respect to time, you can construct a new description of it with respect to frequency. This is shown in Fig. 1–4.

Figure 1–4 depicts a three-dimensional waveform space with amplitude as one axis and time and frequency as the other two axes. The time and amplitude axes

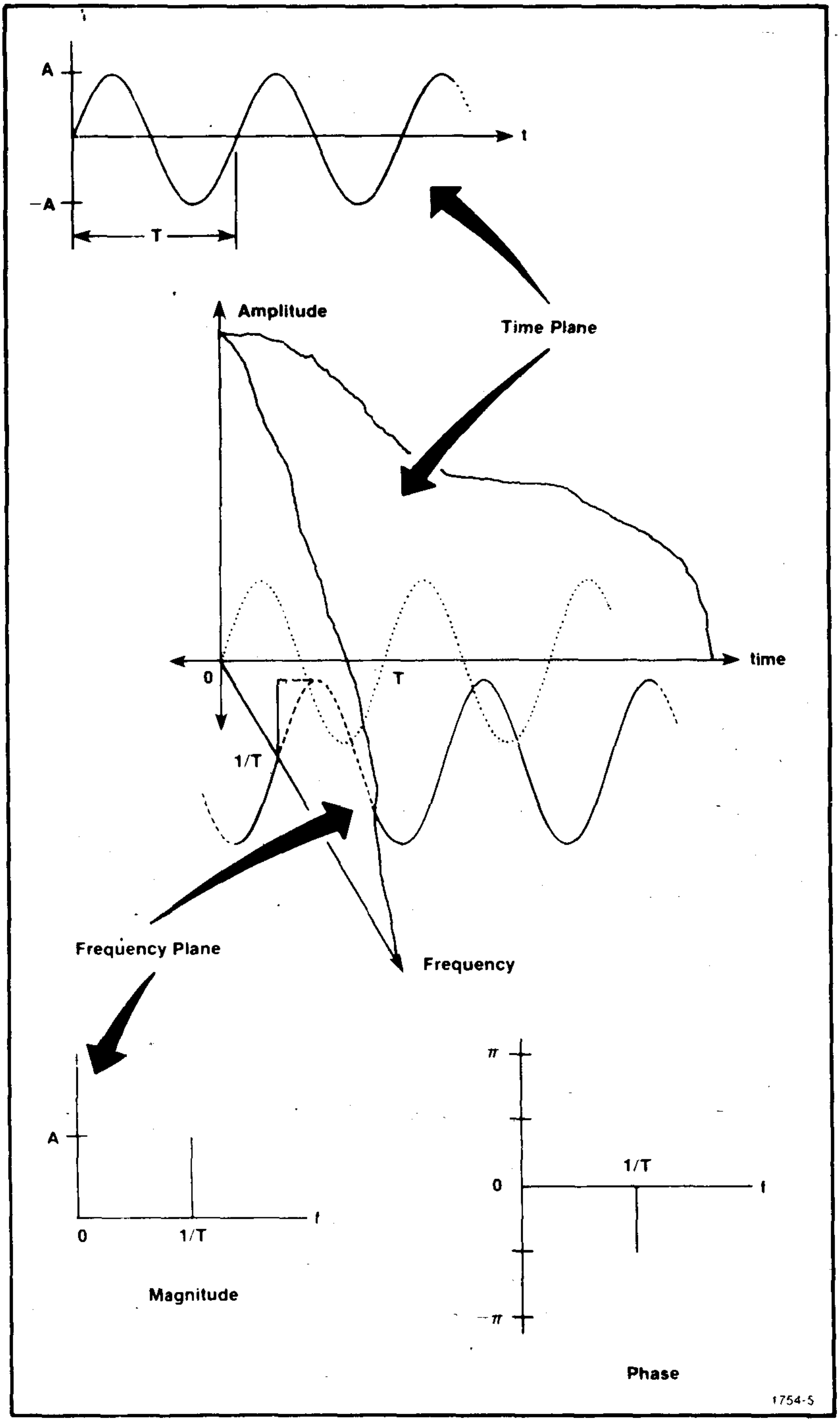


Figure 1–4 Time and frequency description of a sine wave.

define something that can be called a *time plane*. In the same manner, the frequency and amplitude axes define a *frequency plane* that is normal to the time plane.

The time history of a sinusoid, such as that in Fig. 1–3, can be treated as a projection on the time plane. In concept, the sinusoid can be thought of as actually existing at some distance from the time plane. This distance is measured along the frequency axis and is equal to the reciprocal of the waveform period.

Similarly, the sinusoid also has a projection onto the frequency plane. This projection takes the form of an impulse (a pulse of instantaneous rise and fall) with an amplitude equal to the sinusoid's amplitude. Because of symmetry, it is necessary to project only the peak amplitude rather than the full peak-to-peak swing. This is shown in Fig. 1–4 by the positive amplitude impulse on the magnitude diagram. The position of this impulse on the frequency axis coincides with the frequency of the sinusoid. (For now, just consider an impulse to be a line.)

The single impulse in the magnitude diagram defines both the amplitude and frequency of the sinusoid. With only this information, the sinusoid can be reconstructed in the time domain. Some additional information is needed, however, to fix the sinusoid's position relative to the zero time reference. This additional information is provided by a phase diagram, which also consists of an impulse located on a frequency axis. The amplitude of this latter impulse indicates the amount of phase associated with the sinusoid.

Phase diagrams for sinusoids can be determined by looking at the positive peak closest to time zero. For the case of Fig. 1–4, the positive peak occurs after time zero by an amount equal to one-fourth the period. There are 360° in a cycle or period, and the peak is shifted by one-fourth of this. So the phase in Fig. 1–4 is 360°/4, or 90°. Since the positive peak occurs after time zero, the sinusoid is said to be *delayed*. As a matter of convention, delay is denoted by negative phase. If the closest positive peak had been located before time zero, then the sinusoid would have been advanced. An advance is denoted by positive phase. The conventions are further illustrated in Fig. 1–5, and more examples are provided in Fig. 1–6.

In looking at Fig. 1–6, it should be pointed out that the total range of shift is −180° to +180°, or 360°. With no reference point fixed to the sinusoid, an actual shift out of the $360° = 2\pi$ range corresponds to a shift within the 2π range. For example, a sinusoid advanced by $360° + 90° = 450°$ is not generally distinguishable from the same sinusoid advanced by just 90°, so it can be represented as having just a 90° shift. This system of representing phase within a 2π range is referred to as *modulo* 2π *phase*. If on the other hand, a reference can be attached to the sinusoid, then shifts beyond the 2π range can be represented as such. This latter approach is referred to as a continuous phase representation and is detailed later in Part II.