PRINCIPLES AND APPLICATIONS IN ENGINEERING SERIES

# ELECTRICAL MEASUREMENT, SIGNAL PROCESSING, and DISPLAYS

Edited by

JOHN G. WEBSTER

CRC PRESS

# ELECTRICAL MEASUREMENT, SIGNAL PROCESSING, and DISPLAYS

# ELECTRICAL MEASUREMENT, SIGNAL PROCESSING, and DISPLAYS

Edited by
JOHN G. WEBSTER

CRC PRESS

Boca Raton London New York Washington, D.C.

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742


Version Date: 20131104

International Standard Book Number-13: 978-0-203-00940-6 (eBook - PDF)



**Visit the Taylor & Francis Web site at**
**http://www.taylorandfrancis.com**

**and the CRC Press Web site at**
**http://www.crcpress.com**

# Preface

## Introduction

The purpose of *Electrical Measurement, Signal Processing, and Displays* is to provide a reference that is both concise and useful for engineers in industry, scientists, designers, managers, research personnel and students, as well as many others who have measurement problems. The book covers an extensive range of topics that comprise the subject of measurement, instrumentation, and sensors.

The book describes the use of instruments and techniques for practical measurements required in electrical measurements. It includes sensors, techniques, hardware, and software. It also includes information processing systems, automatic data acquisition, reduction and analysis and their incorporation for control purposes.

Chapters include descriptive information for professionals, students, and workers interested in measurement. Chapters include equations to assist engineers and scientists who seek to discover applications and solve problems that arise in fields not in their specialty. They include specialized information needed by informed specialists who seek to learn advanced applications of the subject, evaluative opinions, and possible areas for future study. Thus, *Electrical Measurement, Signal Processing, and Displays* serves the reference needs of the broadest group of users — from the advanced high school science student to industrial and university professionals.

## Organization

The book is organized according to the measurement problem. Section I covers electromagnetic variables measurement such as voltage, current, and power. Section II covers signal processing such as amplifiers, filters, and compatibility. Section III covers displays such as cathode ray tubes, liquid crystals, and plasma displays.

**John G. Webster**
*Editor-in-Chief*

# Editor-in-Chief

**John G. Webster** received the B.E.E. degree from Cornell University, Ithaca, NY, in 1953, and the M.S.E.E. and Ph.D. degrees from the University of Rochester, Rochester, NY, in 1965 and 1967, respectively.

He is Professor of Electrical and Computer Engineering at the University Wisconsin-Madison. In the field of medical instrumentation he teaches undergraduate and graduate courses, and does research on RF cardiac ablation and measurement of vigilance.

He is author of *Transducers and Sensors*, an IEEE/EAB Individual Learning Program (Piscataway, NJ: IEEE, 1989). He is co-author, with B. Jacobson, of *Medicine and Clinical Engineering* (Englewood Cliffs, NJ: Prentice-Hall, 1977), with R. Pallás-Areny, of *Sensors and Signal Conditioning* (New York: Wiley, 1991), and with R. Pallás-Areny, of *Analog Signal Conditioning* (New York: Wiley, 1999). He is editor of *Encyclopedia of Medical Devices and Instrumentation* (New York: Wiley, 1988), *Tactile Sensors for Robotics and Medicine* (New York: Wiley, 1988), *Electrical Impedance Tomography* (Bristol, UK: Adam Hilger, 1990), *Teaching Design in Electrical Engineering* (Piscataway, NJ: Educational Activities Board, IEEE, 1990), *Prevention of Pressure Sores: Engineering and Clinical Aspects* (Bristol, UK: Adam Hilger, 1991), *Design of Cardiac Pacemakers* (Piscataway, NJ: IEEE Press, 1995), *Design of Pulse Oximeters* (Bristol, UK: IOP Publishing, 1997), *Medical Instrumentation: Application and Design, Third Edition* (New York: Wiley, 1998), and *Encyclopedia of Electrical and Electronics Engineering* (New York, Wiley, 1999). He is co-editor, with A.M. Cook, of *Clinical Engineering: Principles and Practices* (Englewood Cliffs, NJ: Prentice-Hall, 1979) and *Therapeutic Medical Devices: Applications and Design* (Englewood Cliffs, NJ: Prentice-Hall, 1982), with W.J. Tompkins, of *Design of Microcomputer-Based Medical Instrumentation* (Englewood Cliffs, NJ: Prentice-Hall, 1981) and *Interfacing Sensors to the IBM PC* (Englewood Cliffs, NJ: Prentice-Hall, 1988, and with A.M. Cook, W.J. Tompkins, and G.C. Vanderheiden, *Electronic Devices for Rehabilitation* (London: Chapman & Hall, 1985).

Dr. Webster has been a member of the IEEE-EMBS Administrative Committee and the NIH Surgery and Bioengineering Study Section. He is a fellow of the Institute of Electrical and Electronics Engineers, the Instrument Society of America, and the American Institute of Medical and Biological Engineering. He is the recipient of the AAMI Foundation Laufman-Greatbatch Prize and the ASEE/Biomedical Engineering Division, Theo C. Pilkington Outstanding Educator Award.

# Advisory Board

# Contributors


**A. Roberto Ambrosini**
Institute of Radioastronomy
Bologna, Italy

**Jeff P. Anderson**
LTV Steel Corporation
Independence, Ohio

**Pasquale Arpaia**
Università di Napoli Federico II
Naples, Italy

**Francesco Avallone**
Università di Napoli Federico II
Naples, Italy

**Aldo Baccigalupi**
Università di Napoli Federico II
Naples, Italy

**William A. Barrow**
Planar Systems
Beaverton, Oregon

**Cipriano Bartoletti**
Institute of Radioastronomy
Bologna, Italy

**David M. Beams**
University of Texas at Tyler
Tyler, Texas

**K. Beilenhoff**
Institut für
Hochfrequenztechnik,
Technische Universität
Darmstadt, Germany

**Richard J. Blotzer**
Cleveland, Ohio

**C. Bortolotti**
Institute of Radioastronomy
Bologna, Italy

**Arnaldo Brandolini**
Politecnico di Milano
Milan, Italy

**Saps Buchman**
Stanford University
Stanford, California

**Barrett S. Caldwell**
Purdue University
West Lafayette, Indiana

**Robert B. Campbell**
Sandia National Laboratories
Albuquerque, New Mexico

**Robert M. Crovella**
Plano, Texas

**N. D'Amico**
Institute of Radioastronomy
Bologna, Italy

**Claudio de Capua**
Università di Napoli Federico II
Naples, Italy

**Alfons Dehé**
Siemens AG
Munich, Germany

**Achim Dreher**
German Aerospace Center
Wessling, Germany

**Halit Eren**
Curtin University of Technology
Bentley, Australia

**Alessandro Ferrero**
Politecnico di Milano
Milan, Italy

**K. Fricke**
Institut für
Hochfrequenztechnik,
Technische Universität
Darmstadt, Germany

**Alessandro Gandelli**
Politecnico di Milano
Milan, Italy

**Daryl Gerke**
Kimmel Gerke Associates, Ltd.
Mesa, Arizona

**W. A. Gillespie**
University of Abertay
Dundee, Scotland

**James Goh**
Curtin University of Technology
Perth, Australia

**G. Grueff**
Institute of Radioastronomy
Bologna, Italy

**H. L. Hartnagel**
Institut für
Hochfrequenztechnik,
Technische Universität
Darmstadt, Germany

**Michael B. Heaney**
Palo Alto Research Center
Palo Alto, California

**Albert D. Helfrick**
Embry-Riddle Aeronautical
University
Daytona Beach, Florida

**David A. Hill**
National Institute of Standards
and Technology
Boulder, Colorado

**Rahman Jamal**
National Instruments Germany
Munich, Germany

**Motohisa Kanda**
National Institute of Standards
and Technology
Boulder, Colorado

**Mohammad A. Karim**
City College of New York
New York, New York

**William Kimmel**
Kimmel Gerke Associates, Ltd.
Mesa, Arizona

**H. Klingbeil**
Institut für
Hochfrequenztechnik,
Technische Universität
Darmstadt, Germany

**V. Krozer**
Institut für
Hochfrequenztechnik,
Technische Universität
Darmstadt, Germany

**Carmine Landi**
Università de L'Aquila
L'Aquila, Italy

**W. Marshall Leach, Jr.**
Georgia Institute of Technology
Atlanta, Georgia

**Yufeng Li**
Samsung Information Systems
San Jose, California

**E. B. Loewenstein**
National Instruments
Austin, Texas

**Michael Z. Lowenstein**
Harmonics Limited
Mequou, Wisconsin

**Albert Lozano-Nieto**
The Pennsylvania State
University
Lehman, Pennsylvania

**Steven A. Macintyre**
Macintyre Electronic Design
Herndon, Virginia

**Allan M. MacLeod**
University of Abertay
Dundee, Scotland

**Sergio Mariotti**
Institute of Radioastronomy
Bologna, Italy

**P. F. Martin**
University of Abertay
Dundee, Scotland

**Edward McConnell**
National Instruments
Austin, Texas

**Robert T. McGrath**
The Pennsylvania State
University
University Park, Pennsylvania

**Douglas P. McNutt**
The MacNauchtan Laboratory
Colorado Springs, Colorado

**John Mester**
Stanford University
Stanford, California

**Jeffrey P. Mills**
Illinois Institute of Technology
Chicago, Illinois

**Devendra K. Misra**
University of Wisconsin
Milwaukee, Wisconsin

**William C. Moffatt**
Sandia National Laboratories
Albuquerque, New Mexico

**Stelio Montebugnoli**
Institute of Radioastronomy
Fontano, Italy

**Jerry Murphy**
Hewlett Packard Company
Geneva, Switzerland

**Steven A. Murray**
California State University
Fullerton, California

**A. Orfei**
Institute of Radioastronomy
Bologna, Italy

**Peter O'Shea**
Royal Melbourne Institute of Technology
Melbourne, Australia

**Ramón Pallás-Areny**
Universitat Politecnica de Catalyna
Barcelona, Spain

**Ronney B. Panerai**
University of Leicester
Leicester, England

**Luca Podestà**
University of Rome La Sapienza
Rome, Italy

**Rodney Pratt**
University of South Australia
Adelaide, Australia

**Gordon W. Roberts**
McGill University
Montreal, Canada

**Giancarlo Sacerdoti**
University of Rome La Sapienza
Rome, Italy

**Kalluri R. Sarma**
Honeywell, Inc.
Phoenix, Arizona

**Christopher J. Sherman**
Merrimack, New Hampshire

**Robert Steer**
Frequency Devices
Haverhill, Massachusetts

**Timothy J. Sumner**
Imperial College
London, England

**Peter H. Sydenham**
University of South Australia
Adelaide, Australia

**Michał Szyper**
University of Mining and Metallurgy
Cracow, Poland

**G. Tomassetti**
University of L'Aquila
L'Aquila, Italy

**Michael F. Toner**
Northern Telecom Ltd.
Nepean, Ontario, Canada

**Ramanapathy Veerasingam**
The Pennsylvania State University
University Park, Pennsylvania

**Herman Vermariën**
Vrije Universiteit Brussel
Brussels, Belgium

# Contents

## SECTION I Electromagnetic Variables Measurement

# I

# Electromagnetic Variables Measurement

# 1

# Voltage Measurement

Alessandro Ferrero
*Politecnico di Milano*

Jerry Murphy
*Hewlett Packard Company*

Cipriano Bartoletti
*Institute of Radioastronomy*

Luca Podestà
*University of Rome La Sapienza*

Giancarlo Sacerdoti
*University of Rome La Sapienza*



## 1.1 Meter Voltage Measurement

*Alessandro Ferrero*

Instruments for the measurement of electric voltage are called *voltmeters*. Correct insertion of a voltmeter requires the connection of its terminals to the points of an electric circuit across which the voltage has to be measured, as shown in Figure 1.1. To a first approximation, the electric equivalent circuit of a voltmeter can be represented by a resistive impedance $Z_v$ (or a pure resistance $R_v$ for dc voltmeters). This means that any voltmeter, once connected to an electric circuit, draws a current $I_v$ given by:

$$I_v = \frac{U}{Z_v} \tag{1.1}$$

where $U$ is the measured voltage. The higher the value of the internal impedance, the higher the quality of the voltmeter, since it does not significantly modify the status of the electric circuit under test.

Different operating principles are used to measure an electric voltage. The mechanical interaction between currents, between a current and a magnetic field, or between electrified conductors was widely adopted in the past to generate a mechanical torque proportional to the voltage or the squared voltage to be measured. This torque, balanced by a restraining torque, usually generated by a spring, causes the instrument pointer, which can be a mechanical or a virtual optical pointer, to be displaced by an angle proportional to the driving torque, and hence to the voltage or the squared voltage to be measured. The value of the input voltage is therefore given by the reading of the pointer displacement on a graduated scale. The thermal effects of a current flowing in a conductor are also used for measuring electric voltages, although they have not been adopted as widely as the previous ones. More recently, the widespread diffusion of semiconductor devices led to the development of a completely different class of voltmeters: *electronic* voltmeters. They basically attain the required measurement by processing the input signal by means of electronic semiconductor devices. According to the method, analog or digital, the input signal is processed, the electronic voltmeters can be divided into *analog* electronic voltmeters and *digital*

0-8493-1733-9/04/$0.00+$1.50

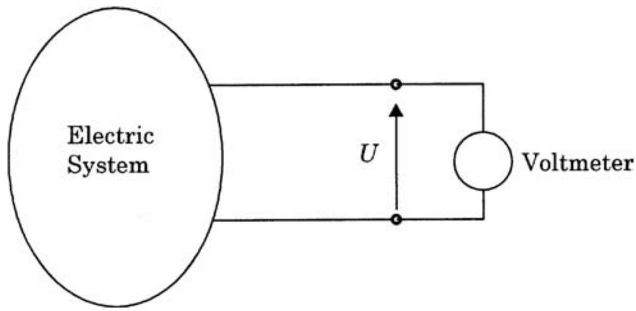



**FIGURE 1.1** Voltmeter insertion.

**TABLE 1.1** Classification of Voltage Meters

| Class | Operating principle | Subclass | Application field |
|---|---|---|---|
| Electromagnetic | Interaction between currents and magnetic fields | Moving magnet | Dc voltage |
| | | Moving coil | Dc voltage |
| | | Moving iron | Dc and ac voltage |
| Electrodynamic | Interactions between currents | — | Dc and ac voltage |
| Electrostatic | Electrostatic interactions | — | Dc and ac voltage |
| Thermal | Current's thermal effects | Direct action | Dc and ac voltage |
| | Indirect action | | Dc and ac voltage |
| Induction | Magnetic induction | — | Ac voltage |
| Electronic | Signal processing | Analog | Dc and ac voltage |
| | | Digital | Dc and ac voltage |

electronic voltmeters. Table 1.1 shows a rough classification of the most commonly employed voltmeters, according to their operating principle and their typical application field.

This chapter section briefly describes the most commonly employed voltmeters, both electromechanical and electronic.

## Electromechanical Voltmeters

Electromechanical voltmeters measure the applied voltage by transducing it into a mechanical torque. This can be accomplished in different ways, basically because of the interactions between currents (*electrodynamic voltmeters*), between a current and a magnetic field (*electromagnetic voltmeters*), between electrified conductors (*electrostatic voltmeters*, or *electrometers*), and between currents induced in a conducting vane (*induction voltmeters*). According to the different kinds of interactions, different families of instruments can be described, with different application fields. Moving-coil electromagnetic voltmeters are restricted to the measurement of dc voltages; moving-iron electromagnetic, electrodynamic, and electrostatic voltmeters can be used to measure both dc and ac voltages; while induction voltmeters are restricted to ac voltages.

The most commonly employed electromechanical voltmeters are the electromagnetic and electrodynamic ones. Electrostatic voltmeters have been widely employed in the past (and are still employed) for the measurement of high voltages, both dc and ac, up to a frequency on the order of several megahertz. Induction voltmeters have never been widely employed, and their present use is restricted to ac voltages.

Therefore, only the electromagnetic, electrodynamic, and electrostatic voltmeters will be described in the following.

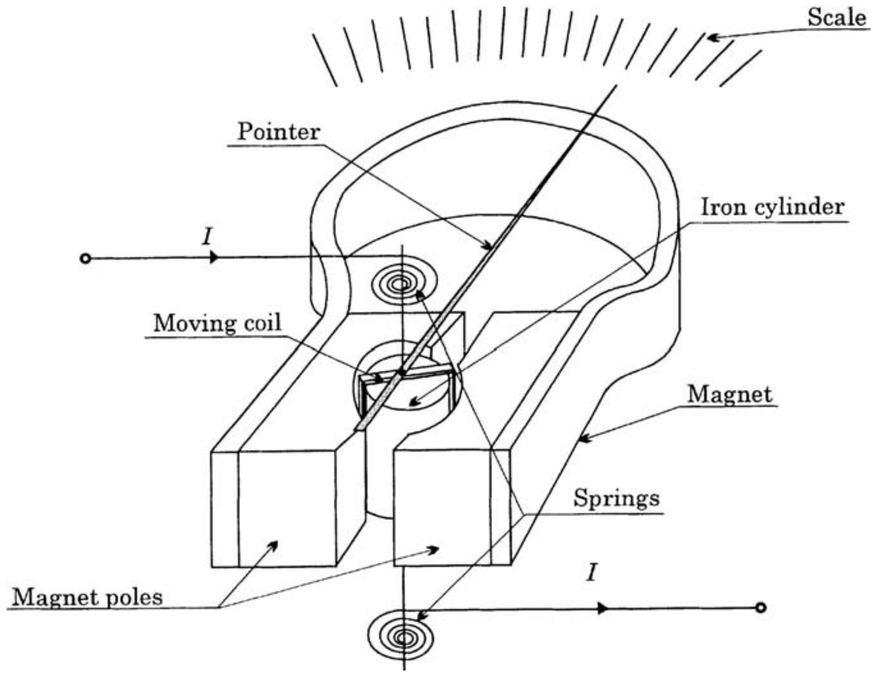



**FIGURE 1.2** Dc moving-coil meter.

## Electromagnetic Voltmeters

### Dc Moving-Coil Voltmeters.

The structure of a dc moving-coil meter is shown in Figure 1.2. A small rectangular pivoted coil is wrapped around an iron cylinder and placed between the poles of a permanent magnet. Because of the shape of the poles of the permanent magnet, the induction magnetic field $B$ in the air gap is radial and constant.

Suppose that a dc current $I$ is flowing in the coil, the coil has $N$ turns, and that the length of the sides that cut the magnetic flux (active sides) is $l$; the current interacts with the magnetic field $B$ and a force $F$ is exerted on the conductors of the active sides. The value of this force is given by:

$$F = NBlI \tag{1.2}$$

Its direction is given by the right-hand rule. Since the two forces applied to the two active sides of the coil are directed in opposite directions, a torque arises in the coil, given by:

$$T_\mathrm{i} = Fd = NBldI \tag{1.3}$$

where $d$ is the coil width. Since $N$, $B$, $l$, $d$ are constant, Equation 1.3 leads to:

$$T_\mathrm{i} = k_\mathrm{i} I \tag{1.4}$$

showing that the mechanical torque exerted on the coil is directly proportional to the current flowing in the coil itself.

Because of $T_\mathrm{i}$, the coil rotates around its axis. Two little control springs, with $k_\mathrm{r}$ constant, provide a restraining torque $T_\mathrm{r}$. The two torques balance when the coil is rotated by an angle $\delta$ so that:

$$k_\mathrm{i} I = k_\mathrm{r} \delta \tag{1.5}$$

which leads to:

$$\delta = \frac{k_\mathrm{i}}{k_\mathrm{r}} I \tag{1.6}$$

Equation 1.6 shows that the rotation angle of the coil is directly proportional to the dc current flowing in the coil. If a pointer with length $h$ is keyed on the coil axes, a displacement $l = h\delta$ can be read on the instrument scale. Therefore, the pointer displacement is proportional to the current flowing in the coil, according to the following relationship:

$$\lambda = h \frac{k_\mathrm{i}}{k_\mathrm{r}} I \tag{1.7}$$

This instrument is hence intrinsically a current meter. A voltmeter can be obtained by connecting an additional resistor in series with the coil. If the coil resistance is $R_\mathrm{c}$, and the resistance of the additional resistor is $R_\mathrm{a}$, the current flowing in the coil when the voltage $U$ is applied is given by:

$$I = \frac{U}{R_\mathrm{a} + R_\mathrm{c}} \tag{1.8}$$

and therefore the pointer displacement is given by:

$$\lambda = h\delta = h \frac{k_\mathrm{i}}{k_\mathrm{r}} I = h \frac{k_\mathrm{i}}{k_\mathrm{r}\left(R_\mathrm{a} + R_\mathrm{c}\right)} U \tag{1.9}$$

and is proportional to the applied voltage. Because of this proportionality, moving-coil dc meters show a proportional-law scale, where the applied voltage causes a proportional angular deflection of the pointer.

Because of the operating principle expressed by Equation 1.3, these voltmeters can measure only dc voltages. Due to the inertia of the mechanical part, ac components typically do not cause any coil rotation, and hence these meters can be also employed to measure the dc component of a variable voltage. They have been widely employed in the past for the measurement of dc voltages up to some thousands volts with a relative measurement uncertainty as low as 0.1% of the full-scale value. At present, they are being replaced by electronic voltmeters that feature the same or better accuracy at a lower cost.

***Dc Galvanometer.***

*General characteristics.* A galvanometer is used to measure low currents and low voltages. Because of the high sensitivity that this kind of measurement requires, galvanometers are widely employed as null indicators in all dc balance measurement methods (like the bridge and potentiometer methods) [1, 2].

A dc galvanometer is, basically, a dc moving-coil meter, and the relationship between the index displacement and the current flowing in the moving coil is given by Equation 1.7. The instrument constant:

$$k_\mathrm{a} = h \frac{k_\mathrm{i}}{k_\mathrm{r}} \tag{1.10}$$

is usually called the galvanometer *current constant* and is expressed in mm $\mu A^{-1}$. The galvanometer *current sensitivity* is defined as $1/k_\mathrm{a}$ and is expressed in $\mu A\ mm^{-1}$.

According to their particular application field, galvanometers must be chosen with particular care. If $k_\mathrm{a}$ is taken into account, note that once the full-scale current and the corresponding maximum pointer displacement are given, the value of the ratio $hk_\mathrm{i}/k_\mathrm{r}$ is also known. However, the single values of $h$, $k_\mathrm{i}$,

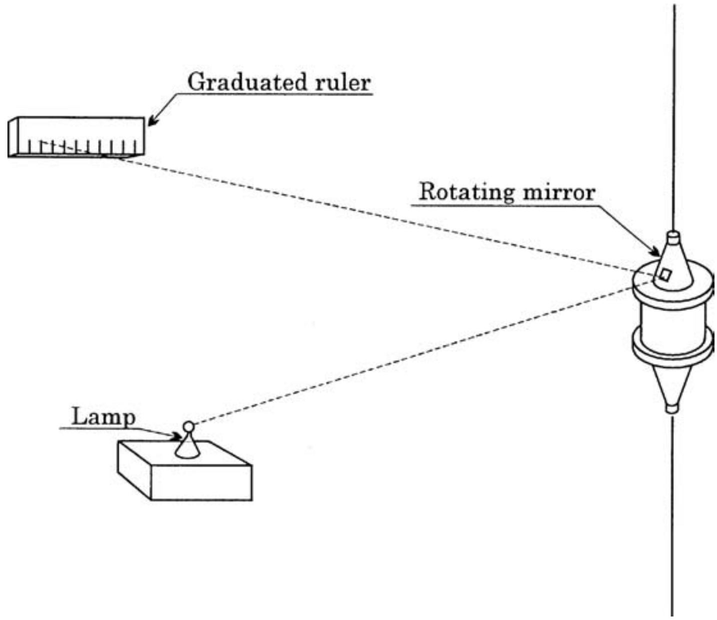



**FIGURE 1.3** Virtual optical pointer structure in a dc galvanometer.

and $k_r$ can assume any value and are usually set in order to reduce the friction effects. In fact, if the restraining friction torque $T_f$ is taken into account in the balance equation, Equation 1.5 becomes:

$$k_i I = k_r \frac{\lambda}{h} \pm T_f \tag{1.11}$$

where the ± sign shows that the friction torque does not have its own sign, but always opposes the rotation.

The effects of $T_f$ can be neglected if the driving torque $hk_i I$ and the restraining torque $k_r \lambda$ are sufficiently greater than $T_f$. Moreover, since the galvanometer is employed as a null indicator, a high sensitivity is needed; hence, $k_a$ must be as high as possible. According to Equations 1.10 and 1.11, this requires high values of $hk_i$ and low values of $k_r$. A high value of $h$ means a long pointer; a high value of $k_i$ means a high driving torque, while a low value of $k_r$ means that the inertia of the whole moving system must be low.

The pointer length can be increased without increasing the moving system inertia by employing virtual optical pointers: a little, light concave mirror is fixed on the moving-coil axis and is lit by an external lamp. The reflected light hits a translucid, graduated ruler, so that the mirror rotation can be observed (Figure 1.3). In this way, a virtual pointer is obtained, whose length equals the distance between the mirror and the graduated ruler.

The reduction of the moving system inertia is obtained by reducing the weight and dimension of the moving coil, and reducing the spring constant. This is usually done by suspending the moving coil with a thin fiber of conducting material (usually bronze). Thus, the friction torque is practically removed, and the restraining spring action is given by the fiber torsion.

According to Equations 1.3 and 1.4, the driving torque can be increased by increasing the coil flux linkage. Three parameters can be modified to attain this increase: the induction field $B$, the coil section $ld$, and the number of turns $N$ of the coil winding.

The induction field $B$ can be increased by employing high-quality permanent magnets, with high coercive force, and minimizing the air gap between the magnet's poles. This minimization prevents the use of moving coils with a large section. Moreover, large coil sections lead to heavier coils with greater inertia, which opposes the previous requirement of reduced inertia. For this reason, the coil section is usually rectangular (although a square section maximizes the flux linkage) and with $l > d$.

If the galvanometer is used to measure a low voltage $U$, the *voltage sensitivity*, expressed in $\mu$V mm$^{-1}$ is the inverse of:

$$k_v = \frac{\lambda}{U} \tag{1.12}$$

where $k_v$ is called the galvanometer's *voltage constant* and is expressed in mm $\mu$V$^{-1}$.

*Mechanical characteristics*. Due to the low inertia and low friction, the galvanometer moving system behaves as an oscillating mechanical system. The oscillations around the balance position are damped by the electromagnetic forces that the oscillations of the coil in the magnetic field exert on the coil active sides. It can be proved [1] that the oscillation damping is a function of the coil circuit resistance: that is, the coil resistance $r$ plus the equivalent resistance of the external circuit connected to the galvanometer.

In particular, the damping effect is nil if the coil circuit is open, and maximum if the coil is short-circuited. In practical situations, a resistor is connected in series with the moving coil, whose resistance is selected in such a way to realize a critical damping of the coil movement. When this situation is obtained, the galvanometer is said to be *critically damped* and reaches its balance position in the shortest time, without oscillations around this position.

*Actual trends*. Moving-coil dc galvanometers have been widely employed in the past when they represented the most important instrument for high-sensitivity measurements. In more recent years, due to the development of the electronic devices, and particularly high-gain, low-noise amplifiers, the moving-coil galvanometers are being replaced by electronic galvanometers, which feature the same, or even better, performance than the electromagnetic ones.

## Electrodynamic Voltmeters

### Ac Moving-Coil Voltmeters.

The structure of an ac moving-coil meter is shown in Figure 1.4. It basically consists of a pivoted moving coil, two stationary field coils, control springs, a pointer, and a calibrated scale. The stationary coils are series connected and, when a current $i_f$ is applied, a magnetic field $B_f$ is generated along the axis of the stationary coils, as shown in Figure 1.5. A magnetic flux is therefore generated, whose instantaneous values are given by:

$$\varphi_f(t) = k' m_f i_f(t) \tag{1.13}$$

where $m_f$ is the number of turns of the stationary coil and $k'$ is a proportionality factor. When a current $i_m$ is applied to the moving coil, a torque arises, whose instantaneous values are proportional to the product of $\varphi_f$ and $i_m$ instantaneous values:

$$T_i(t) = k'' \varphi_f(t) i_m(t) = k i_f(t) i_m(t) \tag{1.14}$$

The driving torque is therefore proportional to the instantaneous product of the currents flowing in the two coils. Due to this driving torque, the moving element is displaced by an angle ($\delta t$), until the spring restraining torque $T_s(t) = k_s \delta(t)$ balances the driving torque. The moving element rotation is thus given by:

$$\delta(t) = \frac{k}{k_s} i_f(t) i_m(t) \tag{1.15}$$

and, if the pointer length is $h$, the following pointer displacement can be read on the scale:

$$\lambda(t) = h \frac{k}{k_s} i_f(t) i_m(t) \tag{1.16}$$

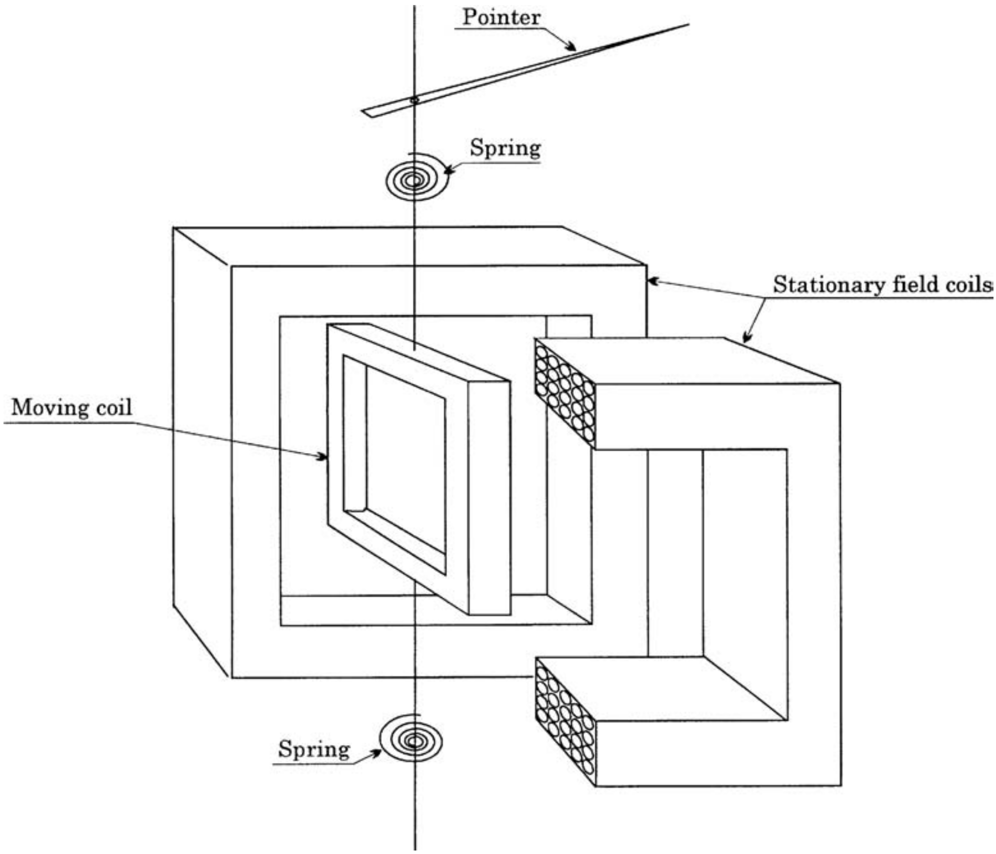



**FIGURE 1.4** Ac moving-coil meter.

The proportionality factor $k$ is generally not constant, since it depends on the mutual inductance between the two coils, and thus on their number of turns, shape, and relative position. However, if the two coils are carefully designed and placed, the magnetic field can be assumed to be constant and radial in the rotation area of the moving coil. Under this condition, $k$ is virtually constant.

Because the bandwidth of the moving element is limited to a few hertz, due to its inertia, the balance position is proportional to the average value of the driving torque when the signal bandwidth exceeds this limit. If $i_f$ and $i_m$ currents are sinusoidal, with $I_f$ and $I_m$ rms values, respectively, and with a relative phase displacement $\beta$, the driving torque average value is given by:

$$\overline{T_i} = kI_f I_m \cos\beta \tag{1.17}$$

and thus, the pointer displacement in Equation 1.16 becomes:

$$\lambda = h\frac{k}{k_S} I_f I_m \cos\beta \tag{1.18}$$

In order to realize a voltmeter, the stationary and moving coils are series connected, and a resistor, with resistance $R$, is also connected in series to the coils. If $R$ is far greater than the resistance of the two coils, and if it is also far greater than the coil inductance, in the frequency operating range of the voltmeter, the rms value of the coils' currents is given by:

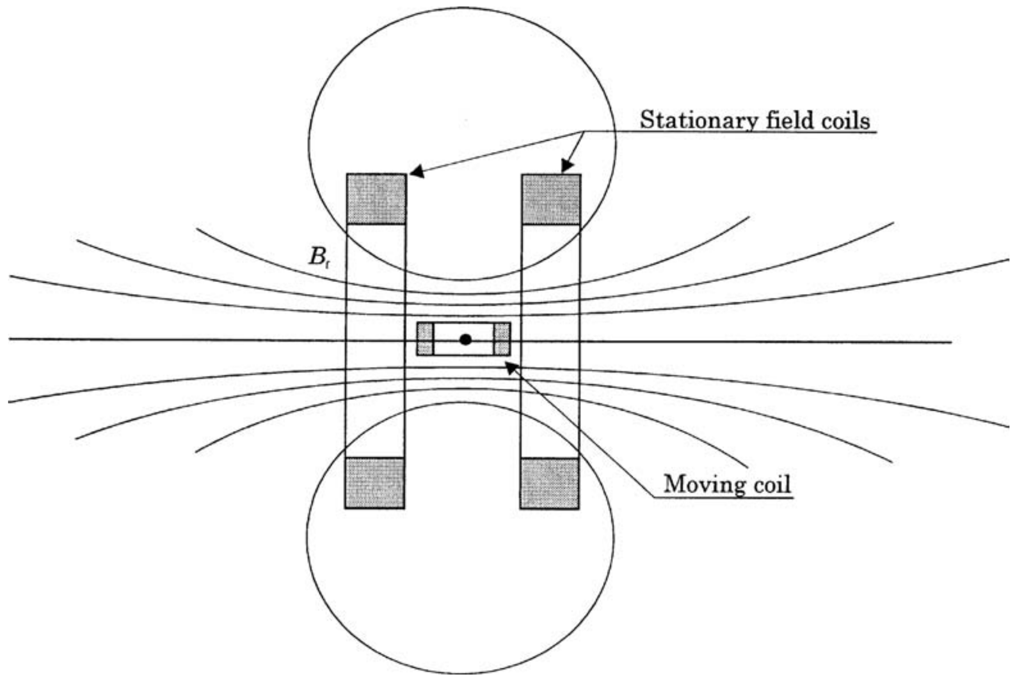



**FIGURE 1.5** Magnetic field generated by the field coils in an ac moving-coil meter.

$$I_f = I_m = \frac{U}{R} \tag{1.19}$$

$U$ being the applied voltage rms value. From Equation 1.18, the pointer displacement is therefore given by:

$$\lambda = h \frac{k}{k_S} \frac{U^2}{R^2} = k_v U^2 \tag{1.20}$$

Because of Equation 1.20, the voltmeter features a square-law scale, with $k_v$ constant, provided that the coils are carefully designed, and that the coils' inductance can be neglected with respect to the resistance of the coils themselves and the series resistor. This last condition determines the upper limit of the input voltage frequency.

These voltmeters feature good accuracy (their uncertainty can be as low as 0.2% of the full-scale value), with full-scale values up to a few hundred volts, in a frequency range up to 2 kHz.

## Electrostatic Voltmeters

The action of electrostatic instruments is based on the force exerted between two charged conductors. The conductors behave as a variable plate air capacitor, as shown in Figure 1.6. The moving plate, when charged, tends to move so as to increase the capacitance between the plates. The energy stored in the capacitor, when the applied voltage is $U$ and the capacitance is $C$, is given by:

$$W = \frac{1}{2} C U^2 \tag{1.21}$$

This relationship is valid both under dc and ac conditions, provided that the voltage rms value $U$ is considered for ac voltage.

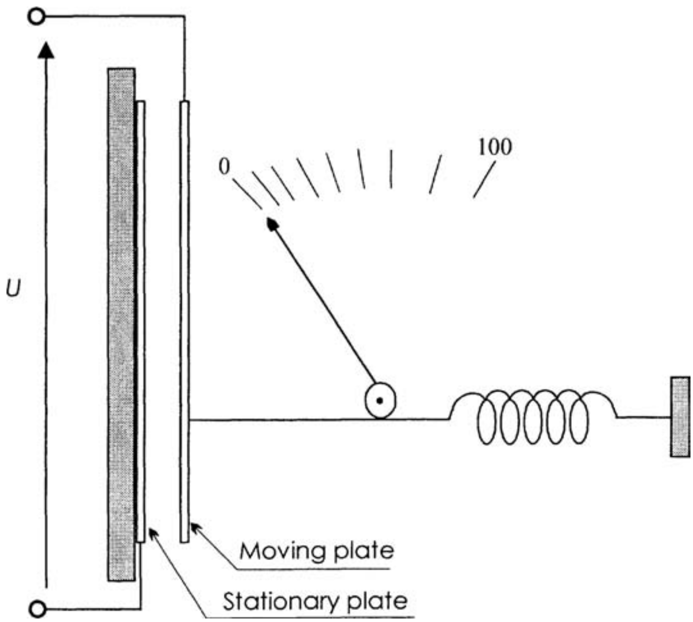



**FIGURE 1.6** Basic structure of an electrostatic voltmeter.

When the moving plate is displaced horizontally by d$s$, while the voltage is held constant, the capacitor energy changes in order to equal the work done in moving the plate. The resulting force is:

$$F = \frac{\mathrm{d}W}{\mathrm{d}s} = \frac{U^2}{2}\frac{\mathrm{d}C}{\mathrm{d}s} \tag{1.22}$$

For a rotable system, Equation 1.21 leads similarly to a resulting torque:

$$T = \frac{\mathrm{d}W}{\mathrm{d}\vartheta} = \frac{U^2}{2}\frac{\mathrm{d}C}{\mathrm{d}\vartheta} \tag{1.23}$$

If the action of a control spring is also considered, both Equations 1.22 and 1.23 show that the balance position of the moving plate is proportional to the square of the applied voltage, and hence electrostatic voltmeters have a square-law scale. These equations, along with Equation 1.21, show that these instruments can be used for the measurement of both dc and ac rms voltages. However, the force (or torque) supplied by the instrument schematically represented in Figure 1.6 is generally very weak [2], so that its use is very impractical.

### The Electrometer.

A more useful configuration is the quadrant electrometer, shown in Figure 1.7. Four fixed plates realize four quadrants and surround a movable vane suspended by a torsion fiber at the center of the system. The opposite quadrants are electrically connected together, and the potential difference ($U_1 - U_2$) is applied. The moving vane can be either connected to potential $U_1$ or $U_2$, or energized by an independent potential $U_3$.

Let the zero torque position of the suspension coincide with the symmetrical X-X position of the vane. If $U_1 = U_2$, the vane does not leave this position; otherwise, the vane will rotate.

Let $C_1$ and $C_2$ be the capacitances of quadrants 1 and 2, respectively, relative to the vane. They both are functions of $\vartheta$ and, according to Equation 1.23, the torque applied to the vane is given by:

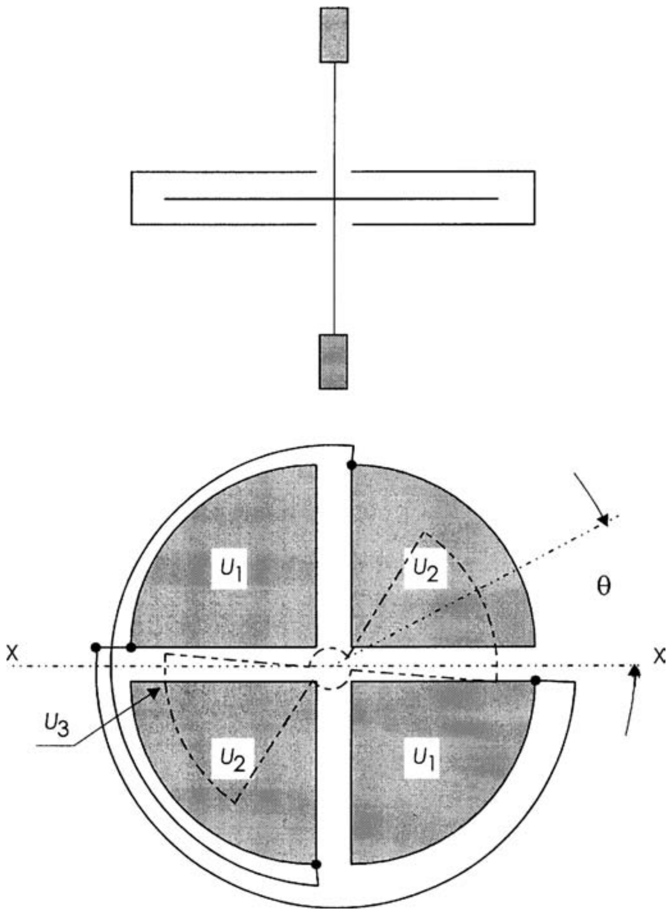



**FIGURE 1.7** Quadrant electrometer structure.

$$T = \frac{(U_3 - U_1)^2}{2}\frac{dC_1}{d\vartheta} + \frac{(U_3 - U_2)^2}{2}\frac{dC_2}{d\vartheta} \tag{1.24}$$

Since the vane turns out of one pair of quadrants as much as it turns into the other, the variations of $C_1$ and $C_2$ can be related by:

$$-\frac{dC_1}{d\vartheta} = \frac{dC_2}{d\vartheta} = k_1 \tag{1.25}$$

Taking into account the suspension restraining torque $T_r = k_2\vartheta$, the balance position can be obtained by Equations 1.24 and 1.25 as:

$$\vartheta = \frac{k_1}{2k_2}\left[(U_3 - U_2)^2 - (U_3 - U_1)^2\right] \tag{1.26}$$

If the vane potential $U_3$ is held constant, and is large compared to the quadrant potentials $U_1$ and $U_2$, Equation 1.26 can be simplified as follows:

$$\vartheta = \frac{k_1}{k_2}U_3(U_1 - U_2) \tag{1.27}$$

Equation 1.27 shows that the deflection of the vane is directly proportional to the voltage difference applied to the quadrants. This method of use is called the *heterostatic* method.

If the vane is connected to quadrant 1, $U_3 = U_1$ follows, and Equation 1.26 becomes

$$\vartheta = \frac{k_1}{2k_2}\left(U_1 - U_2\right)^2 \tag{1.28}$$

Equation 1.28 shows that the deflection of the vane is proportional to the square of the voltage difference applied to the quadrants, and hence this voltmeter has a square-law scale. This method of use is called the *idiostatic* method, and is suitable for the direct measurement of dc and ac voltages without an auxiliary power source.

The driving torque of the electrometer is extremely weak, as in all electrostatic instruments. The major advantage of using this kind of meter is that it allows for the measurement of dc voltages without drawing current by the voltage source under test. Now, due to the availability of operational amplifiers with extremely high input impedance, they have been almost completely replaced by electronic meters with high input impedance.

## Electronic Voltmeters

Electronic meters process the input signal by means of semiconductor devices in order to extract the information related to the required measurement [3, 4]. An electronic meter can be basically represented as a three-port element, as shown in Figure 1.8.

The input signal port is an input port characterized by high impedance, so that the signal source has very little load. The measurement result port is an output port that provides the measurement result (in either an analog or digital form, depending on the way the input signal is processed) along with the power needed to energize the device used to display the measurement result. The power supply port is an input port which the electric power required to energize the meter internal devices and the display device flows through.

One of the main characteristics of an electronic meter is that it requires an external power supply. Although this may appear as a drawback of electronic meters, especially where portable meters are concerned, note that, this way, the energy required for the measurement is no longer drawn from the signal source.

The high-level performance of modern electronic devices yields meters that are as accurate (and sometimes even more accurate) as the most accurate electromechanical meters. Because they do not require the extensive use of precision mechanics, they are presently less expensive than electromechanical meters, and are slowly, but constantly, replacing them in almost all applications.

Depending on the way the input signal is processed, electronic meters are divided into *analog* and *digital* meters. Analog meters attain the required measurement by analog, continuous-time processing of the input signal. The measurement result can be displayed both in analog form using, for example, an electromechanical meter; or in digital form by converting the analog output signal into digital form.

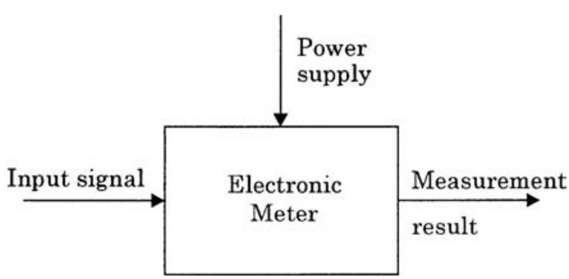



**FIGURE 1.8** Electronic meter.

Digital meters attain the required measurement by digital processing of the input signal. The measurement result is usually displayed in digital form. Note that the distinction between analog and digital meters is not due to the way the measurement result is displayed, but to the way the input signal is processed.

## Analog Voltmeters

An electronic analog voltmeter is based on an electronic amplifier and an electromechanical meter to measure the amplifier output signal. The amplifier operates to make a dc current, proportional to the input quantity to be measured, flow into the meter. This meter is hence a dc moving-coil milliammeter.

Different full-scale values can be obtained using a selectable-ratio voltage divider if the input voltage is higher than the amplifier dynamic range, or by selecting the proper amplifier gain if the input voltage stays within the amplifier dynamic range.

The main features of analog voltmeters are high input impedance, high possible gain, and wide possible bandwidth for ac measurements. The relative measurement uncertainty can be lower than 1% of full-scale value. Because of these features, electronic analog voltmeters can have better performance than the electromechanical ones.

### Dc Analog Voltmeters

Figure 1.9 shows the circuit for an electronic dc analog voltmeter. Assuming that the operational amplifier exhibits ideal behavior, current $I_m$ flowing in the milliammeter A is given by:

$$I_m = I_o + I_2 = \frac{U_o}{R_o} + \frac{U_o}{R_2} = -U_i \frac{R_2}{R_1} \frac{R_2 + R_o}{R_2 R_o} = -\frac{U_1}{R_1}\left(1 + \frac{R_2}{R_o}\right) \tag{1.29}$$

If $R_1 = R_2$, and the same resistances are far greater than $R_o$, Equation 1.29 can be simplified to:

$$I_m = -\frac{U_1}{R_o} \tag{1.30}$$

Equation 1.30 shows that the milliammeter reading is directly proportional to the input voltage through resistance $R_o$ only. This means that, once the milliammeter full-scale value is set, the voltmeter full-scale value can be changed, within the dynamic range of the amplifier, by changing the $R_o$ value. This way, the meter full-scale value can be changed without changing its input impedance.

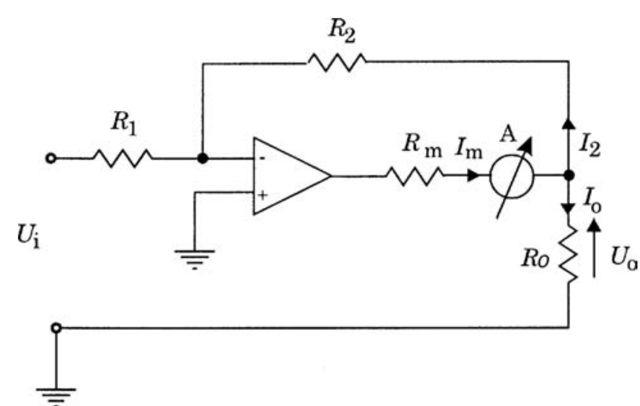



**FIGURE 1.9** Electronic dc analog voltmeter schematics.

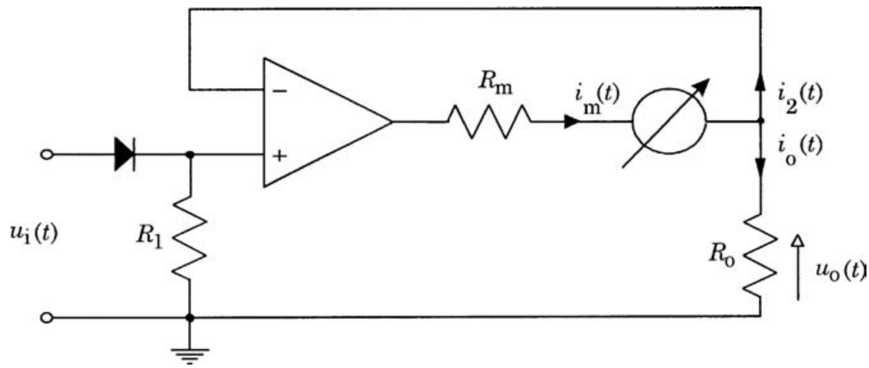



**FIGURE 1.10** Electronic, rectifier-based ac analog voltmeter schematics.

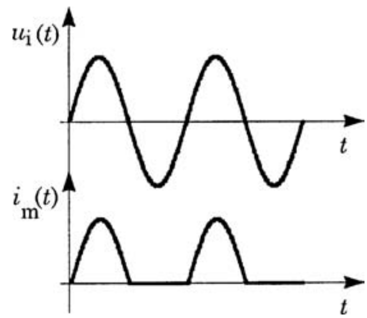



**FIGURE 1.11** Signal waveforms in a rectifier-based ac analog voltmeter when the input voltage is sinusoidal.

### Rectifier-Based Ac Analog Voltmeters.

Analog meters for ac voltages can be obtained starting from the dc analog voltmeters, with a rectifying input stage. Figure 1.10 shows how the structure in Figure 1.9 can be modified in order to realize an ac voltmeter.

Because of the high input impedance of the electronic amplifier, $i_2(t) = 0$, and the current $i_m(t)$ flowing in the milliammeter A is the same as current $i_o(t)$ flowing in the load resistance. Since the amplifier is connected in a voltage-follower configuration, the output voltage is given by:

$$u_o(t) = u_i(t) \tag{1.31}$$

Due to the presence of the input diode, current $i_m(t)$ is given by:

$$i_m(t) = \frac{u_i(t)}{R_o} \tag{1.32}$$

when $u_i(t) > 0$, and

$$i_m(t) = 0 \tag{1.33}$$

when $u_i(t) \le 0$. If $u_i(t)$ is supposed to be a sine wave, the waveform of $i_m(t)$ is shown in Figure 1.11.

The dc moving-coil milliammeter measures the average value $\bar{I}_m$ of $i_m(t)$, which, under the assumption of sinusoidal signals, is related to the rms value $U_i$ of $u_i(t)$ by:

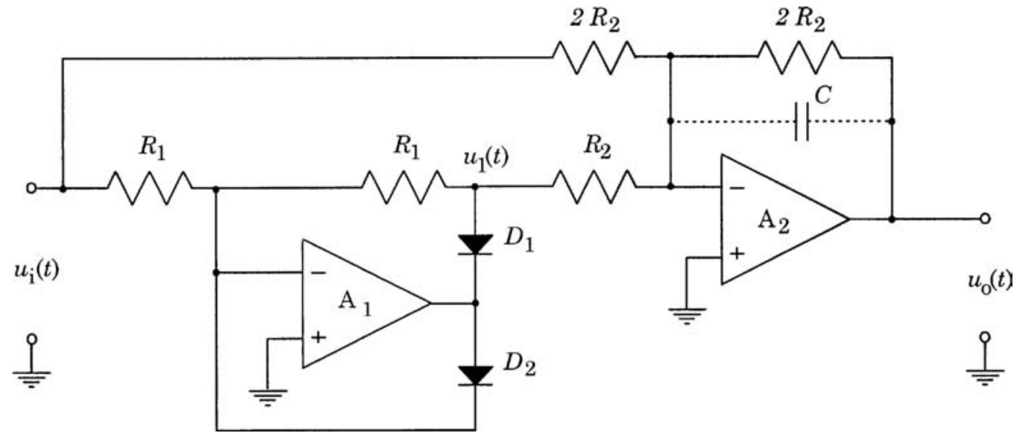



**FIGURE 1.12** Electronic, full-wave rectifier-based ac analog voltmeter schematics.

$$\bar{I}_m = \frac{2\sqrt{2}}{\pi R_o} U_i \tag{1.34}$$

The performance of the structure in Figure 1.10 can be substantially improved by considering the structure in Figure 1.12 which realizes a full-wave rectifier. Because of the presence of diodes $D_1$ and $D_2$, the output of amplifier A1 is given by:

$$u_1(t) = \begin{cases} -u_i(t) & \text{for } u_i(t) \geq 0 \\ 0 & \text{for } u_i(t) < 0 \end{cases} \tag{1.35}$$

where $u_i(t)$ is the circuit input voltage.

If capacitor $C$ is supposed to be not connected, amplifier $A_2$ output voltage is:

$$u_o(t) = -\left[u_i(t) + 2u_1(t)\right] \tag{1.36}$$

which gives:

$$u_o(t) = \begin{cases} u_i(t) & \text{for } u_i(t) \geq 0 \\ -u_i(t) & \text{for } u_i(t) < 0 \end{cases} \tag{1.37}$$

thus proving that the circuit in Figure 1.12 realizes a full-wave rectifier.

If $u_i(t)$ is a sine wave, the waveforms of $u_i(t)$, $u_1(t)$, and $u_o(t)$ are shown in Figure 1.13.

Connecting capacitor $C$ in the feedback loop of amplifier A2 turns it into a first-order low-pass filter, so that the circuit output voltage equals the average value of $u_o(t)$:

$$\bar{U}_o = \overline{\left|u_i(t)\right|} \tag{1.38}$$

In the case of sinusoidal input voltage with rms value $U_i$, the output voltage is related to this rms value by:

$$\bar{U}_o = \frac{2\sqrt{2}}{\pi} U_i \tag{1.39}$$

$\bar{U}_o$ can be measured by a dc voltmeter.

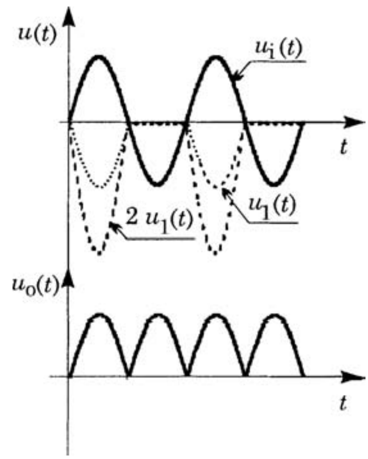



**FIGURE 1.13** Signal waveforms in a fullwave rectifier-based ac analog voltmeter when the input voltage is sinusoidal.

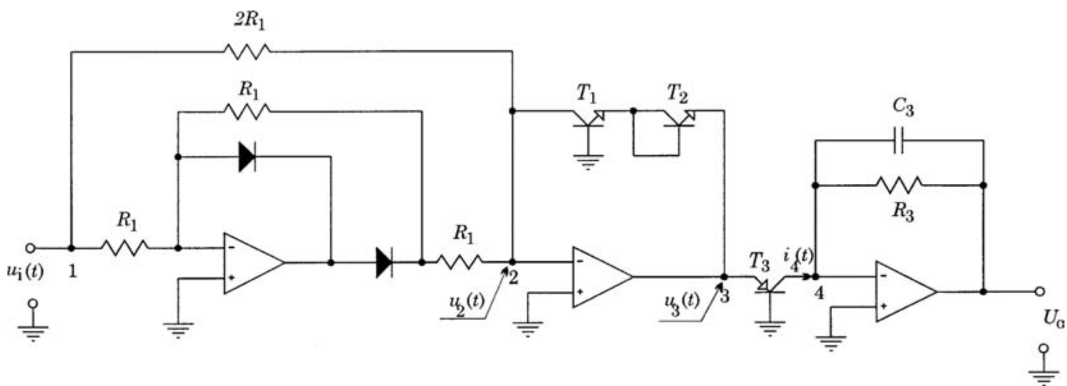



**FIGURE 1.14** True rms electronic ac voltmeter schematics.

Both meters in Figures 1.10 and 1.12 are actually average detectors. However, due to Equations 1.34 and 1.39, their scale can be labeled in such a way that the instrument reading gives the rms value of the input voltage, provided it is sinusoidal. When the input voltage is no longer sinusoidal, an error arises that depends on the signal form factor.

### True rms Analog Voltmeters.

The rms value $U_i$ of a periodic input voltage signal $u_i(t)$, with period $T$, is given by:

$$U_i = \sqrt{\frac{1}{T}\int_0^T u_i^2(t)dt} \tag{1.40}$$

The electronic circuit shown in Figure 1.14 provides an output signal $U_o$ proportional to the squared rms value of the input signal $u_i(t)$. The circuit section between nodes 1 and 2 is a full-wave rectifier. Hence, node 2 potential is given by:

$$u_2(t) = |u_i(t)| \tag{1.41}$$

The circuit section between nodes 2 and 4 is a log multiplier. Because of the logarithmic characteristic of the feedback path due to the presence of T1 and T2, node 3 potential is given by:

$$u_3(t) = 2k_1 \log[u_2(t)] = k_1 \log[u_2^2(t)] = k_1 \log\left[|u_i(t)|^2\right] = k_1 \log[u_i^2(t)] \tag{1.42}$$

and, due to the presence of $T_3$, the current flowing in node 4 is given by:

$$i_4(t) = k_2 \exp[u_3(t)] = k_3 u_i^2(t) \tag{1.43}$$

The circuit section after node 4 is a low-pass filter that extracts the dc component of the input signal. Therefore, the circuit output voltage is given by:

$$U_o = \frac{k}{T}\int_0^T u_i^2(t)dt = kU_i^2 \tag{1.44}$$

thus providing an output signal proportional to the squared rms value of the input signal $u_i(t)$ in accordance with Equation 1.40. Quantities $k_1$, $k_2$, and $k$ depend on the values of the elements in the circuit in Figure 1.14. Under circuit operating conditions, their values can be considered constant, so that $k_1$, $k_2$, and $k$ can be considered constant also.

If carefully designed, this circuit can feature an uncertainty in the range of ±1% of full scale, for signal frequencies up to 100 kHz.

## Digital Voltmeters

A digital voltmeter (DVM) attains the required measurement by converting the analog input signal into digital, and, when necessary, by discrete-time processing of the converted values. The measurement result is presented in a digital form that can take the form of a digital front-panel display, or a digital output signal. The digital output signal can be coded as a decimal BCD code, or a binary code.

The main factors that characterize DVMs are speed, automatic operation, and programmability. In particular, they presently offer the best combination of speed and accuracy if compared with other available voltage-measuring instruments. Moreover, the capability of automatic operations and programmability make DVMs very useful in applications where flexibility, high speed, and computer controllability are required. A typical application field is therefore that of automatically operated systems.

When a DVM is directly interfaced to a digital signal processing (DSP) system and used to convert the analog input voltage into a sequence of sampled values, it is usually called an analog-to-digital converter (ADC).

DVMs basically differ in the following ways: (1) number of measurement ranges, (2) number of digits, (3) accuracy, (4) speed of reading, and (5) operating principle.

The basic measurement ranges of most DVMs are either 1 V or 10 V. It is however possible, with an appropriate preamplifier stage, to obtain full-scale values as low as 0.1 V. If an appropriate voltage divider is used, it is also possible to obtain full-scale values as high as 1000 V.

If the digital presentation takes the form of a digital front-panel display, the measurement result is presented as a decimal number, with a number of digits that typically ranges from 3 to 6. If the digital representation takes the form of a binary-coded output signal, the number of bits of this representation typically ranges from 8 to 16, though 18-bit ADCs are available.

The accuracy of a DVM is usually correlated to its resolution. Indeed, assigning an uncertainty lower than the 0.1% of the range to a three-digit DVM makes no sense, since this is the displayed resolution of the instrument. Similarly, a poorer accuracy makes the three-digit resolution quite useless. Presently, a six-digit DVM can feature an uncertainty range, for short periods of time in controlled environments, as low as the 0.0015% of reading or 0.0002% of full range.

The speed of a DVM can be as high as 1000 readings per second. When the ADC is considered, the conversion rate is taken into account instead of the speed of reading. Presently, the conversion rate for

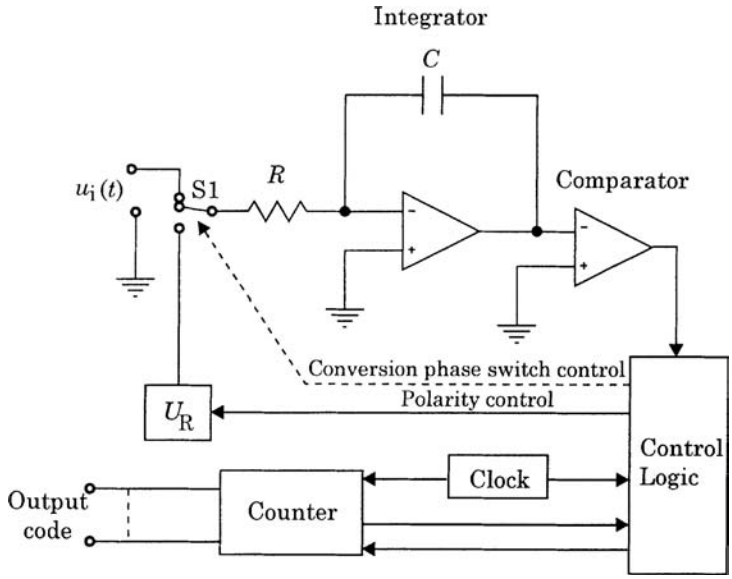



**FIGURE 1.15** Dual slope DVM schematics.

12-bit, successive approximation ADCs can be on the order of 10 MHz. It can be on the order of 100 MHz for lower resolution, flash ADCs [5].

DVMs can be divided into two main operating principle classes: the *integrating* types and the *non-integrating* types [3]. The following sections give an example for both types.

### Dual Slope DVM.

Dual slope DVMs use a counter and an integrator to convert an unknown analog input voltage into a ratio of time periods multiplied by a reference voltage. The block diagram in Figure 1.15 shows this operating principle. The switch S1 connects the input signal to the integrator for a fixed period of time $t_f$. If the input voltage is positive and constant, $u_i(t) = U_i > 0$, the integrator output represents a negative-slope ramp signal (Figure 1.16). At the end of $t_f$, S1 switches and connects the output of the voltage reference $U_R$ to the integrator input. The voltage reference output is negative for a positive input voltage. The integrator output starts to increase, following a positive-slope ramp (Figure 1.16). The process stops when the ramp attains the 0 V level, and the comparator allows the control logic to switch S1 again. The period of time $t_v$ the ramp takes to increase to 0 V is variable and depends on the ramp peak value attained during period $t_f$.

The relationship between the input voltage $U_i$ and the time periods $t_v$ and $t_f$ is given by:

$$\frac{1}{RC}\int_0^{t_f} U_i dt = \frac{t_v}{RC} U_R \tag{1.45}$$

that, for a constant input voltage $U_i$, leads to:

$$U_i = U_R \frac{t_v}{t_f} \tag{1.46}$$

Since the same integrating circuit is used, errors due to comparator offset, capacitor tolerances, long-term counter clock drifts, and integrator nonlinearities are eliminated. High resolutions are therefore possible, although the speed of reading is low (in the order of milliseconds).

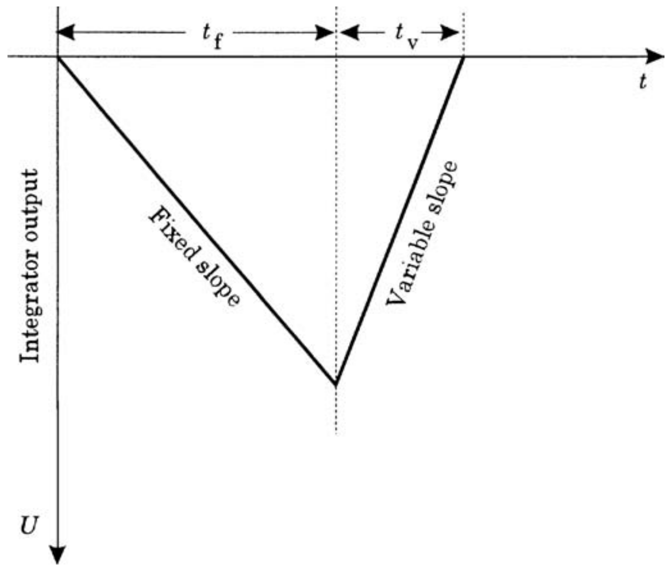



**FIGURE 1.16** Integrator output signal in a dual slope DVM.

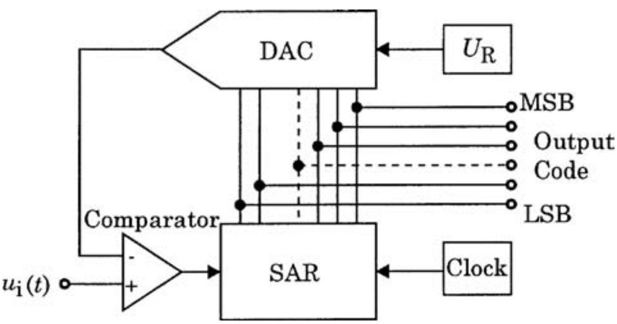



**FIGURE 1.17** Successive approximation ADC schematics.

Slowly varying voltages can be also measured by dual slope DVMs. However, this requires that the input signal does not vary for a quantity greater than the DVM resolution during the reading time. For high-resolution DVMs, this limits the DVM bandwidth to a few hertz.

### Successive Approximation ADC.

The successive approximation technique represents the most popular technique for the realization of ADCs. Figure 1.17 shows the block diagram of this type of converter. The input voltage is assumed to have a constant value $U_i$ and drives one input of the comparator. The other comparator's input is driven by the output of the digital-to-analog converter (DAC), which converts the binary code provided by the successive approximation register (SAR) into an analog voltage. Let $n$ be the number of bits of the converter, $U_R$ the voltage reference output, and $C$ the code provided by the SAR. The DAC output voltage is then given by:

$$U_c = \frac{C}{2^n} U_R \tag{1.47}$$

When the conversion process starts, the SAR most significant bit (MSB) is set to logic 1. The DAC output, according to Equation 1.47, is set to half the reference value, and hence half the analog input

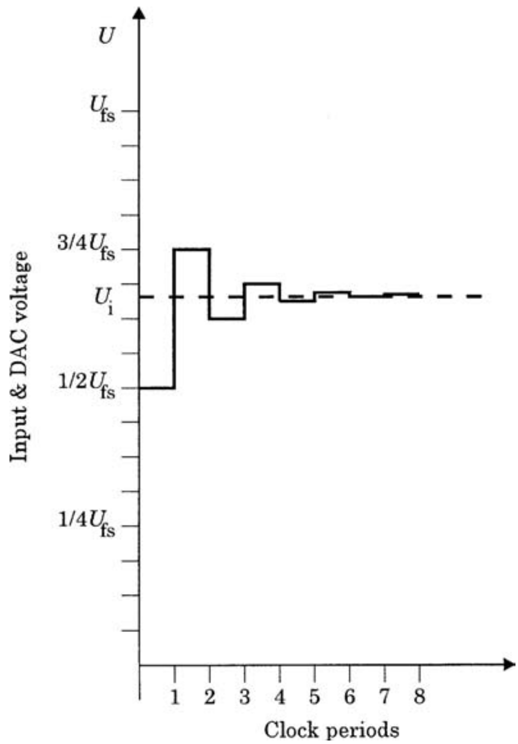



**FIGURE 1.18** DAC output signal in a successive approximation ADC.

full-scale range. The comparator determines whether the DAC output is above or below the input signal. The comparator output controls the SAR in such a way that, if the input signal is above the DAC output, as shown in Figure 1.18, the SAR MSB is retained and the next bit is set to logic 1.

If now the input signal is below the DAC output (Figure 1.18), the last SAR bit set to logic 1 is reset to logic 0, and the next one is set to logic 1. The process goes on until the SAR least significant bit (LSB) has been set. The entire conversion process takes time $t_c = nT_c$, where $T_c$ is the clock period. At the end of conversion, the SAR output code represents the digitally converted value of the input analog voltage $U_i$.

According to Equation 1.47, the ADC resolution is $U_R/2n$, which corresponds to 1 LSB. The conversion error can be kept in the range ±1/2 LSB. Presently, a wide range of devices is available, with resolution from 8 to 16 bits, and conversion rates from 100 ms to below 1 ms.

Varying voltages can be sampled and converted into digital by the ADC, provided the input signal does not vary by a quantity greater than $U_R/2n$ during the conversion period $t_c$. The maximum frequency of an input sine wave that satisfies this condition can be readily determined starting from given values of $n$ and $t_c$.

Let the input voltage of the ADC be an input sine wave with peak-to-peak voltage $U_{pp} = U_R$ and frequency $f$. Its maximum variation occurs at the zero-crossing time and, due to the short conversion period $t_c$, is given by $2\pi f t_c U_{pp}$. To avoid conversion errors, it must be:

$$2\pi f t_c U_{pp} \le \frac{U_R}{2^n} \tag{1.48}$$

Since $U_{pp} = U_R$ is assumed, this leads to:

$$f \le \frac{1}{2^n 2\pi t_c} \tag{1.49}$$

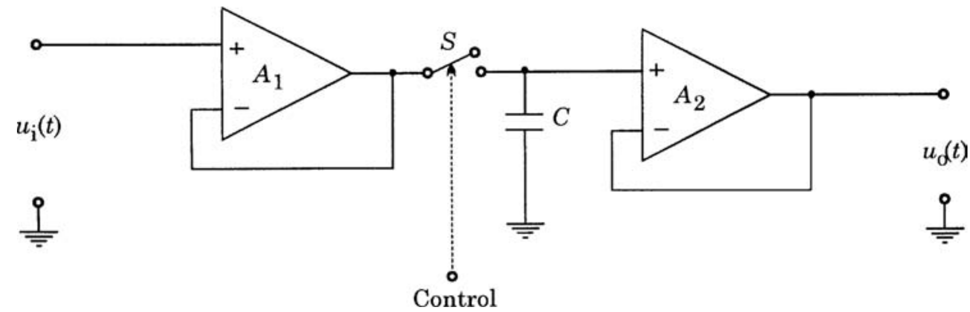



**FIGURE 1.19** Sample and Hold schematics.

If $t_c = 1$ ms and $n = 12$, Equation 1.49 leads to $f \leq 38.86$ Hz. However, ADCs can still be employed with input signals whose frequency exceeds the value given by Equation 1.49, provided that a *Sample and Hold* circuit is used to keep the input voltage constant during the conversion period.

The *Sample and Hold* circuit is shown in Figure 1.19. When the electronic switch $S$ is closed, the output voltage $u_o(t)$ follows the input voltage $u_i(t)$. When switch $S$ is open, the output voltage is the same as the voltage across capacitor $C$, which is charged at the value assumed by the input voltage at the time the switch was opened. Due to the high input impedance of the operational amplifier $A_2$, if a suitable value is chosen for capacitor $C$, its discharge transient is slow enough to keep the variation of the output voltage below the ADC resolution.

## Ac Digital Voltmeters.

True rms ac voltmeters with digital reading can be obtained using an electronic circuit like the one in Figure 1.14 to convert the rms value into a dc voltage signal, and measuring it by means of a DVM. However, this structure cannot actually be called a digital structure, because the measurement is attained by means of analog processing of the input signal.

A more modern approach, totally digital, is shown in Figure 1.20. The input signal $u_i(t)$ is sampled at constant sampling rate $f_s$, and converted into digital by the ADC. The digital samples are stored in the memory of the digital signal processor (DSP) and then processed in order to evaluate Equation 1.40 in a numerical way. Assuming that the input signal is periodic, with period $T$, and its frequency spectrum is upper limited by harmonic component of order $N$, the sampling theorem is satisfied if at least $(2N + 1)$ samples are taken over period $T$ in such a way that $(2N + 1)T_s = T$, $T_s = 1/f_s$ being the sampling period [6, 7]. If $u_i(kT_s)$ is the $k^{\text{th}}$ sample, the rms value of the input signal is given by, according to Equation 1.40:

$$U^2 = \sqrt{\frac{1}{2N+1}\sum_{k=0}^{2N} u_i^2(kT_s)} \tag{1.50}$$

This approach can feature a relative uncertainty as low as ±0.1% of full scale, with an ADC resolution of 12 bits. The instrument bandwidth is limited to half the sampling frequency, according to the sampling theorem. When modern ADCs and DSPs are employed, a 500 kHz bandwidth can be obtained. Wider bandwidths can be obtained, but with a lower ADC resolution, and hence with a lower accuracy.

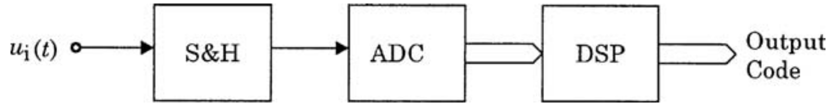



**FIGURE 1.20** Block diagram of a modern digital meter.

**Frequency Response of Ac Voltmeters.**

When the frequency response of ac voltmeters is taken into account, a distinction must be made between the analog voltmeters (both electromechanical and electronic) and digital voltmeters, based on DSP techniques.

The frequency response of the analog meters is basically a low-pass response, well below 1 kHz for most electromechanical instruments, and up to hundreds of kilohertz for electronic instruments.

When digital, DSP-based meters are concerned, the sampling theorem and aliasing effects must be considered. To a first approximation, the frequency response of a digital meter can be considered flat as long as the frequency-domain components of the input signal are limited to a frequency band narrower than half the sampling rate. If the signal components exceed this limit (the so-called Nyquist frequency), the aliasing phenomenon occurs [6]. Because of this phenomenon, the signal components at frequencies higher than half the sampling rate are folded over the lower frequency components, changing them. Large measurement errors occur under this situation.

To prevent the aliasing, a low-pass filter must be placed at the input stage of any digital meter. The filter cut-off frequency must ensure that all frequency components above half the sampling rate are negligible. If the low-pass, anti-aliasing filter is used, the digital DSP-based meters feature a low-pass frequency response also.

## References

1. M.B. Stout, *Basic Electrical Measurements,* Englewood Cliffs, NJ, Prentice-Hall, 1960.
2. I.F. Kinnard, *Applied Electrical Measurements,* New York, John Wiley & Sons, Chapman & Hall, Ltd. London, 1956.
3. B.M. Oliver and J.M. Cage, *Electronic Measurements and Instrumentation,* London, McGraw-Hill, Inc. 1975.
4. T.T. Lang, *Electronics of Measuring Systems,* New York, John Wiley & Sons, 1987.
5. Analog Devices, *Analog-Digital Conversion Handbook,* Englewood Cliffs, NJ, Prentice-Hall, 1986.
6. A.V. Oppenheim and R.W. Schafer, *Digital Signal Processing,* Englewood Cliffs, NJ, Prentice-Hall, 1975.
7. A. Ferrero and R. Ottoboni, High-accuracy Fourier analysis based on synchronous sampling techniques. *IEEE Trans. Instr. Meas.,* 41(6), 780-785, 1992.

# 1.2 Oscilloscope Voltage Measurement

*Jerry Murphy*

Engineers, scientists, and other technical professionals around the world depend on oscilloscopes as one of the primary voltage measuring instruments. This is an unusual situation because the oscilloscope is not the most accurate voltage measuring instrument usually available in the lab. It is the graphical nature of the oscilloscope that makes it so valued as a measurement instrument — not its measurement accuracy.

The oscilloscope is an instrument that presents a graphical display of its input voltage as a function of time. It displays voltage waveforms that cannot easily be described by numerical methods. For example, the output of a battery can be completely described by its output voltage and current. However, the output of a more complex signal source needs additional information such as frequency, duty cycle, peak-to-peak amplitude, overshoot, preshoot, rise time, fall time, and more to be completely described. The oscilloscope, with its graphical presentation of complex waveforms, is ideally suited to this task. It is often described as the "screwdriver of the electronic engineer" because the oscilloscope is the most fundamental tool that technical professionals apply to the problem of trying to understand the details of the operation of their electronic circuit or device. So, what is an oscilloscope?

The oscilloscope is an electronic instrument that presents a high-fidelity graphical display of the rapidly changing voltage at its input terminals.

The most frequently used display mode is voltage vs. time. This is not the only display that could be used, nor is it the display that is best suited for all situations. For example, the oscilloscope could be called on to produce a display of two changing voltages plotted one against the other, such as a Lissajous display. To accurately display rapidly changing signals, the oscilloscope is a high bandwidth device. This means that it must be capable of displaying the high-order harmonics of the signal being applied to its input terminals in order to correctly display that signal.

## The Oscilloscope Block Diagram

The oscilloscope contains four basic circuit blocks: the vertical amplifier, the time base, the trigger, and the display. This section treats each of these in a high-level overview. Many textbooks exist that cover the details of the design and construction of each of these blocks in detail [1]. This discussion will cover these blocks in enough detail so that readers can construct their own mental model of how their operation affects the application of the oscilloscope for their voltage measurement application. Most readers of this book have a mental model of the operation of the automatic transmission of an automobile that is sufficient for its successful operation but not sufficient for the overhaul or redesign of that component. It is the goal of this section to instill that level of understanding in the operation of the oscilloscope. Those readers who desire a deeper understanding will get their needs met in later sections.

Of the four basic blocks of the oscilloscope, the most visible of these blocks is the display with its *cathode-ray tube* (CRT). This is the component in the oscilloscope that produces the graphical display of the input voltage and it is the component with which the user has the most contact. Figure 1.21 shows the input signal is applied to the vertical axis of a cathode ray tube. This is the correct model for an analog oscilloscope but it is overly simplified in the case of the digital oscilloscope. The important thing to learn from this diagram is that the input signal will be operated on by the oscilloscope's vertical axis circuits so that it can be displayed by the CRT. The differences between the analog and digital oscilloscope are covered in later sections.

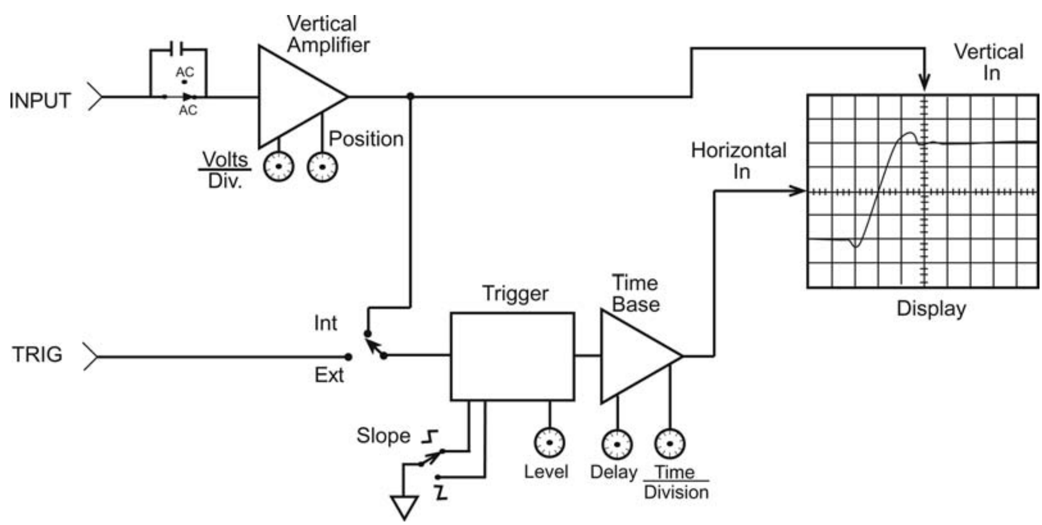



**FIGURE 1.21** Simplified oscilloscope block diagram that applies to either analog or digital oscilloscopes. In the case of the digital oscilloscope, the vertical amplifier block will include the ADC and high-speed waveform memory. For the analog scope the vertical block will include delay lines with their associated drivers and a power amplifier to drive the CRT plates.

The *vertical amplifier* conditions the input signal so that it can be displayed on the CRT. The vertical amplifier provides controls of volts per division, position, and coupling, allowing the user to obtain the desired display. This amplifier must have a high enough bandwidth to ensure that all of the significant frequency components of the input signal reach the CRT.

The *trigger* is responsible for starting the display at the same point on the input signal every time the display is refreshed. It is the stable display of a complex waveform that allows the user of an oscilloscope to make judgments about that waveform and its implications as to the operation of the device under test.

The final piece of the simplified block diagram is the *time base*. This circuit block is also known as the horizontal system in some literature. The time base is the part of the oscilloscope that causes the input signal to be displayed as a function of time. The circuitry in this block causes the CRT beam to be deflected from left to right as the input signal is being applied to the vertical deflection section of the CRT. Controls for time-per-division and position (or delay) allow the user of the oscilloscope to adjust the display for the most useful display of the input signal. The time-per-division controls of most oscilloscopes provide a wide range of values, ranging from a few nanoseconds ($10^{-9}$ s) to seconds per division. To get a feeling for the magnitude of the dynamic range of the oscilloscope's time base settings, keep in mind that light travels about 1 m in 3 ns.

## The Oscilloscope as a Voltage Measurement Instrument

That the oscilloscope's vertical axis requires a wide bandwidth amplifier and its time base is capable of displaying events that are as short as a few nanoseconds apart, indicates that the oscilloscope can display rapidly changing voltages. Voltmeters, on the other hand, are designed to give their operator a numeric readout of steady-state or slowly changing voltages. Voltmeters are not well suited for displaying voltages that are changing levels very quickly. This can be better understood by examination of the operation of a voltmeter as compared to that of an oscilloscope. The analog voltmeter uses the magnetic field produced by current flowing through a coil to move the pointer against the force of a spring. This nearly linear deflection of the voltmeter pointer is calibrated by applying known standard voltages to its input. Therefore, if a constant voltage is applied to the coil, the pointer will move to a point where the magnetic force being produced by the current flowing in its coil is balanced by the force of the spring. If the input voltage is slowly changing, the pointer will follow the changing voltage. This mechanical deflection system limits the ability of this measurement device to the measurement of steady-state or very low-frequency changes in the voltage at its input terminals. Higher-frequency voltmeters depend on some type of conversion technique to change higher frequencies to a dc signal that can be applied to the meter's deflection coil. For example, a diode is used to rectify ac voltages to produce a dc voltage that corresponds to the average value of the ac voltage at the input terminals in average responding ac voltmeters.

The digital voltmeter is very much like the analog meter except that the mechanical displacement of the pointer is replaced with a digital readout of the input signal. In the case of the digital voltmeter, the input signal is applied to an analog-to-digital converter (ADC) where it is compared to a reference voltage and digitized. This digital value of the input signal is then displayed in a numerical display. The ADC techniques applied to voltmeters are designed to produce very accurate displays of the same signals that were previously measured with analog meters. The value of a digital voltmeter is its improved measurement accuracy as compared to that of its analog predecessors.

The oscilloscope will display a horizontal line displaced vertically from its zero-voltage level when a constant, or dc voltage is applied to its input terminals. The magnitude of this deflection of the oscilloscope's beam vertically from the point where it was operating with no input being applied is how the oscilloscope indicates the magnitude of the dc level at its input terminals. Most oscilloscopes have a graticule as a part of their display and the scope's vertical axis is calibrated in volts per division of the graticule. As one can imagine, this is not a very informative display of a dc level and perhaps a voltmeter with its numeric readout is better suited for such applications.

There is more to the scope–voltmeter comparison than is obvious from the previous discussion. That the oscilloscope is based on a wide-bandwidth data-acquisition system is the major difference between

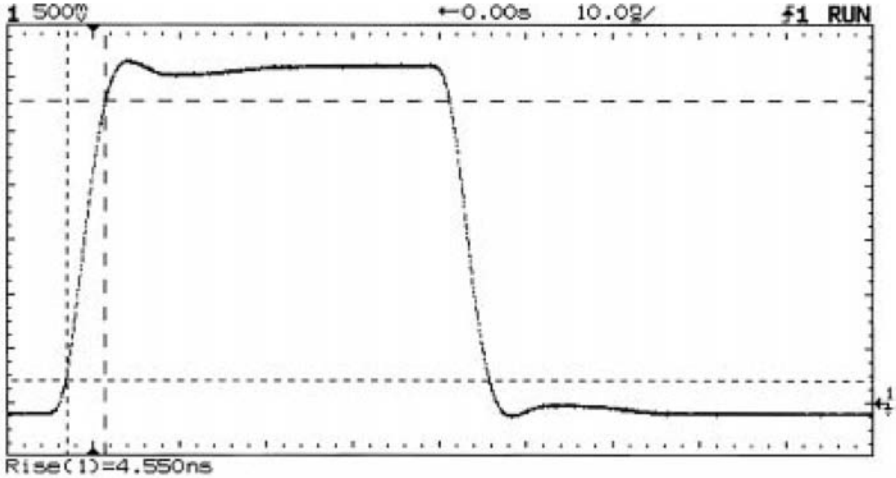



**FIGURE 1.22** A typical complex waveform. This waveform is described by measurements of its amplitude, offset, risetime, falltime, overshoot, preshoot, and droop.

these two measurement instruments. The oscilloscope is designed to produce a high fidelity display of rapidly changing signals. This puts additional constraints on the design of the oscilloscope's vertical system that are not required in the voltmeter. The most significant of these constraints is that of a constant group delay. This is a rather complex topic that is usually covered in network analysis texts. It can be easily understood if one realizes the effect of group delay on a complex input signal.

Figure 1.22 shows such a signal. The amplitude of this signal is a dc level and the rising edge is made up of a series of high-frequency components. Each of these high-frequency components is a sine wave of specific amplitude and frequency. Another example of a complex signal is a square wave with a frequency of 10 MHz. This signal is made up of a series of odd harmonics of that fundamental frequency. These harmonics are sine waves of frequencies of 10 MHz, 30 MHz, 50 MHz, 70 MHz, etc. So, the oscilloscope must pass all of these high-frequency components to the display with little or no distortion. Group delay is the measure of the propagation time of each component through the vertical system. A constant group delay means that each of these components will take the same amount of time to propagate through the vertical system to the CRT, independent of their frequencies. If the higher-order harmonics take more or less time to reach the scope's deflection system than the lower harmonics, the resulting display will be a distorted representation of the input signal. Group delay (in seconds) is calculated by taking the first derivative of an amplifier's phase-vs.-frequency response (in radians/(l/s)). If the amplifier has a linearly increasing phase shift with frequency, the first derivative of its phase response will be a horizontal line corresponding to the slope of the phase plot (in seconds). Amplifier systems that have a constant group delay are known as Gaussian amplifiers. They have this name because their pass band shape resembles that of the bell curve of a Gaussian distribution function (Figure 1.23). One would think that the oscilloscope's vertical amplifier should have a flat frequency response, but this is not the case because such amplifiers have nonconstant group delay [1].

The oscilloscope's bandwidth specification is based on the frequency where the vertical deflection will be –3 dB (0.707) of the input signal. This means that if a constant 1 V sine wave is applied to the oscilloscope's input, and the signal's frequency is adjusted to higher and higher frequencies, the oscilloscope's bandwidth will be that frequency where its display of the input signal has been reduced to be 0.707 V. Noticable errors in ampitude measurements will start at 20% of the scope's bandwidth. The oscilloscope's error-free display of complex waveforms gives it poor voltage accuracy. For the measurement of dc and single frequency signals such as sine waves, other instruments can produce more accurate measurements.

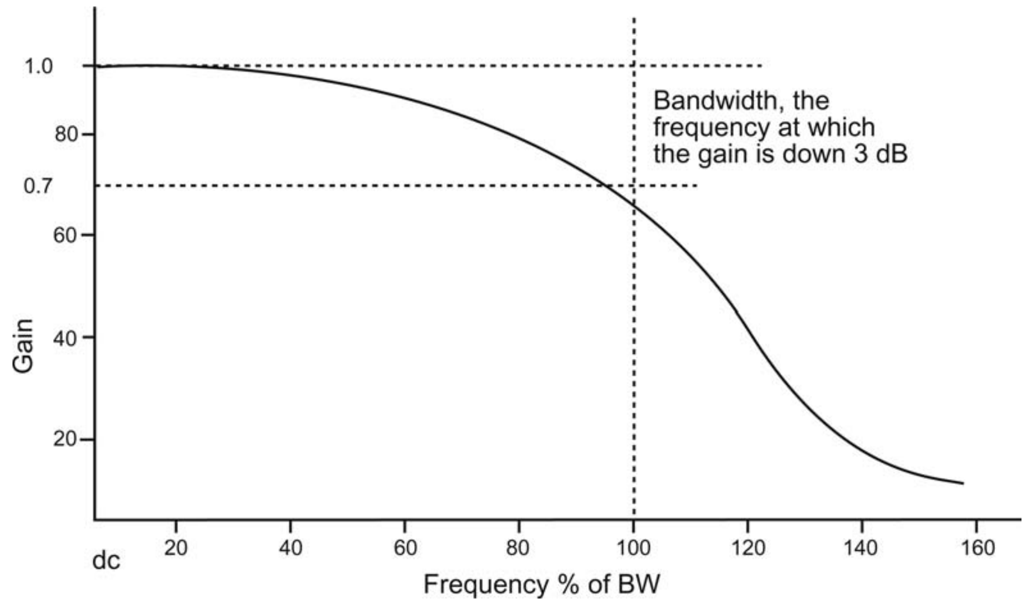



**FIGURE 1.23** The Gaussian frequency response of the oscilloscope's vertical system which is not flat in its pass band. Amplitude measurements made at frequencies greater than 20% of the scope's bandwidth will be in error.

*Conclusion:* The voltmeter makes the most accurate measurements of voltages that are dc, slowly changing, or can be converted to a dc analog of their ac content. The oscilloscope is not the most accurate voltage measurement instrument, but it is well suited to measurements of voltages that are changing very rapidly as a function of time. Oscilloscopes are the instrument of choice for observing and characterizing these complex voltages.

## Analog or Digital

The world of oscilloscopes is divided into two general categories: analog and digital. The first oscilloscopes were analog. These products are based on the direct-view vector cathode-ray tube (DVVCRT or CRT for short). The analog oscilloscope applies the input signal to the vertical deflection plates of the CRT where it causes the deflection of a beam of high-energy electrons moving toward the phosphor-coated faceplate. The electron beam generates a lighted spot where it strikes the phosphor. The intensity of the light is directly related to the density of the electrons hitting a given area of the phosphor. Because this analog operation is not based on any digitizing techniques, most people have little trouble creating a very accurate and simple mental model in their minds of its operation.

The analog oscilloscope produces a display of the input signal that is bright and easy to see under most conditions. It can also contain as many as 16 shades of gray-scale information. This means that an event that occurs less frequently will appear at a lower intensity in the display than another event that occurs more frequently. This oscilloscope does not produce a continous display of the input signal. It is blind during retrace and trigger hold-off times. Because the display depends on the production of visible light from the phosphor being excited by an electron beam, the display must be refreshed frequently. This makes the analog oscilloscope a low-dead-time display system that can follow rapidly changing signals. Also, there is little lag time in front panel control settings.

The analog oscilloscope is not without its shortcomings. The strength of the analog oscilloscope is its CRT, but this is also the source of its weaknesses. The biggest problem with analog scopes is their dependence on a display that is constantly being refreshed. This means that these scopes do not have any waveform storage. If the input signal fails to repeat frequently, the display will simply be a flash of light

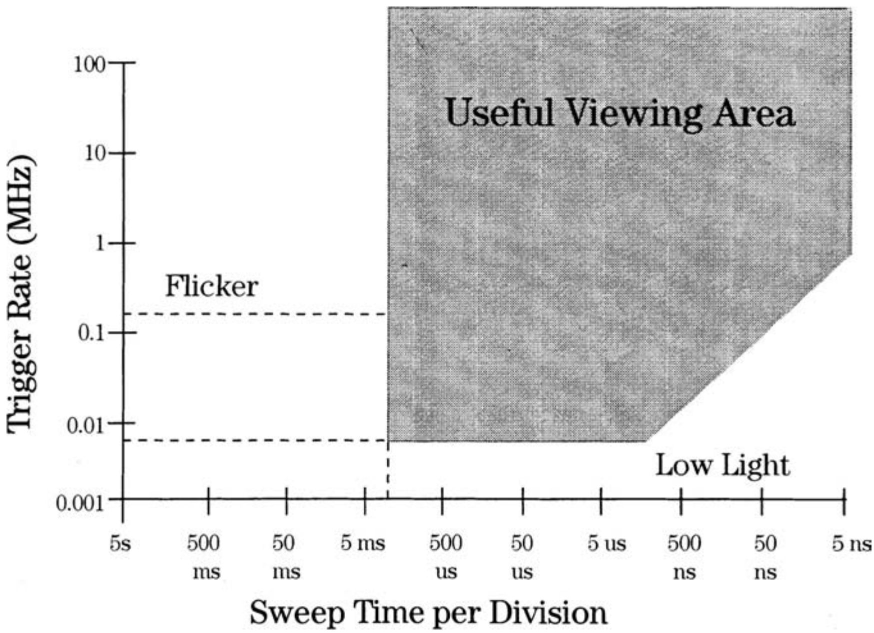



**FIGURE 1.24** The operating range of the analog oscilloscope. This is a plot of input signal repetition rate from the lower limit of single shot to the full bandwidth of the scope plotted against sweep speed. The shaded area is the area where the analog oscilloscope will produce a usable display.

when the beam sweeps by the phosphor. If the signal's repetition rate falls below 100 Hz, the display will flicker annoyingly. Figure 1.24 shows a plot of the range of an input signal's repetition frequency range from a single-shot event to the full bandwidth of a scope vs. the scope's sweep speeds. The result is a map of the scope's operational area. Figure 1.24 shows that the analog oscilloscope fails to map onto the full range of possible input signals and sweep speeds.

Another problem of the analog oscilloscope is its inability to display information ahead of its trigger. This is a problem in applications where the only suitable trigger is at the end of the event of interest. Another limitation of analog scopes is their timing accuracy. The time base of the analog scope is based on the linearity of a voltage ramp. There are other sources of errors in the analog oscilloscope's horizontal axis, but the sweep nonlinearity is the major contributor. This results in these scopes having a timing accuracy of typically ±3% of their full-scale setting. Therefore, if the time base is set to 100 ns/div, in order to view a 100 ns wide pulse, the full scale will be 1000 ns or 1 ms. The accuracy of this pulse width measurement will be ±30 ns or ±30% of the pulse width!

The digital oscilloscope or digital storage oscilloscope (DSO) differs from its analog counterpart in that the input signal is converted to digital data and therefore it can be managed by an embedded microprocessor. The waveform data can have correction factors applied to remove errors in the scope's acquisition system and can then be stored, measured, and/or displayed. That the input signal is converted from analog to digital and manipulations are performed on it by a microprocessor results in people not having a good mental model of the digital oscilloscope's operation. This would not be a problem except for the fact that the waveform digitizing process is not totally free from errors, and a lack of a correct mental model of the scope's operation on the part of its user can increase the odds of a measurement error. To make matters worse, various manufacturers of these products make conflicting claims, making it easy to propagate incorrect mental models of the digital scope's operation. It is the intention of this presentation to give the information needed to create a mental model of the operation of these devices that will enable the user to perform error-free measurements with ease.

The digital storage oscilloscope offers many advantages over its analog counterpart. The first is accuracy. The voltage measurement accuracy of the digital oscilloscope is better than that of an analog scope because the microprocessor can apply correction factors to the data to correct for errors in the calibration of the scope's vertical system. The timing accuracy of a digital oscilloscope is an order of magnitude better than that of an analog scope. The digital scope can store the waveform data for comparison to other test results or uploading to a computer for analysis or project documentation. The digital oscilloscope does not depend on the input signal being continuously updated to produce an easily viewable display. A single-shot event is displayed at the same brightness level as a signal that repeats in time periods corresponding to the full bandwidth of the scope.

The disadvantages of the digital oscilloscope are its more complex operation, aliasing, and display performance. The analog-to-digital conversion process [1] is used to convert the input signal into a series of discrete values, or samples, uniformly spaced in time, which can be stored in memory. Voltage resolution is determined by the total number of codes that can be produced. A larger number permits a smoother and more accurate reproduction of the input waveform but increases both the cost and difficulty in achieving a high sample frequency. Most digital oscilloscopes provide 8-bit resolution in their ADC. As the ADC's sampling speed is increased, the samples will be closer together, resulting in smaller gaps in the waveform record.

All digital scopes are capable of producing an aliased display. Some models are more prone to this problem than others, but even the best will alias under the right conditions. An alias is a lower frequency false reproduction of the input signal resulting from under-sampling, i.e., sampling less than the Nyquist frequency. The display of the digital scope is based on computer display technology. This results in a display that is very bright and easy to see, even under conditions where an analog scope would have difficulty in producing a viewable display. The disadvantage of the digital scope's display is its lower horizontal resolution. Most of the scopes on the market have a raster scan display with a resolution of 500 lines, less than half the resolution of an analog scope's display. This is not a problem in most applications. It could become a factor where very complex waveforms, such as those found in TV systems, are being analyzed. Many digital scopes have display systems that exhibit large dead- or blind-times. Scopes based on a single CPU will be able to display their waveform data only after the CPU has finished all of its operations. This can result in a display that is unresponsive to front panel control inputs as well as not being able to follow changes in the input signal.

Table 1.2 shows that both analog and digital oscilloscopes have relative advantages and disadvantages. All the major producers of oscilloscopes are pushing the development of digital scopes in an attempt to overcome their disadvantages. All the major producers of these products believe that the future is digital. However, a few manufacturers produce scopes that are both analog and digital. These products appear to have the best of both worlds; however, they have penalties with respect to both cost and complexity of operation.

**TABLE 1.2** A Comparison of Analog and Digital Oscilloscopes

| | Analog Oscilloscope | Digital Oscilloscope |
|---|---|---|
| Operation | Simple | Complex |
| Front panel controls | Direct access knobs | Knobs and menus |
| Display | Real-time vector | Digital raster scan |
| Gray scales | >16 | >4 |
| Horizontal resolution | >1000 lines | 500 lines |
| Dead-time | Short | Can be long |
| Aliasing | No | Yes |
| Voltage accuracy | ±3% of full scale | ±3% of full scale |
| Timing accuracy | ±3% of full scale | ±0.01% of full scale |
| Single shot capture | None | Yes |
| Glitch capture | Limited | Yes |
| Waveform storage | None | Yes |
| Pretrigger viewing | None | Yes |
| Data out to a computer | No | Yes |

One of the driving forces making scope manufacturers believe that the future of the digital oscilloscope is bright is that modern electronic systems are becoming ever more digital in nature. Digital systems place additional demands on the oscilloscope that exceed the capabilities of the analog scope. For example, often in digital electronic systems, there is a need to view fast events that occur at very slow or infrequent rates. Figure 1.24 shows that these events fail to be viewable on analog scopes. Another common problem with digital systems is the location of trigger events. Often the only usable trigger is available at the end of the event being viewed. Analog scopes can only display events that occur after a trigger event. The rapid growth of digital electronics that occurred in the late 1990s is being attributed to the lowering of the cost of single-chip microcontrollers. These devices, which contain a complete microprocessor on one integrated circuit, are responsible for the "electronics everywhere" phenomenon, where mechanical devices are becoming electronic as well as those devices that were previously electrical in nature. In 1996, Hewlett Packard introduced a new class of oscilloscope designed to meet the unique needs of the microcontrollerbased applications. This new class of oscilloscope is known as the mixed signal oscilloscope or MSO [2].

## Voltage Measurements

Voltage measurements are usually based on comparisons of the waveform display to the oscilloscope's graticule. Measurements are made by counting the number of graticule lines between the end-points of the desired measurement and then multiplying that number by the sensitivity setting. This was the only measurement available to most analog scope users, and it is still used by those performing troubleshooting with their digital scope as a time-saving step. (Some late-model analog oscilloscopes incorporate cursors to enhance their measurement ability.) For example, a waveform that is 4.5 divisions high at a vertical sensitivity of 100 mV/div would be 450 mV high.

Switching the scope's coupling between ac and dc modes will produce a vertical shift in the waveform's position that is a measurement of its dc component. This technique can be applied to either analog or digital scopes. Simply note the magnitude of the change in waveform position and multiply by the channel's sensitivity.

Additional measurements can be performed with an analog oscilloscope but they usually require more skill on the part of the operator. For example, if the operator can determine the location of the top and base of a complex waveform, its amplitude can be measured. Measurements based on percentages can be made using the scope's vernier to scale the waveform so that its top and bottom are 5 divisions apart. Then, each division represents 20% of the amplitude of the waveform being studied. The use of the vernier, which results in the channel being uncalibrated, prevents performance of voltage measurements. Many analog scopes have a red light to warn the operator that the scope is uncalibrated when in vernier mode.

The digital oscilloscope contains an embedded microprocessor that automates the measurement. This measurement automation is based on a histogramming technique, where a histogram of all the voltages levels in the waveform are taken from the oscilloscope's waveform data. The histogram is a plot of the voltage levels in the waveform plotted against the number of samples found at each voltage level. Figure 1.25 shows the histogramming technique being applied to the voltage measurements of complex waveforms.

## Understanding the Specifications

The oscilloscope's vertical accuracy is one place that a person's mental model of the scope's operation can lead to measurement trouble. For example, the oscilloscope's vertical axis has a frequency response that is not flat across its pass band. However, as noted above, the scope has a Gaussian frequency response to produce the most accurate picture of complex signals. This means that the oscilloscope's accuracy specification of ±3% is a dc-only specification. If one were to attempt to measure the amplitude of a signal whose frequency is equal to the bandwidth of the scope, one would have to add another 29.3% to the error term, for a total error of ±32.3%. This is true for both analog and digital oscilloscopes. This limitation can be overcome by carefully measuring the frequency response of the oscilloscope's vertical

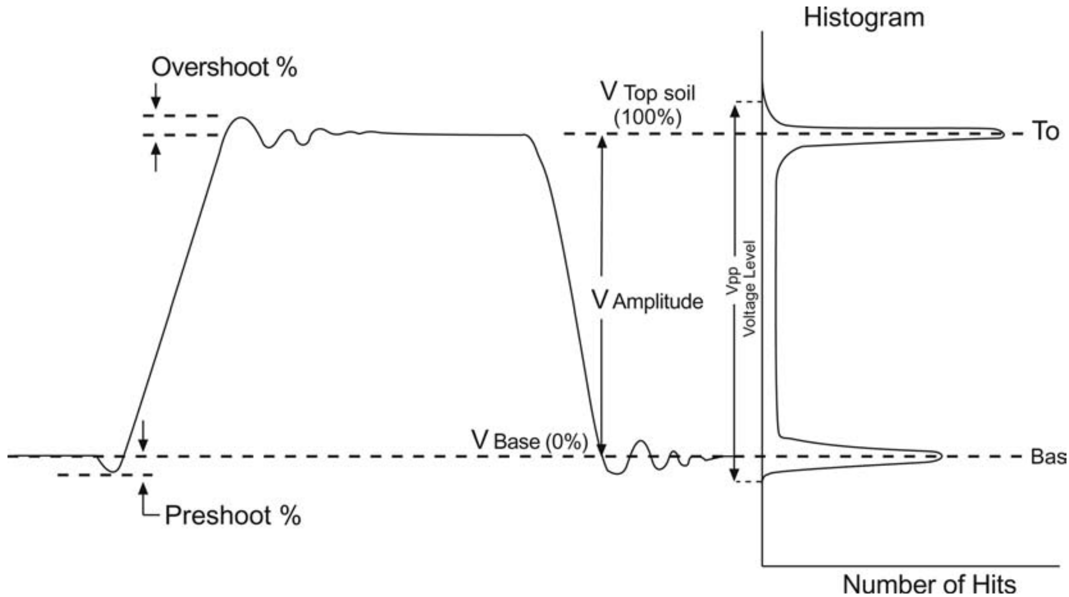



**FIGURE 1.25** Voltage histograms as applied by a digital oscilloscope. The complex waveform is measured by use of the voltage histogram. This histogram is a plot of each voltage level in the display and the number of data points at that level.

channels. One will need to repeat this process every time the scope is serviced or calibrated, because the various high-frequency adjustments that may need to be made in the scope's vertical axis will affect the scope's frequency response. One is probably asking, why don't the manufacturers do this for me? The answer is twofold. The first is cost, and the second is that this is not the primary application of an oscilloscope. There are other instruments that are much better suited to the measurement of high-frequency signals. The spectrum analyzer would be this author's first choice.

Additionally, the vertical accuracy is a full-scale specification. This means that at 1 V/div, the full-scale value is typically 8 V. The measurement error for a scope with a ±3% specification under these conditions will be ±0.24 V. If the signal being measured is only 1 V high, the resulting measurement will be ±24% of reading. Check the manual for the scope being used, as some manufacturers will specify full-scale as being 10 or even 10.2 divisions. This will increase the error term because the full-scale term is larger.

In digital oscilloscopes, the vertical accuracy is often expressed as a series of terms. These attempt to describe the analog and digital operations the scope performs on the input signal. Terms might include digitizing resolution, gain, and offset (sometimes called as position). They also might be called out as single and dual cursor accuracies. The single cursor accuracy is a sum of all three terms. In the dual cursor case, where the voltage measurement is made between two voltage cursors, the offset term will cancel out, leaving only the digitizing resolution and gain errors. For example, the Hewlett Packard model 54603B has a single cursor accuracy specification of ±1.2% of full scale, ±0.5% of position value, and a dual cursor specification of ±0.4% of full scale.

**HINT:** Always try to make the voltage measurements on the largest possible vertical and widest possible display of the signal.

The horizontal accuracy specifications of analog and digital scopes are very different; however, both are based on a full-scale value. In the analog scope, many manufacturers limit accuracy specifications to only the center eight divisions of their display. This means that a measurement of a signal that starts or ends in either the first or ninth graticule, will be even more error prone than stated in the scope's specifications. To the best of this author's knowledge, this limitation does not apply to digital scopes. The horizontal specifications of digital scopes are expressed as a series of terms. These might include the crystal accuracy, horizontal display resolution, and trigger placement resolution. These can be listed as cursor accuracy. For example, the Hewlett Packard model 54603B has a horizontal cursor accuracy specification of ±0.01% ±0.2% full-scale ±200 ps. In this example, the first term is the crystal accuracy, the second is the display resolution (500 lines), and the final term is twice the trigger placement error. By comparing the analog and digital scopes' horizontal specifications, it can be seen that in either case, the measurement is more accurate if it can be made at full screen. The digital scope is more accurate than its analog counterpart.

Digital scopes also have acquisition system specifications. Here is another place where the operator's mental model of the operation of a digital scope can produce measurement errors. All manufacturers of digital scopes specify the maximum sampling speed of their scope's acquisition system as well as its memory depth and number of bits. The scope's maximum sampling speed does not apply to all sweep speeds, only memory depth and number of bits applies to all sweep speeds. The scope's maximum sampling speed applies only to its fastest sweep speeds.

The complexity of the digital scope results from the problem of having to sample the input. There is more to be considered than Nyquist's Sampling Theorem in the operation of a digital scope. For example, how does the scope's maximum sampling rate relate to the smallest time interval that the scope can capture and display? A scope that samples at 100 MSa $s^{-1}$ takes a sample every 10 ns; therefore, in principle, it cannot display any event that is less than 10 ns wide because that event will fall between the samples. In practice, however, this limit can — under certain circumstances — be extended. If the scope is operating in an "equivalent time" or "random repetitive" mode and if the signal is repetitive, even if very infrequently, the scope will be able to capture any event that is within its vertical system bandwidth. Figure 1.26 shows an infrequently occurring pulse that is 25 ns wide embedded into a data stream being captured and displayed on an oscilloscope with a maximum sampling speed of 20 MSa $s^{-1}$ (sampling interval of 50 ns). Figure 1.26(b) shows this pulse at a faster sweep speed. An analog scope would produce a similar display of this event, with the infrequent event being displayed at a lower intensity than the rest of the trace. Notice that the infrequent event does not break the baseline of the trace.

The correct mental model of the digital scope's ability to capture signals needs to be based on the scope's bandwidth, operating modes, and timing resolution. It is the timing resolution that tells the operator how closely spaced the samples can be in the scope's data record.

The most common flaw in many mental models of the operation of a digital scope is related to its maximum sampling speed specification. As noted, the maximum sampling speed specification applies only to the scope's fastest sweep speeds. Some scope manufacturers will use a multiplex A/D system that operates at its maximum sampling speed only in single-channel mode. The scope's memory depth determines its sampling speed at the sweep speed being used for any specific measurement. The scope's memory depth is always equal to the scope's horizontal full-scale setting. For scopes with no off-screen memory, this is 10′ the time base setting. If the scope has off-screen memory, this must be taken into account. For example, assume that one has two scopes with a maximum sampling speed of 100 MSa $s^{-1}$. One scope has a memory depth of 5 K points and the other only 1 K. At a sweep speed of 1 ms per division, both scopes will be able to store data into their memory at their full sampling speed, and each will be storing 100 data points per division, for a total of 1000 data points being stored. The scope with the 5 K memory will have a data point in one of every 5 memory locations, and the scope with the 1 K memory will have a data point in every memory location. If one reduces the sweep speed to 5 ms/div, the deeper memory scope will now fill every one of its memory locations with data points separated by 10 ns. The scope with only 1 K of memory would produce a display only 2 divisions wide if its sampling speed is not reduced. Scope designers believe that scope users expect to see a full-length sweep at every sweep speed. Therefore, the 1 K scope must reduce its sampling speed to one sample every 50 ns, or 20 MSa $s^{-1}$, to be able to fill its memory with a full sweep width of data. This 5:1 ratio of sampling speeds between these two scopes will be maintained as their time bases are set to longer and longer sweeps. For example, at 1 s/div, the 5 K scope will be sampling at 500 samples per second, while the 1 K scope will be sampling at only 100 samples per second. One can determine a scope's sampling speed for any specific time base setting from Equation 1.51.

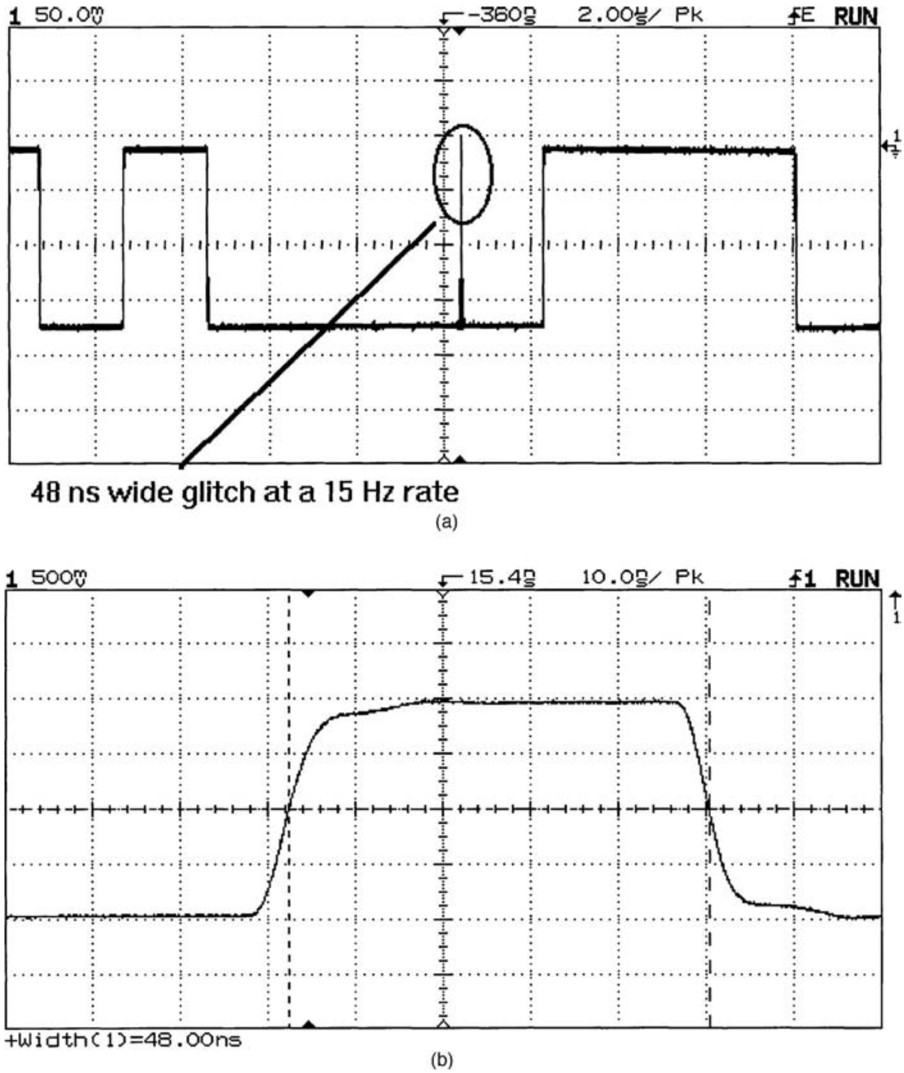



**FIGURE 1.26** An infrequently occurring event as displayed on a digital oscilloscope with random repetitive sampling. (a) The event embedded in a pulse train. (b) Shows the same event at a faster sweep speed. The fact that the waveform baseline is unbroken under the narrow pulse indicates that it does not occur in every sweep. The width of this pulse is less than half the scope's sampling period in (b). Both traces are from a Hewlett Packard model 54603B dual channel 60 MHz scope.

$$S\left(\text{samples/second}\right) = \frac{\text{memory depth}\left(\text{samples}\right)}{\text{full-scale time base}\left(\text{seconds}\right)}, \tag{1.51}$$

$$\text{or the scope's maximum sampling speed, whichever is less}$$

One must look closely at the application to determine if a specific scope is best suited to that application. As a rule, the deeper the memory, the faster the scope will be able to sample the signal at any given time base setting. Memory depth is not free. High-speed memory required to be able to store the data out of the scope's A/D is costly, and deeper memory takes longer to fill, thus reducing the scope's display update rate. Most scopes that provide memory depths of 20 K or more will also give the user a memory depth selection control so that the user can select between fast and deep. (In 1996, Hewlett Packard Co.

introduced two scopes based on an acquisition technology known as MegaZoom (TM) [10] that removes the need for a memory depth control.) A correct mental model for the sampling speed of a digital scope is based on Equation 1.51 and not just on the scope's maximum performance specifications.

Some digital oscilloscopes offer a special sampling mode known as *peak detection*. Peak detection is a special mode that has the effect of extending the scope's sampling speed to longer time records. This special mode can reduce the possibility of an aliased display. The performance of this special mode is specified as the minimum pulse width that the peak detection system can capture. There are several peak detection systems being used by the various manufacturers. Tektronix has an analog-based peak detection system in some of its models, while Hewlett Packard has a digital system in all of its models. Both systems perform as advertised, and they should be evaluated in the lab to see which system best meets one's needs. There is a downside to peak detection systems and that is that they display high-frequency noise that might not be within the bandwidth of the system under test. Figure 1.27 shows a narrow pulse being captured by peak detection and being missed when the peak detection is off.

What effect does display dead-time have on the oscilloscope's voltage measurement capabilities? Display dead-time applies to both analog and digital oscilloscopes, and it is that time when the oscilloscope is not capturing the input signal. This is also a very important consideration in the operation of a digital scope because it determines the scope's ability to respond to front-panel control commands and to follow changing waveforms. A digital scope that produces an incorrect display of an amplitude-modulated signal is not following this rapidly changing signal because its display update rate is too low. Sampling speed is not related to display update rate or dead-time. Display dead-time is a function of the scope's ability to process the waveform data from its A/D and plot it on the display. Every major oscilloscope manufacturer has been working on this problem. Tektronix offers a special mode on some of its products known as InstaVu (TM) [4]. This special mode allows these scopes to process up to 400,000 waveforms per second to their display. Hewlett Packard has developed a multiple parallel processor technology [5] in the HP 54600 series of benchtop scopes that provides a high-speed, low dead-time display in a lowcost instrument. These instruments can plot 1,500,000 points per second to their display and they have no dead-time at their slower sweep speeds. LeCroy has been applying the Power PC as an embedded processor for its scopes to increase display throughput. There are other special modes being produced by other vendors, so be sure to understand what these can do before selecting an oscilloscope. Figure 1.28 shows the effect of display update rate on a rapidly changing waveform. An amplitude-modulated signal is displayed with a high-speed display and with the display speed reduced by the use of hold-off.

## Triggering

The trigger of the oscilloscope has no direct effect on the scope's ability to measure a voltage except that the trigger does enable the oscilloscope to produce a stable display of the voltage of interest. Ref. [6] presents a thorough discussion of this subject.

## Conclusion

The mental model that oscilloscope users have created in their minds of the oscilloscope's operation can be helpful in reducing measurement errors. If the operator's mental model is based on the following facts, measurement errors can be minimized:

- Oscilloscopes have a frequency response that affects measurement accuracy.
- Digital scopes are more accurate than analog scopes.
- Analog scopes do not have continuous displays.
- Oscilloscope accuracy specifications always contain a percent of full-scale term.
- Measurements should be made at the largest possible deflection in order to minimize errors.
- Maximum sampling speed is available only at the scope's fastest sweep speeds.
- Deeper memory depth allows faster sampling at more sweep speeds.

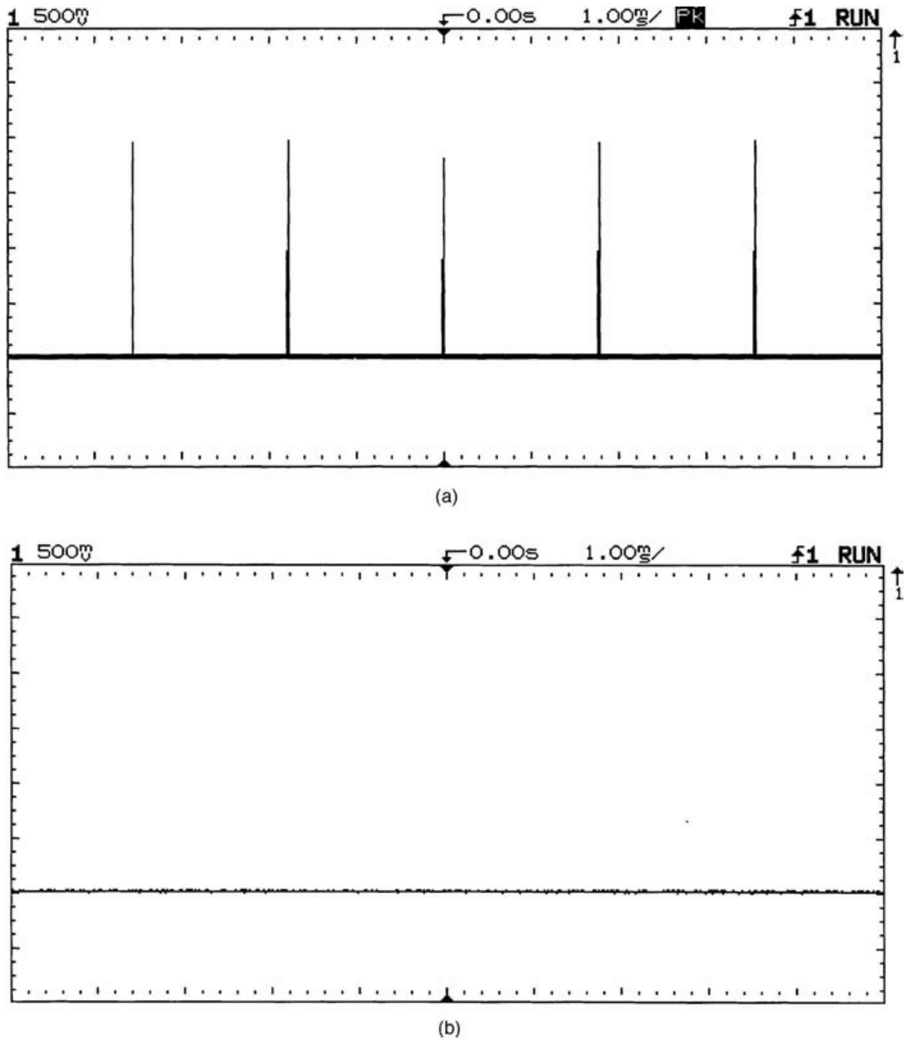



**FIGURE 1.27** Peak detection. This special mode has the effect of increasing the scope's sampling speed at time base settings where it would be decimated. In operation, each memory location contains either the maximum or minimum value of the waveform at that location in time. (a) A series of 300 ns wide pulses being captured at a slow sweep speed; (b) the same setup with peak detection disabled. These narrow pulses would appear as intermittent pulses if the scope could be seen in operation with peak detection disabled.

- All digital scopes can produce aliases, some more than others.
- Display dead-time is an important characteristic of digital scopes that is often not specified.
- Display dead-time affects measurement accuracy because it can cause a distorted display.
- The scope with the highest maximum sampling speed specification might not be the most accurate or have the lowest display dead-time.
- The operator must have some knowledge of the signals being measured to be able to make the best possible measurements.

The person who has the mental model of the oscilloscope that takes these factors into account will be able to purchase the scope that is best suited to his/her application and not spend too much money on unnecessary performance. In addition, that person will be able to make measurements that are up to the full accuracy capabilities of the scope.

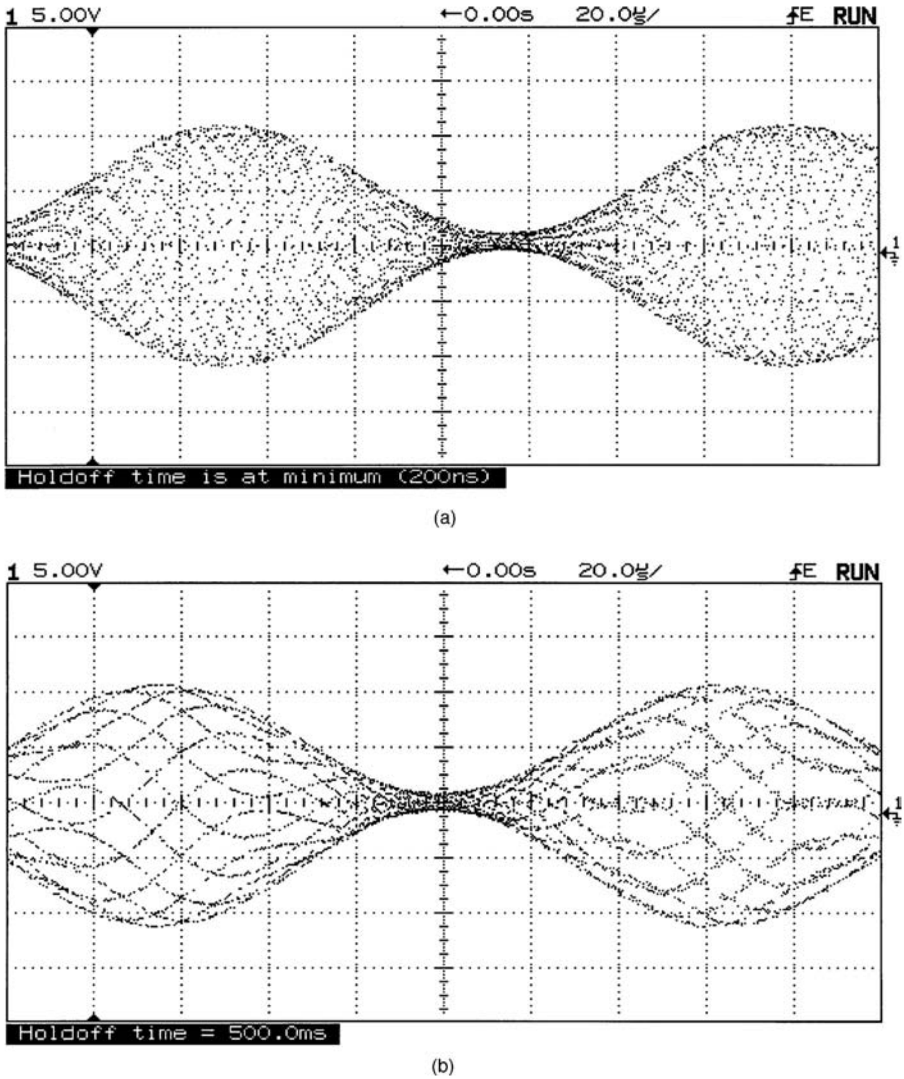



**FIGURE 1.28** Display dead-time. The time that an oscilloscope is blind to the input signal has an effect on the scope's ability to correctly display rapidly changing signals. (a) An amplitude-modulated signal with a high-speed display; (b) the same signal with the dead-time increased by use of hold-off.

## Selecting the Oscilloscope

There are ten points to consider when selecting an oscilloscope. This author has published a thorough discussion of these points [7] and they are summarized as follows:

1. **Analog or Digital?** There are a few places where the analog scope might be the best choice, and the reader can make an informed selection based on the information presented here.
2. **How much bandwidth?** This is a place where the person selecting an oscilloscope can save money by not purchasing more bandwidth than is needed. When analog oscilloscopes were the only choice, many people were forced to purchase more bandwidth than they needed because they needed to view infrequent or low repetition signals. High-bandwidth analog scopes had brighter CRTs so that they were able to display high-frequency signals at very fast time base settings. At a sweep speed of 5 ns/div, the phosphor is being energized by the electron beam for 50 ns, so the

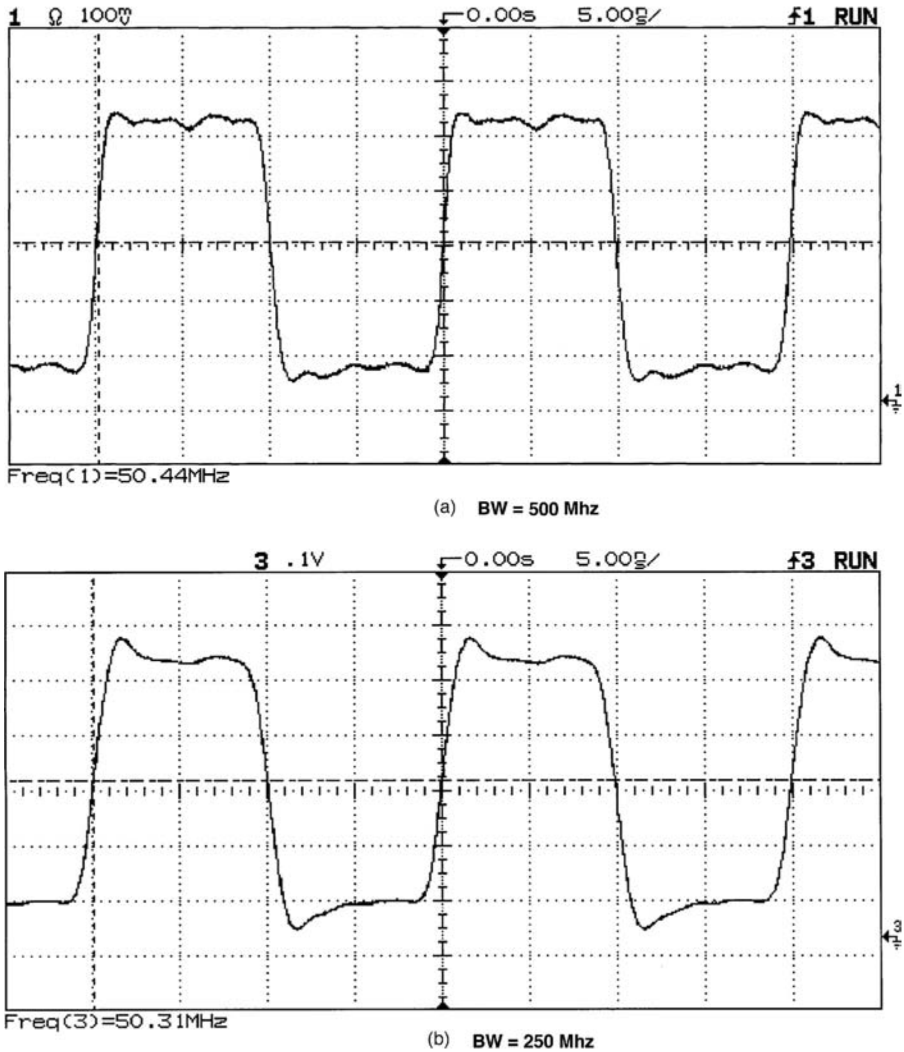


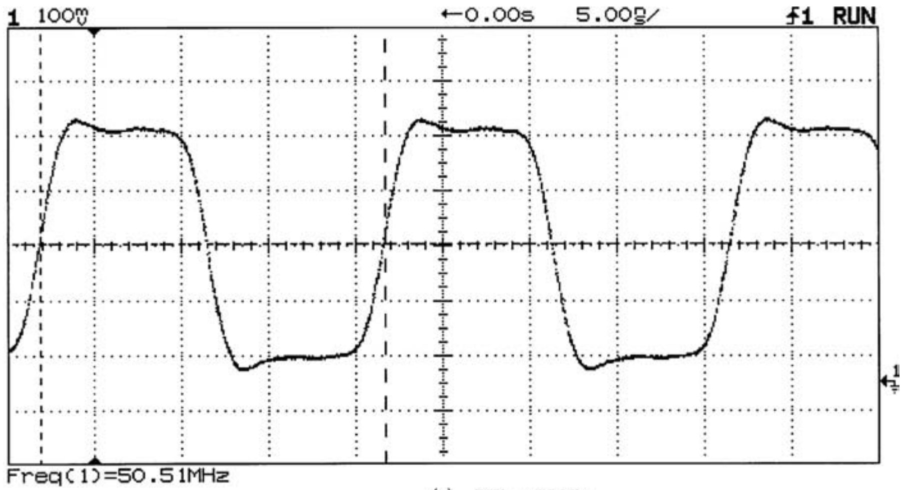

**FIGURE 1.29** The effect of the scope's bandwidth is shown in this set of waveforms. The same 50 MHz square wave is shown as it was displayed on scopes of 500 MHz in Figure 1.28(a) all the way down to 20 MHz in Figure 1.29(e). Notice that the 100 MHz scope produced a usable display although it was missing the high-frequency details of the 500 MHz display. The reason that the 100 MHz scope looks so good is the fact that its bandwidth is slightly greater than 100 MHz. This performance, which is not specified on any data sheet, is something to look for in any evaluation.

electron beam had to be very high energy to produce a visible trace. This situation does not apply to digital scopes. Now, one needs to be concerned only with the bandwidth required to make the measurement. Figure 1.29 shows the effect of oscilloscope bandwidth on the display of a 50 MHz square wave.

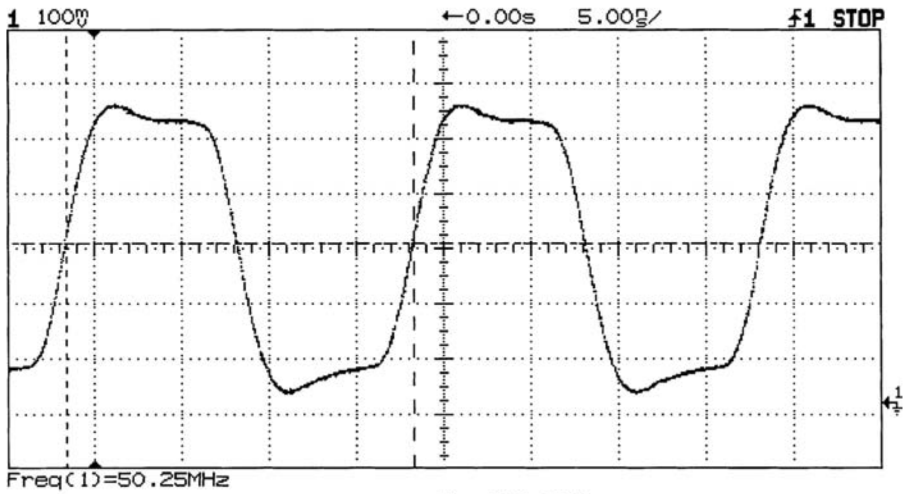
The oscilloscope's bandwidth should be >2× the fundamental highest frequency signal to be measured.

The bandwidth of the scope's vertical system can affect the scope's ability to correctly display narrow pulses and to make time interval measurements. Because of the scope's Gaussian frequency response, one can determine its ability to correctly display a transient event in terms of risetime with Equation 1.52.

$$t_r = 0.35/\mathrm{BW} \tag{1.52}$$




(c) **BW = 100 Mhz**




(d) **BW = 60 Mhz**

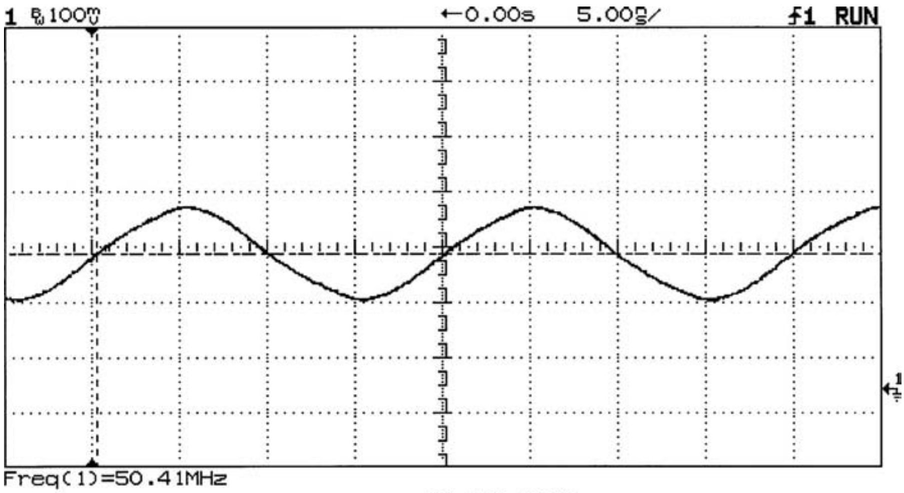
**FIGURE 1.29** (continued)

Therefore, a 100 MHz scope will have a risetime of 3.5 ns. This means that if the scope were to have a signal at its input with zero risetime edges, it would be displayed with 3.5 ns edges. This will affect the scope's measurements in two ways. First is narrow pulses. Figure 1.30 shows the same 5 ns wide pulse being displayed on oscilloscopes of 500 MHz and 60 MHz bandwidths, and the effect of the lower bandwidth on this event that is closest to the risetime of the slower scope is apparent.

The second is fast time interval measurements. A measurement of signal risetime is an example. The observed risetime on the scope's display is according to Equation 1.53.

$$t_{\text{observed}} = \left(t_{\text{signal}}{}^2 + t_{\text{scope}}{}^2\right)^{1/2} \tag{1.53}$$

If a 10 ns risetime were to be measured with a 100 MHz scope, one would obtain a measurement of 10.6 ns based on Equation 1.53. The scope would have made this measurement with a 6% reading error before any other factors, such as time base accuracy, are considered.




(e) **BW = 20 Mhz**

**FIGURE 1.29** (continued)

The scope's risetime should be at least no more than 1/5 of the shortest time interval to be measured. For time interval measurements, this should be >1/10.

3. **How many channels?** Most oscilloscopes in use today are dual-channel models. In addition, there are models described as being 2+2 and four channels. This is one time where 2+2 is not equal to 4. The 2+2 models have limited features on two of their channels and cost less than 4-channel models. Most oscilloscope suppliers will hold the 4-channel description only for models with four full-featured channels, but the user should check the model under consideration so as to be sure if it is a 4- or 2+2 model. Either of the four channel classes is useful for applications involving the testing and development of digital-based systems where the relationship of several signals must be observed.

   Hewlett Packard introduced a new class of oscilloscopes that is tailored for the applications involving both analog and digital technologies, or mixed-signal systems. The mixed signal oscilloscope (MSO) [4] provides 2 scope channels and 16 logic channels so that it can display both the analog and digital operation of a mixed-signal system on its display.
4. **What sampling speed?** Do not simply pick the scope with the highest banner specification. One needs to ask, what is the sampling speed at the sweep speeds that my application is most likely to require? As observed in Equation 1.51 the scope's sampling speed is a function of memory depth and full-scale time base setting. If waveforms are mostly repetitive, one can save a lot of money by selecting an oscilloscope that provides equivalent time or random repetitive sampling.
5. **How much memory?** As previously discussed, memory depth and sampling speed are related. The memory depth required depends on the time span needed to measure and the time resolution required. The longer the time span to be captured and the finer the resolution required, the more memory one will need. High-speed waveform memory is expensive. It takes time to process a longer memory, so the display will have more dead-time in a long memory scope than a shallow memory model. All the suppliers of deep memory scopes provide a memory depth control. They provide this control so that the user can choose between a high-speed display and deep memory for the application at hand. Hewlett Packard introduced MegaZoom (TM) technology [3] in 1996; it produces a high-speed low dead-time display with deep memory all the time.
6. **Triggering?** All scope manufacturers are adding new triggering features to their products. These features are important because they allow for triggering on very specific events. This can be a valuable troubleshooting tool because it will let the user prove whether a suspected condition

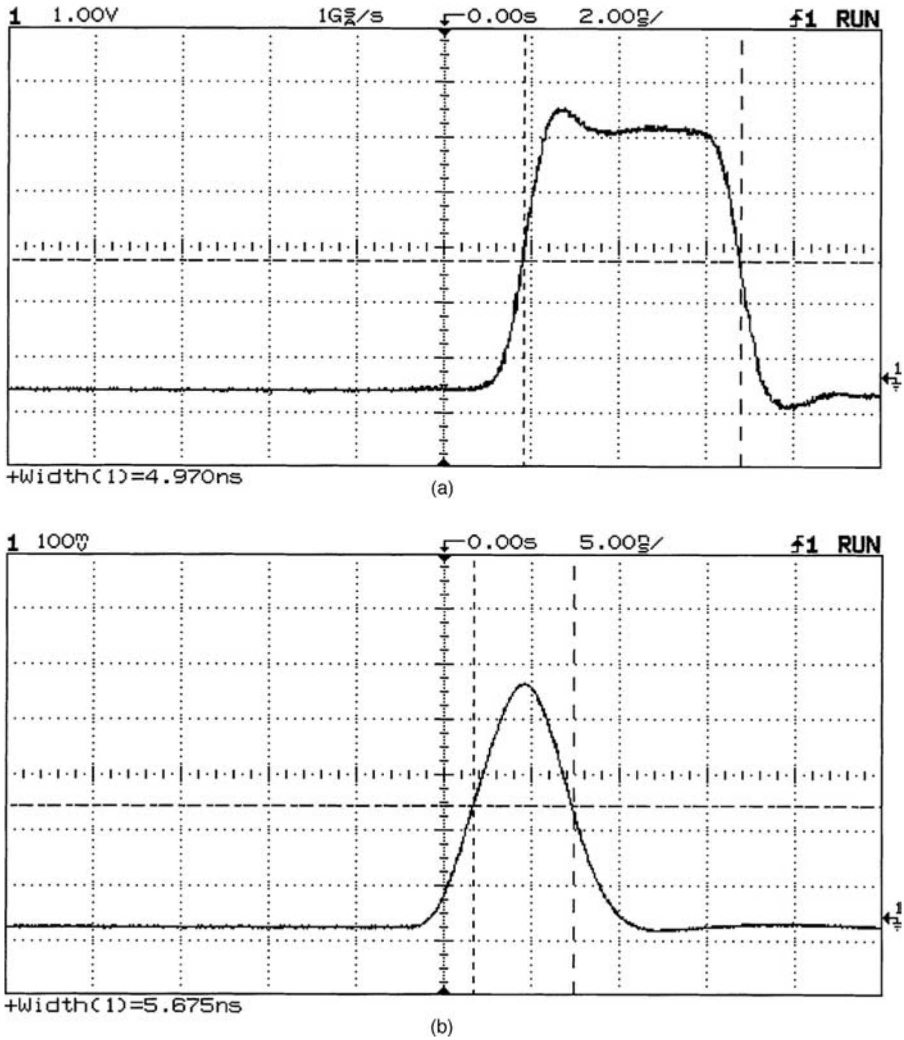



**FIGURE 1.30** Bandwidth and narrow events. (a) A 5 ns wide pulse as displayed on a 500 MHz scope; (b) the same pulse displayed on a 60 MHz scope. The 60 MHz scope has a risetime of 5.8 ns, which is longer than the pulse width. This results in the pulse shape being incorrectly displayed and its amplitude being in error.

exists or not. Extra triggering features add complexity to the scope's user interface; so be sure to try them out to make sure that they can be applied.

7. **Trustworthy display?** Three factors critically affect a scope's ability to display the unknown and complex signals that are encountered in oscilloscope applications. If the user loses confidence in the scope's ability to correctly display what is going on at its probe tip, productivity will take a real hit. These are display update rate, dead-time, and aliasing.

   Because all digital scopes operate on sampled data, they are subject to aliasing. An alias is a false reconstruction of the signal caused by under-sampling the original. An alias will always be displayed as a lower frequency than the actual signal. Some vendors employ proprietary techniques to minimize the likelihood of this problem occurring. Be sure to test any scope being considered for purchase on your worse-case signal to see if it produces a correct or aliased display. Do not simply test it with a single-shot signal that will be captured at the scope's fastest sweep speed because this will fail to test the scope's ability to correctly display signals that require slower sweep speeds.

**TABLE 1.3** Major Suppliers of Oscilloscopes and their Web Addresses

| Vendor | Description | Web address |
|---|---|---|
| B&K Precision<br>6460 W. Cortland St.<br>Chicago, IL 60635 | Analog and digital scopes and Metrix scopes in France | http://bkprecision.com |
| Boonton Electronics Corp.<br>25 Estmans Road<br>P.O. Box 465<br>Parsippany, NJ 07054-0465 | U.S. importer for Metrix analog, mixed analog, and digital scopes from France | http://www.boonton.com |
| Fluke<br>P.O. Box 9090<br>Everett, WA 98206-9090 | Hand-held, battery-powered scopes (ScopeMeter), analog scopes, and CombiScopes(R) | http://www.fluke.com |
| Gould<br>Roebuck Road, Hainault,<br>Ilford, Essex IG6 3UE, England | 200 MHz DSO products | http://www.gould.co.uk |
| Hewlett Packard Co.<br>Test & Measurement<br>Mail Stop 51LSJ<br>P.O. Box 58199<br>Santa Clara, CA 95052-9952 | A broad line of oscilloscopes and the Mixed Signal oscilloscope for technical professionals | http://www.tmo.hp.com/tmo/pia search on "oscilloscopes" |
| LeCroy Corp.<br>700 Chestnut Ridge Road<br>Chestnut Ridge, NY 10977 | Deep memory oscilloscopes for the lab | http://www.lecroy.com |
| Tektronix Inc.<br>Corporate Offices<br>26600 SW Parkway<br>P.O. Box 1000<br>Watsonville, OR 97070-1000 | The broad line oscilloscope supplier with products ranging from hand-held to high-performance lab scopes | http://www.tek.com/measurement search on "oscilloscopes" |
| Yokogawa Corp. of America<br>Corporate offices<br>Newnan, GA<br>1-800-258-2552 | Digital oscilloscopes for the lab | http://www.yca.com |

8. **Analysis functions?** Digital oscilloscopes with their embedded microprocessors have the ability to perform mathematical operations that can give additional insight into waveforms. These operations often include addition, subtraction, multiplication, integration, and differentiation. An FFT can be a powerful tool, but do not be misled into thinking it is a replacement for a spectrum analyzer. Be sure to check the implementation of these features in any scope being considered. For example, does the FFT provide a selection of window functions? Are these analysis functions implemented with a control system that only their designer could apply?
9. **Computer I/O?** Most of the digital scopes on the market today can be interfaced to a PC. Most of the scope manufacturers also provide some software that simplifies the task of making the scope and PC work together. Trace images can be incorporated into documents as either PCX or TIF files. Waveform data can be transferred to spreadsheet applications for additional analysis. Some scope models are supplied with a disk drive that can store either waveform data or trace images.
10. **Try it out?** Now one has the information to narrow oscilloscope selection to a few models based on bandwidth, sampling speed, memory depth, and budget requirements. Contact the scope vendors (Table 1.3) and ask for an evaluation unit. While the evaluation unit is in the lab, look for the following characteristics:

- Control panel responsiveness: Does the scope respond quickly to inputs or does it have to think about it for a while?
- Control panel layout: Are the various functions clearly labeled? Does the user have to refer to the manual even for simple things?
- Display speed: Turn on a couple of automatic measurements and check that the display speed remains fast enough to follow changing signals.
- Aliasing: Does the scope produce an alias when the time base is reduced from fast to slow sweep speeds? How does the display look for the toughest signal?

The oscilloscope is undergoing a period of rapid change. The major manufacturers of oscilloscopes are no longer producing analog models and the digital models are evolving rapidly. There is confusion in the oscilloscope marketplace because of the rapid pace of this change. Hopefully, this discussion will prove valuable to the user in selecting and applying oscilloscopes in the lab in the years to come.

### References

1. A. DeVibiss, Oscilloscopes, in C.F. Coombs, Jr. (ed.), *Electronic Instrument Handbook,* 2nd ed., New York, McGraw-Hill, 1995.
2. R. A. Witte, A family of instruments for testing mixed-signal circuits and systems, *Hewlett Packard J.,* April 1996, Hewlett Packard Co., Palo Alto, CA.
3. M.S. Holcomb, S.O. Hall, W.S. Tustin, P.J. Burkart, and S.D. Roach, Design of a mixed signal oscilloscope, *Hewlett Packard J.,* April 1996, Hewlett Packard Co., Palo Alto, CA.
4. InstaVu acquisition mode, *Tektronix Measurement Products Catalog,* Tektronix Inc., Beaverton, OR, 1996, 69.
5. M.S. Holcomb and D.P. Timm, A high-throughput acquisition architecture for a 100 MHz digitizing oscilloscope, *Hewlett Packard J.,* February 1992, Hewlett Packard Co., Palo Alto,CA.
6. R.A. Witte, *Elecronic Test Instruments, Theory and Applications,* Englewood Cliffs, NJ, Prentice-Hall, 1993.
7. J. Murphy, Ten points to ponder in picking an oscilloscope, *IEEE Spectrum,* 33(7), 69-77, 1996.

## 1.3 Inductive and Capacitive Voltage Measurement

*Cipriano Bartoletti, Luca Podestà, and Giancarlo Sacerdoti*

This chapter section addresses electrical measurements where the voltage range to be measured is very large — from $10^{-10}$ V to $10^{7}$ V. The waveform can be continuous, periodic, or impulsive. If it is periodic, the spectrum components can vary for different situations, and within the same electric power network there may be subharmonic components. In impulsive voltage measurement, it is often important to get maximum value, pulse length, etc. Capacitive and inductive voltage sensors are mainly utilized in low-frequency electric measurements.

### Capacitive Sensors

The voltage to be measured can be reduced by means of capacitive dividers (Figure 1.31). Capacitive dividers are affected by temperature and frequency and therefore are not important, at least in Europe. Capacitive sensors detect voltage by different methods:

1. Electrostatic force (or torque)
2. Kerr or Pockels effect
3. Josephson effect
4. Transparency through a liquid crystal device
5. Change in refractive index of the optic fiber or in light pipe

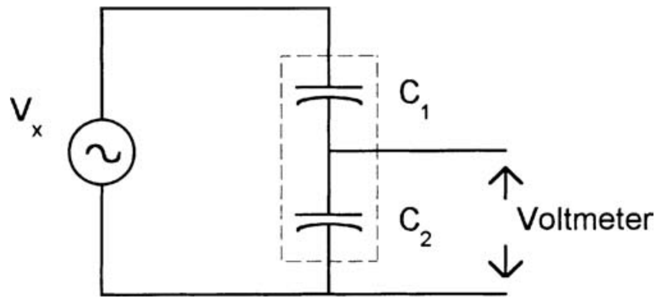



**FIGURE 1.31** Schematic arrangement of a capacitive divider.

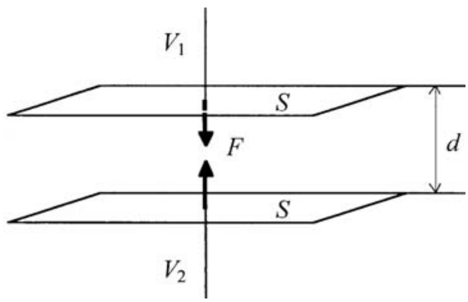



**FIGURE 1.32** Force between two electrodes with an applied voltage.

1. The relations that rule the listed capacitive voltage sensors are reported below. The force between two electrodes is (Figure 1.32):

$$F = \varepsilon_0 \frac{S}{d}\left(V_1 - V_2\right)^2 \tag{1.54}$$

where $\varepsilon_0$ = Dielectric constant
$S$ = Area of the electrode
$d$ = Distance
$V_1$, $V_2$ = Potentials of the electrodes

The *torque* between electrostatic voltmeter quadrants (Figure 1.33) is given by:

$$T = \frac{1}{2}\frac{\partial C}{\partial \theta}\left(V_1 - V_2\right)^2 \tag{1.55}$$

where $C$ = Capacitance
$\theta$ = Angle between electrodes

To get the torque from the rate of change (derivative) of electrostatic energy vs. the angle is easy. Obtaining the torque by mapping the electric field is difficult and requires long and complex field computing.

2. The rotation of the polarization plane of a light beam passing through a KDP crystal under the influence of an electric field (*Pockels effect*) is expressed by (Figure 1.34):

$$\theta = k_\pi l\left(V_1 - V_2\right) \tag{1.56}$$

where $k_\pi$ = Electro-optic constant
$l$ = Length of crystal

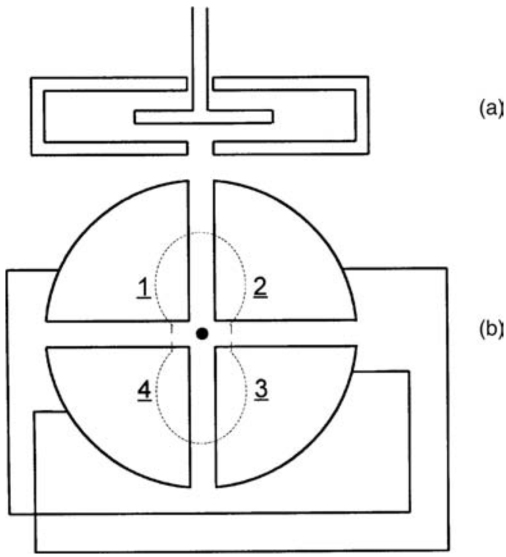



**FIGURE 1.33** Scheme of an electrostatic voltmeter. (a) Lateral view; (b) top view: (1), (2), (3), (4) are the static electrodes; the moving vane is shown in transparency.

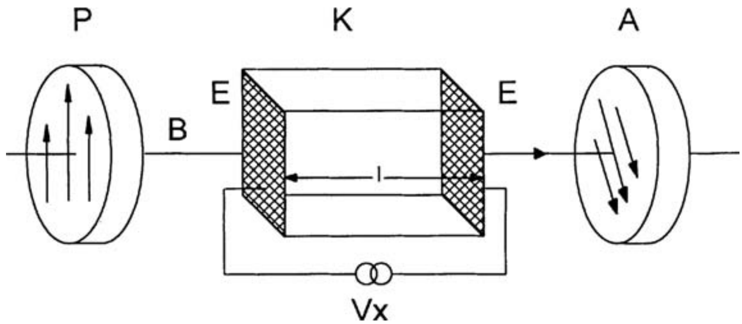



**FIGURE 1.34** Scheme of an electrooptic KDP device. The parts are labeled as: (B) a light beam, (P) a polarizer, (A) an analyzer, (K) a KDP crystal, with the voltage to be measured $V$x applied to its (E) transparent electrodes.

One obtains a rotation of $\pi/2$ by applying a voltage of the order of 1 kV to a KDP crystal of a few centimeters in length.

If a light beam passes through a light pipe that performs the *Kerr effect*, one observes a quadratic dependence of the rotation vs. $V$.

$$\theta \equiv kE^2 \equiv k'V^2 \tag{1.57}$$

3. The *Josephson effect* consists of translation of a voltage into a periodical signal of a certain frequency, carried out by a special capacitive sensor. There is an array of $N$ layers of Josephson superconducting junctions; the frequency of emitted signal, when a voltage $V$ is applied, is given by:

$$v = \frac{2eV}{Nh} \tag{1.58}$$

4. The *transparency* of a liquid crystal device depends on the difference of potential applied. There are liquid crystal devices working in transmission or in reflection. A change in transparency is obtained when a difference of potential of a few volts is applied.

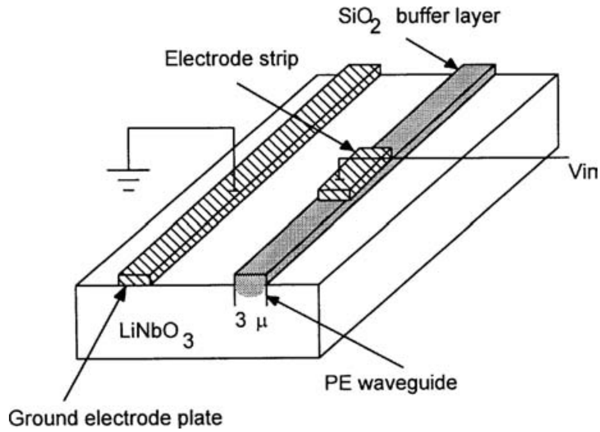



**FIGURE 1.35** Li-Nb optical wave guide device.

5. The *change in refractive index* due to the presence of an electric field can be detected by:
   - Interferometric methods (where the velocity of light is equal to $c/n$)
   - Change in light intensity in a beam passing through an optical wave guide device like Li-Nb (Figure 1.35).

By means of method 1, many kinds of instruments (voltmeters) can be realized. Methods 2 through 5 are used in research laboratories but are not yet used in industrial measurements.

## Inductive Sensors

### Voltage Transformers (VTs)

Voltage transformers have two different tasks:

- Reduction in voltage values for meeting the range of normal measuring instruments or protection relays
- Insulation of the measuring circuit from power circuits (necessary when voltage values are over 600 V)

Voltage transformers are composed of two windings — one primary and one secondary winding. The primary winding must be connected to power circuits; the secondary to measuring or protection circuits. Electrically, these two windings are insulated but are connected magnetically by the core.

One can define:

$$\textit{Nominal ratio} = K_n = \frac{V_{1\text{n}}}{V_{2\text{n}}} \tag{1.59}$$

as the ratio between the magnitude of primary and secondary rated voltages.

$$\textit{Actual ratio} = K = \frac{V_1}{V_2} \tag{1.60}$$

as the ratio between the magnitudes of primary and secondary actual voltages.

*Burden* is the value of the apparent power (normally at $\cos\varphi = 0.8$) that can be provided on the secondary circuit (instruments plus connecting cables).

Burden limits the maximum value of secondary current and then the minimum value of impedance of the secondary circuit is:

TABLE 1.4 Angle and Ratio Error Limit Table Accepted by CEI-IEC Standards

| Class | Percentage voltage (ratio) error (±) | Phase displacement Minutes (±) | Centiradians (±) |
|---|---|---|---|
| 0.1 | 0.1 | 5 | 0.15 |
| 0.2 | 0.2 | 10 | 0.3 |
| 0.5 | 0.5 | 20 | 0.6 |
| 1 | 1 | 40 | 1.2 |
| 3 | 3 | — | — |
| 3P | 3 | 120 | 3,5 |
| 6P | 6 | 240 | 7 |

$$Z_{\min} = \frac{V_{2n}^2}{A_n} \tag{1.61}$$

where $A_n$ = VT burden

For example, if $A_n$ = 25 VA and $V_{2n}$ = 100 V, one obtains:

$$Z_{\min} = \frac{100}{0.25} = 400\ \mathrm{W} \tag{1.62}$$

There are two kinds of errors:

1. *Ratio error* = $\text{Ratio error} = h_{\%} = \dfrac{K_n - K}{K}$ (1.63)
2. *Angle error* = the phase displacement between the primary voltage and the secondary voltage (positive if the primary voltage lags the secondary one).

Voltage transformers are subdivided into accuracy classes related to the limits in ratio and angle error (according to CEI and IEC normative classes 0.1, 0.2, 0.5, 1, 3; see Table 1.4). To choose the voltage transformer needed, the following technical data must be followed:

- Primary and secondary voltage (rated transformation ratio). Normally, the secondary value is 100 V.
- Accuracy class and rated burden in VA: e.g., cl. 0.5 and $A_n$ = 10 VA.
- Rated working voltage and frequency
- Insulation voltage
- Voltage factor: the ratio between maximum operating voltage permitted and the rated voltage. The standard voltage factor is 1.2 $V_n$ (i.e., the actual primary voltage) for an unlimited period of time (with VT connected with phases), and is 1.9 $V_n$ for a period of 8 h for VT connected between phase and neutral.
- Thermal power is the maximum burden withstood by VT (errors excluded).

For extremely high voltage values, both capacitive dividers and voltage transformers are normally used, as shown in Figure 1.36. The capacitive impedance must compensate for the effect of the transformer's internal inductive impedance at the working frequency.

## Other Methods

The ac voltage inductive sensors act by interaction between a magnetic field (by an electromagnet excited by voltage to be measured) and the eddy current induced in an electroconductive disk, producing a force or a torque. This can be achieved by the scheme shown in Figure 1.37. The weight of many parts of the

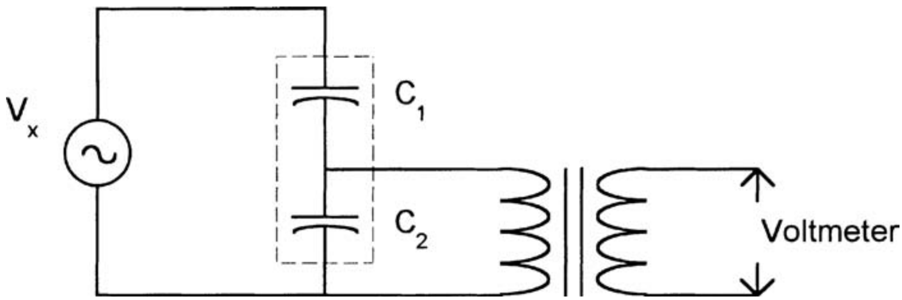



**FIGURE 1.36** Capacitive divider and voltage transformer device for extremely high voltage.

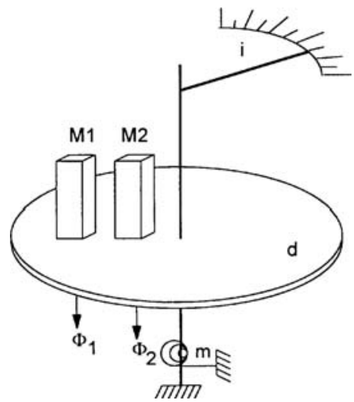



**FIGURE 1.37** Schematic inductive voltmeter. The parts are labeled as: (i) index, (d) metallic disk, (M1) and (M2) electromagnets, (m) spring, (Φ1) and (Φ2) generated fluxes.

indicator can be some tens of grams. The power absorbed is on the order of a few watts. The precision is not high, but it is possible to get these sensors or instruments as they are similar to the widely produced induction energy meters. They are quite robust and are priced between $50 and $100, but they are not widely used. The relation between torque and voltage is quadratic:

$$T = k_i V^2 \tag{1.64}$$

The proportionality factor $k_i$ depends on magnet characteristics and disk geometry.

G.E.C., Landys & Gyr, A.B.B., Schlumberger, etc. are the major companies that furnish components and instruments measuring voltage by inductive and capacitive sensors.

## Defining Terms

**CEI:** Comitato Elettrotecnico Italiano.
**IEC:** International Electric Committee.
**KDP:** Potassium dihydrogen phosphate.
**Li-Nb:** ($LiNbO_3$) lithium niobate.

## Further Information

J. Moeller and G. Rosenberger, Instrument Transformers for HV Systems, Siemens Power Engineering III (1981) Special Issue, *High-Voltage Technology.*

G. Sacerdoti, O. Jappolo, and R. Paggi, *Misure Elettriche, Vol. I Strumenti*, Bologna, Zanichelli, 1994.

G. Zingales, *Metodi e Strumenti per le Misure Elettriche*, Bologna, UTET, 1976.

# References

## 1 1. Voltage Measurement

1. M.B. Stout, Basic Electrical Measurements, Englewood Cliffs, NJ, Prentice-Hall, 1960.

2. I.F. Kinnard, Applied Electrical Measurements, New York, John Wiley & Sons, Chapman & Hall, Ltd. London, 1956.

3. B.M. Oliver and J.M. Cage, Electronic Measurements and Instrumentation, London, McGraw-Hill, Inc. 1975.

4. T.T. Lang, Electronics of Measuring Systems, New York, John Wiley & Sons, 1987.

5. Analog Devices, Analog-Digital Conversion Handbook, Englewood Cliffs, NJ, Prentice-Hall, 1986.

6. A.V. Oppenheim and R.W. Schafer, Digital Signal Processing, Englewood Cliffs, NJ, Prentice-Hall, 1975.

7. A. Ferrero and R. Ottoboni, High-accuracy Fourier analysis based on synchronous sampling techniques. IEEE Trans. Instr. Meas., 41(6), 780-785, 1992.

1.2 Oscilloscope Voltage Measurement

Jerry Murphy

Engineers, scientists, and other technical professionals around the world depend on oscilloscopes as one

of the primary voltage measuring instruments. This is an unusual situation because the oscilloscope is

not the most accurate voltage measuring instrument usually available in the lab. It is the graphical nature

of the oscilloscope that makes it so valued as a measurement instrument — not its measurement accuracy.

The oscilloscope is an instrument that presents a graphical display of its input voltage as a function

of time. It displays voltage waveforms that cannot easily be described by numerical methods. For example,

the output of a battery can be completely described by its output voltage and current. However, the output of a more complex signal source needs additional information such as frequency, duty cycle, peak

to-peak amplitude, overshoot, preshoot, rise time, fall time, and more to be completely described. The

oscilloscope, with its graphical presentation of complex waveforms, is ideally suited to this task. It is

often described as the “screwdriver of the electronic engineer” because the oscilloscope is the most

fundamental tool that technical professionals apply to the problem of trying to understand the details

of the operation of their electronic circuit or device. So, what is an oscilloscope?

The oscilloscope is an electronic instrument that presents a high-fidelity graphical display of the rapidly

changing voltage at its input terminals.

The most frequently used display mode is voltage vs. time. This is not the only display that could be

used, nor is it the display that is best suited for all situations. For example, the oscilloscope could be

called on to produce a display of two changing voltages plotted one against the other, such as a Lissajous

display. To accurately display rapidly changing signals, the oscilloscope is a high bandwidth device. This

means that it must be capable of displaying the high-order harmonics of the signal being applied to its

input terminals in order to correctly display that signal.

The Oscilloscope Block Diagram

The oscilloscope contains four basic circuit blocks: the vertical amplifier, the time base, the trigger, and

the display. This section treats each of these in a high-level overview. Many textbooks exist that cover

the details of the design and construction of each of these blocks in detail [1]. This discussion will cover

these blocks in enough detail so that readers can construct their own mental model of how their operation

affects the application of the oscilloscope for their voltage measurement application. Most readers of this

book have a mental model of the operation of the automatic transmission of an automobile that is

sufficient for its successful operation but not sufficient for the overhaul or redesign of that component.

It is the goal of this section to instill that level of understanding in the operation of the oscilloscope.

Those readers who desire a deeper understanding will get their needs met in later sections.

Of the four basic blocks of the oscilloscope, the most visible of these blocks is the display with its

cathode-ray tube (CRT). This is the component in the oscilloscope that produces the graphical display

of the input voltage and it is the component with which the user has the most contact. Figure 1.21 shows

the input signal is applied to the vertical axis of a cathode ray tube. This is the correct model for an

analog oscilloscope but it is overly simplified in the case of the digital oscilloscope. The important thing

to learn from this diagram is that the input signal will be operated on by the oscilloscope's vertical axis

circuits so that it can be displayed by the CRT. The differences between the analog and digital oscilloscope

are covered in later sections.

FIGURE 1.21 Simplified oscilloscope block diagram that applies to either analog or digital oscilloscopes. In the case

of the digital oscilloscope, the vertical amplifier block will include the ADC and high-speed waveform memory. For the analog scope the vertical block will include delay lines with their associated drivers and a power amplifier to

drive the CRT plates.

The vertical amplifier conditions the input signal so that it can be displayed on the CRT. The vertical

amplifier provides controls of volts per division, position, and coupling, allowing the user to obtain the

desired display. This amplifier must have a high enough bandwidth to ensure that all of the significant

frequency components of the input signal reach the CRT.

The trigger is responsible for starting the display at the same point on the input signal every time the

display is refreshed. It is the stable display of a complex waveform that allows the user of an oscilloscope

to make judgments about that waveform and its implications as to the operation of the device under test.

The final piece of the simplified block diagram is the time base. This circuit block is also known as

the horizontal system in some literature. The time base is the part of the oscilloscope that causes the

input signal to be displayed as a function of time. The circuitry in this block causes the CRT beam to be

deflected from left to right as the input signal is being applied to the vertical deflection section of the

CRT. Controls for time-per-division and position (or delay) allow the user of the oscilloscope to adjust

the display for the most useful display of the input signal. The time-per-division controls of most

oscilloscopes provide a wide range of values, ranging from a few nanoseconds ($10^{-9}$ s) to seconds per

division. To get a feeling for the magnitude of the dynamic range of the oscilloscope’s time base settings, keep in mind that light travels about 1 m in 3 ns.

## The Oscilloscope as a Voltage Measurement Instrument

That the oscilloscope's vertical axis requires a wide bandwidth amplifier and its time base is capable of displaying events that are as short as a few nanoseconds apart, indicates that the oscilloscope can display rapidly changing voltages. Voltmeters, on the other hand, are designed to give their operator a numeric readout of steady-state or slowly changing voltages. Voltmeters are not well suited for displaying voltages that are changing levels very quickly. This can be better understood by examination of the operation of a voltmeter as compared to that of an oscilloscope. The analog voltmeter uses the magnetic field produced by current flowing through a coil to move the pointer against the force of a spring. This nearly linear deflection of the voltmeter pointer is calibrated by applying known standard voltages to its input.

Therefore, if a constant voltage is applied to the coil, the pointer will move to a point where the magnetic force being produced by the current flowing in its coil is balanced by the force of the spring. If the input voltage is slowly changing, the pointer will follow the changing voltage. This mechanical deflection system limits the ability of this measurement device to the measurement of steady-state or very low-frequency changes in the voltage at its input terminals. Higher-frequency voltmeters depend on some type of conversion technique to change higher frequencies to a dc signal that can be applied to the meter's deflection coil. For example, a diode is used to rectify ac voltages to produce a dc voltage that corresponds to the average value of the ac voltage at the input terminals in average responding ac voltmeters.

The digital voltmeter is very much like the analog meter except that the mechanical displacement of

the pointer is replaced with a digital readout of the input signal. In the case of the digital voltmeter, the

input signal is applied to an analog-to-digital converter (ADC) where it is compared to a reference voltage

and digitized. This digital value of the input signal is then displayed in a numerical display. The ADC

techniques applied to voltmeters are designed to produce very accurate displays of the same signals that

were previously measured with analog meters. The value of a digital voltmeter is its improved measure

ment accuracy as compared to that of its analog predecessors.

The oscilloscope will display a horizontal line displaced vertically from its zero-voltage level when a

constant, or dc voltage is applied to its input terminals. The magnitude of this deflection of the oscillo

scope’s beam vertically from the point where it was operating with no input being applied is how the

oscilloscope indicates the magnitude of the dc level at its input terminals. Most oscilloscopes have a

graticule as a part of their display and the scope’s vertical axis is calibrated in volts per division of the

graticule. As one can imagine, this is not a very informative display of a dc level and perhaps a voltmeter

with its numeric readout is better suited for such applications.

There is more to the scope–voltmeter comparison than is obvious from the previous discussion. That

the oscilloscope is based on a wide-bandwidth data-acquisition system is the major difference between these two measurement instruments. The oscilloscope is designed to produce a high fidelity display of

rapidly changing signals. This puts additional constraints on the design of the oscilloscope's vertical

system that are not required in the voltmeter. The most significant of these constraints is that of a constant

group delay. This is a rather complex topic that is usually covered in network analysis texts. It can be

easily understood if one realizes the effect of group delay on a complex input signal.

Figure 1.22 shows such a signal. The amplitude of this signal is a dc level and the rising edge is made

up of a series of high-frequency components. Each of these high-frequency components is a sine wave

of specific amplitude and frequency. Another example of a complex signal is a square wave with a

frequency of 10 MHz. This signal is made up of a series of odd harmonics of that fundamental frequency.

These harmonics are sine waves of frequencies of 10 MHz, 30 MHz, 50 MHz, 70 MHz, etc. So, the

oscilloscope must pass all of these high-frequency components to the display with little or no distortion.

Group delay is the measure of the propagation time of each component through the vertical system. A

constant group delay means that each of these components will take the same amount of time to propagate

through the vertical system to the CRT, independent of their frequencies. If the higher-order harmonics

take more or less time to reach the scope's deflection system than the lower harmonics, the resulting

display will be a distorted representation of the input signal. Group delay (in seconds) is calculated by

taking the first derivative of an amplifier's phase-vs.-frequency response (in radians/(l/s)). If the amplifier

has a linearly increasing phase shift with frequency, the first derivative of its phase response will be a

horizontal line corresponding to the slope of the phase plot (in seconds). Amplifier systems that have a

constant group delay are known as Gaussian amplifiers. They have this name because their pass band

shape resembles that of the bell curve of a Gaussian distribution function (Figure 1.23). One would think

that the oscilloscope’s vertical amplifier should have a flat frequency response, but this is not the case

because such amplifiers have nonconstant group delay [1].

The oscilloscope’s bandwidth specification is based on the frequency where the vertical deflection will

be –3 dB (0.707) of the input signal. This means that if a constant 1 V sine wave is applied to the

oscilloscope’s input, and the signal’s frequency is adjusted to higher and higher frequencies, the oscillo

scope’s bandwidth will be that frequency where its display of the input signal has been reduced to be

0.707 V. Noticable errors in ampitude measurements will start at 20% of the scope’s bandwidth. The

oscilloscope’s error-free display of complex waveforms gives it poor voltage accuracy. For the measure

ment of dc and single frequency signals such as sine waves, other instruments can produce more accurate

measurements.

FIGURE 1.22 A typical complex waveform. This waveform is described by measurements of its amplitude, offset,

risetime, falltime, overshoot, preshoot, and droop.

Conclusion: The voltmeter makes the most accurate measurements of voltages that are dc, slowly changing, or can be converted to a dc analog of their ac content. The oscilloscope is not the most accurate

voltage measurement instrument, but it is well suited to measurements of voltages that are changing very

rapidly as a function of time. Oscilloscopes are the instrument of choice for observing and characterizing

these complex voltages.

Analog or Digital

The world of oscilloscopes is divided into two general categories: analog and digital. The first oscilloscopes

were analog. These products are based on the direct-view vector cathode-ray tube (DVVCRT or CRT for

short). The analog oscilloscope applies the input signal to the vertical deflection plates of the CRT where

it causes the deflection of a beam of high-energy electrons moving toward the phosphor-coated faceplate.

The electron beam generates a lighted spot where it strikes the phosphor. The intensity of the light is

directly related to the density of the electrons hitting a given area of the phosphor. Because this analog

operation is not based on any digitizing techniques, most people have little trouble creating a very accurate

and simple mental model in their minds of its operation.

The analog oscilloscope produces a display of the input signal that is bright and easy to see under

most conditions. It can also contain as many as 16 shades of gray-scale information. This means that an

event that occurs less frequently will appear at a lower intensity in the display than another event that

occurs more frequently. This oscilloscope does not produce a continous display of the input signal. It is

blind during retrace and trigger hold-off times. Because the display depends on the production of visible light from the phosphor being excited by an electron beam, the display must be refreshed frequently.

This makes the analog oscilloscope a low-dead-time display system that can follow rapidly changing

signals. Also, there is little lag time in front panel control settings.

The analog oscilloscope is not without its shortcomings. The strength of the analog oscilloscope is its

CRT, but this is also the source of its weaknesses. The biggest problem with analog scopes is their

dependence on a display that is constantly being refreshed. This means that these scopes do not have any

waveform storage. If the input signal fails to repeat frequently, the display will simply be a flash of light

FIGURE 1.23 The Gaussian frequency response of the oscilloscope's vertical system which is not flat in its pass

band. Amplitude measurements made at frequencies greater than 20% of the scope's bandwidth will be in error.

when the beam sweeps by the phosphor. If the signal's repetition rate falls below 100 Hz, the display will

flicker annoyingly. Figure 1.24 shows a plot of the range of an input signal's repetition frequency range

from a single-shot event to the full bandwidth of a scope vs. the scope's sweep speeds. The result is a

map of the scope's operational area. Figure 1.24 shows that the analog oscilloscope fails to map onto the

full range of possible input signals and sweep speeds.

Another problem of the analog oscilloscope is its inability to display information ahead of its trigger.

This is a problem in applications where the only suitable trigger is at the end of the event of interest.

Another limitation of analog scopes is their timing accuracy. The time base of the analog scope is based on the linearity of a voltage ramp. There are other sources of errors in the analog oscilloscope's horizontal axis, but the sweep nonlinearity is the major contributor. This results in these scopes having a timing accuracy of typically ±3% of their full-scale setting. Therefore, if the time base is set to 100 ns/div, in order to view a 100 ns wide pulse, the full scale will be 1000 ns or 1 ms. The accuracy of this pulse width measurement will be ±30 ns or ±30% of the pulse width!

The digital oscilloscope or digital storage oscilloscope (DSO) differs from its analog counterpart in that the input signal is converted to digital data and therefore it can be managed by an embedded microprocessor. The waveform data can have correction factors applied to remove errors in the scope's acquisition system and can then be stored, measured, and/or displayed. That the input signal is converted from analog to digital and manipulations are performed on it by a microprocessor results in people not having a good mental model of the digital oscilloscope's operation. This would not be a problem except for the fact that the waveform digitizing process is not totally free from errors, and a lack of a correct mental model of the scope's operation on the part of its user can increase the odds of a measurement error. To make matters worse, various manufacturers of these products make conflicting claims, making it easy to propagate incorrect mental models of the digital scope's operation. It is the intention of this presentation to give the information needed to create a mental model of the operation of these devices that will enable the user to perform error-free measurements with ease.

The digital storage oscilloscope offers many advantages over its analog counterpart. The first is accu

racy. The voltage measurement accuracy of the digital oscilloscope is better than that of an analog scope

FIGURE 1.24 The operating range of the analog oscilloscope. This is a plot of input signal repetition rate from the

lower limit of single shot to the full bandwidth of the scope plotted against sweep speed. The shaded area is the area

where the analog oscilloscope will produce a usable display.

because the microprocessor can apply correction factors to the data to correct for errors in the calibration

of the scope's vertical system. The timing accuracy of a digital oscilloscope is an order of magnitude

better than that of an analog scope. The digital scope can store the waveform data for comparison to

other test results or uploading to a computer for analysis or project documentation. The digital oscillo

scope does not depend on the input signal being continuously updated to produce an easily viewable

display. A single-shot event is displayed at the same brightness level as a signal that repeats in time periods

corresponding to the full bandwidth of the scope.

The disadvantages of the digital oscilloscope are its more complex operation, aliasing, and display

performance. The analog-to-digital conversion process [1] is used to convert the input signal into a series

of discrete values, or samples, uniformly spaced in time, which can be stored in memory. Voltage

resolution is determined by the total number of codes that can be produced. A larger number permits a smoother and more accurate reproduction of the input waveform but increases both the cost and difficulty

in achieving a high sample frequency. Most digital oscilloscopes provide 8-bit resolution in their ADC.

As the ADC's sampling speed is increased, the samples will be closer together, resulting in smaller gaps

in the waveform record.

All digital scopes are capable of producing an aliased display. Some models are more prone to this

problem than others, but even the best will alias under the right conditions. An alias is a lower frequency

false reproduction of the input signal resulting from under-sampling, i.e., sampling less than the Nyquist

frequency. The display of the digital scope is based on computer display technology. This results in a

display that is very bright and easy to see, even under conditions where an analog scope would have

difficulty in producing a viewable display. The disadvantage of the digital scope's display is its lower

horizontal resolution. Most of the scopes on the market have a raster scan display with a resolution of

500 lines, less than half the resolution of an analog scope's display. This is not a problem in most

applications. It could become a factor where very complex waveforms, such as those found in TV systems,

are being analyzed. Many digital scopes have display systems that exhibit large dead- or blind-times.

Scopes based on a single CPU will be able to display their waveform data only after the CPU has finished

all of its operations. This can result in a display that is unresponsive to front panel control inputs as well

as not being able to follow changes in the input signal.

Table 1.2 shows that both analog and digital oscilloscopes have relative advantages and disadvantages.

All the major producers of oscilloscopes are pushing the development of digital scopes in an attempt to

overcome their disadvantages. All the major producers of these products believe that the future is digital.

However, a few manufacturers produce scopes that are both analog and digital. These products appear

to have the best of both worlds; however, they have penalties with respect to both cost and complexity

of operation. TABLE 1.2 A Comparison of Analog and Digital Oscilloscopes Analog Oscilloscope Digital Oscilloscope Operation Simple Complex Front panel controls Direct access knobs Knobs and menus Display Real-time vector Digital raster scan Gray scales >16 >4 Horizontal resolution >1000 lines 500 lines Dead-time Short Can be long Aliasing No Yes Voltage accuracy ±3% of full scale ±3% of full scale Timing accuracy ±3% of full scale ±0.01% of full scale Single shot capture None Yes Glitch capture Limited Yes Waveform storage None Yes Pretrigger viewing None Yes Data out to a computer No Yes

One of the driving forces making scope manufacturers believe that the future of the digital oscilloscope

is bright is that modern electronic systems are becoming ever more digital in nature. Digital systems place

additional demands on the oscilloscope that exceed the capabilities of the analog scope. For example, often

in digital electronic systems, there is a need to view fast events that occur at very slow or infrequent rates.

Figure 1.24 shows that these events fail to be viewable on analog scopes. Another common problem with

digital systems is the location of trigger events. Often the only usable trigger is available at the end of the

event being viewed. Analog scopes can only display events that occur after a trigger event. The rapid growth

of digital electronics that occurred in the late 1990s is being attributed to the lowering of the cost of single chip microcontrollers. These devices, which contain a complete microprocessor on one integrated circuit,

are responsible for the “electronics everywhere” phenomenon, where mechanical devices are becoming

electronic as well as those devices that were previously electrical in nature. In 1996, Hewlett Packard

introduced a new class of oscilloscope designed to meet the unique needs of the microcontrollerbased

applications. This new class of oscilloscope is known as the mixed signal oscilloscope or MSO [2].

Voltage Measurements

Voltage measurements are usually based on comparisons of the waveform display to the oscilloscope’s

graticule. Measurements are made by counting the number of graticule lines between the end-points of

the desired measurement and then multiplying that number by the sensitivity setting. This was the only

measurement available to most analog scope users, and it is still used by those performing troubleshooting

with their digital scope as a time-saving step. (Some late-model analog oscilloscopes incorporate cursors

to enhance their measurement ability.) For example, a waveform that is 4.5 divisions high at a vertical

sensitivity of 100 mV/div would be 450 mV high.

Switching the scope’s coupling between ac and dc modes will produce a vertical shift in the waveform’s

position that is a measurement of its dc component. This technique can be applied to either analog or

digital scopes. Simply note the magnitude of the change in waveform position and multiply by the

channel’s sensitivity.

Additional measurements can be performed with an analog oscilloscope but they usually require more skill on the part of the operator. For example, if the operator can determine the location of the top and base of a complex waveform, its amplitude can be measured. Measurements based on percentages can be made using the scope's vernier to scale the waveform so that its top and bottom are 5 divisions apart. Then, each division represents 20% of the amplitude of the waveform being studied. The use of the vernier, which results in the channel being uncalibrated, prevents performance of voltage measurements. Many analog scopes have a red light to warn the operator that the scope is uncalibrated when in vernier mode.

The digital oscilloscope contains an embedded microprocessor that automates the measurement. This measurement automation is based on a histogramming technique, where a histogram of all the voltages levels in the waveform are taken from the oscilloscope's waveform data. The histogram is a plot of the voltage levels in the waveform plotted against the number of samples found at each voltage level. Figure 1.25 shows the histogramming technique being applied to the voltage measurements of complex waveforms.

## Understanding the Specifications

The oscilloscope's vertical accuracy is one place that a person's mental model of the scope's operation can lead to measurement trouble. For example, the oscilloscope's vertical axis has a frequency response that is not flat across its pass band. However, as noted above, the scope has a Gaussian frequency response to produce the most accurate picture of complex signals. This means that the oscilloscope's accuracy specification of ±3% is a dc-only specification. If one were to attempt to measure the amplitude of a

signal whose frequency is equal to the bandwidth of the scope, one would have to add another 29.3% to

the error term, for a total error of ±32.3%. This is true for both analog and digital oscilloscopes. This

limitation can be overcome by carefully measuring the frequency response of the oscilloscope's vertical

channels. One will need to repeat this process every time the scope is serviced or calibrated, because the

various high-frequency adjustments that may need to be made in the scope's vertical axis will affect the

scope's frequency response. One is probably asking, why don't the manufacturers do this for me? The

answer is twofold. The first is cost, and the second is that this is not the primary application of an

oscilloscope. There are other instruments that are much better suited to the measurement of high

frequency signals. The spectrum analyzer would be this author's first choice.

Additionally, the vertical accuracy is a full-scale specification. This means that at 1 V/div, the full-scale

value is typically 8 V. The measurement error for a scope with a ±3% specification under these conditions

will be ±0.24 V. If the signal being measured is only 1 V high, the resulting measurement will be ±24%

of reading. Check the manual for the scope being used, as some manufacturers will specify full-scale as

being 10 or even 10.2 divisions. This will increase the error term because the full-scale term is larger.

In digital oscilloscopes, the vertical accuracy is often expressed as a series of terms. These attempt to

describe the analog and digital operations the scope performs on the input signal. Terms might include digitizing resolution, gain, and offset (sometimes called as position). They also might be called out as

single and dual cursor accuracies. The single cursor accuracy is a sum of all three terms. In the dual

cursor case, where the voltage measurement is made between two voltage cursors, the offset term will

cancel out, leaving only the digitizing resolution and gain errors. For example, the Hewlett Packard model

54603B has a single cursor accuracy specification of ±1.2% of full scale, ±0.5% of position value, and a

dual cursor specification of ±0.4% of full scale.

HINT: Always try to make the voltage measurements on the largest possible vertical and widest possible

display of the signal.

The horizontal accuracy specifications of analog and digital scopes are very different; however, both

are based on a full-scale value. In the analog scope, many manufacturers limit accuracy specifications to

only the center eight divisions of their display. This means that a measurement of a signal that starts or

ends in either the first or ninth graticule, will be even more error prone than stated in the scope's

specifications. To the best of this author's knowledge, this limitation does not apply to digital scopes.

The horizontal specifications of digital scopes are expressed as a series of terms. These might include the

FIGURE 1.25 Voltage histograms as applied by a digital oscilloscope. The complex waveform is measured by use

of the voltage histogram. This histogram is a plot of each voltage level in the display and the number of data points

at that level.

crystal accuracy, horizontal display resolution, and trigger placement resolution. These can be listed as

cursor accuracy. For example, the Hewlett Packard model 54603B has a horizontal cursor accuracy

specification of ±0.01% ±0.2% full-scale ±200 ps. In this example, the first term is the crystal accuracy,

the second is the display resolution (500 lines), and the final term is twice the trigger placement error.

By comparing the analog and digital scopes' horizontal specifications, it can be seen that in either case,

the measurement is more accurate if it can be made at full screen. The digital scope is more accurate

than its analog counterpart.

Digital scopes also have acquisition system specifications. Here is another place where the operator's

mental model of the operation of a digital scope can produce measurement errors. All manufacturers of

digital scopes specify the maximum sampling speed of their scope's acquisition system as well as its

memory depth and number of bits. The scope's maximum sampling speed does not apply to all sweep

speeds, only memory depth and number of bits applies to all sweep speeds. The scope's maximum

sampling speed applies only to its fastest sweep speeds.

The complexity of the digital scope results from the problem of having to sample the input. There is

more to be considered than Nyquist's Sampling Theorem in the operation of a digital scope. For example,

how does the scope's maximum sampling rate relate to the smallest time interval that the scope can

capture and display? A scope that samples at 100 MSa $s^{-1}$ takes a sample every 10 ns; therefore, in principle,

it cannot display any event that is less than 10 ns wide because that event will fall between the samples.

In practice, however, this limit can — under certain circumstances — be extended. If the scope is operating in an “equivalent time” or “random repetitive” mode and if the signal is repetitive, even if very infrequently, the scope will be able to capture any event that is within its vertical system bandwidth. Figure 1.26 shows an infrequently occurring pulse that is 25 ns wide embedded into a data stream being captured and displayed on an oscilloscope with a maximum sampling speed of 20 $MSa\ s^{-1}$ (sampling interval of 50 ns). Figure 1.26(b) shows this pulse at a faster sweep speed. An analog scope would produce a similar display of this event, with the infrequent event being displayed at a lower intensity than the rest of the trace. Notice that the infrequent event does not break the baseline of the trace.

The correct mental model of the digital scope’s ability to capture signals needs to be based on the scope’s bandwidth, operating modes, and timing resolution. It is the timing resolution that tells the operator how closely spaced the samples can be in the scope’s data record.

The most common flaw in many mental models of the operation of a digital scope is related to its maximum sampling speed specification. As noted, the maximum sampling speed specification applies only to the scope’s fastest sweep speeds. Some scope manufacturers will use a multiplex A/D system that operates at its maximum sampling speed only in single-channel mode. The scope’s memory depth determines its sampling speed at the sweep speed being used for any specific measurement. The scope’s memory depth is always equal to the scope’s horizontal full-scale setting. For scopes with no off-screen

memory, this is 10¢ the time base setting. If the scope has off-screen memory, this must be taken into

account. For example, assume that one has two scopes with a maximum sampling speed of 100 MSa $s^{-1}$.

One scope has a memory depth of 5 K points and the other only 1 K. At a sweep speed of 1 ms per

division, both scopes will be able to store data into their memory at their full sampling speed, and each

will be storing 100 data points per division, for a total of 1000 data points being stored. The scope with

the 5 K memory will have a data point in one of every 5 memory locations, and the scope with the 1 K

memory will have a data point in every memory location. If one reduces the sweep speed to 5 ms/div,

the deeper memory scope will now fill every one of its memory locations with data points separated by

10 ns. The scope with only 1 K of memory would produce a display only 2 divisions wide if its sampling

speed is not reduced. Scope designers believe that scope users expect to see a full-length sweep at every

sweep speed. Therefore, the 1 K scope must reduce its sampling speed to one sample every 50 ns, or

20 MSa $s^{-1}$, to be able to fill its memory with a full sweep width of data. This 5:1 ratio of sampling speeds

between these two scopes will be maintained as their time bases are set to longer and longer sweeps. For

example, at 1 s/div, the 5 K scope will be sampling at 500 samples per second, while the 1 K scope will

be sampling at only 100 samples per second. One can determine a scope's sampling speed for any specific

time base setting from Equation 1.51. (1.51)

One must look closely at the application to determine if a specific scope is best suited to that application.

As a rule, the deeper the memory, the faster the scope will be able to sample the signal at any given time

base setting. Memory depth is not free. High-speed memory required to be able to store the data out of

the scope's A/D is costly, and deeper memory takes longer to fill, thus reducing the scope's display update

rate. Most scopes that provide memory depths of 20 K or more will also give the user a memory depth

selection control so that the user can select between fast and deep. (In 1996, Hewlett Packard Co.

FIGURE 1.26 An infrequently occurring event as displayed on a digital oscilloscope with random repetitive sam

pling. (a) The event embedded in a pulse train. (b) Shows the same event at a faster sweep speed. The fact that the

waveform baseline is unbroken under the narrow pulse indicates that it does not occur in every sweep. The width

of this pulse is less than half the scope's sampling period in (b). Both traces are from a Hewlett Packard model 54603B

dual channel 60 MHz scope. S samples second memory depth samples full-scale time base seconds or the scope's maximum sampling speed, whichever is less ( ) = ( ) ( ) ,

introduced two scopes based on an acquisition technology known as MegaZoom (TM) [10] that removes

the need for a memory depth control.) A correct mental model for the sampling speed of a digital scope

is based on Equation 1.51 and not just on the scope's maximum performance specifications.

Some digital oscilloscopes offer a special sampling mode known as peak detection. Peak detection is a

special mode that has the effect of extending the scope's sampling speed to longer time records. This

special mode can reduce the possibility of an aliased display. The performance of this special mode is specified as the minimum pulse width that the peak detection system can capture. There are several peak detection systems being used by the various manufacturers. Tektronix has an analog-based peak detection system in some of its models, while Hewlett Packard has a digital system in all of its models. Both systems perform as advertised, and they should be evaluated in the lab to see which system best meets one's needs. There is a downside to peak detection systems and that is that they display high-frequency noise that might not be within the bandwidth of the system under test. Figure 1.27 shows a narrow pulse being captured by peak detection and being missed when the peak detection is off.

What effect does display dead-time have on the oscilloscope's voltage measurement capabilities? Display dead-time applies to both analog and digital oscilloscopes, and it is that time when the oscilloscope is not capturing the input signal. This is also a very important consideration in the operation of a digital scope because it determines the scope's ability to respond to front-panel control commands and to follow changing waveforms. A digital scope that produces an incorrect display of an amplitude-modulated signal is not following this rapidly changing signal because its display update rate is too low. Sampling speed is not related to display update rate or dead-time. Display dead-time is a function of the scope's ability to process the waveform data from its A/D and plot it on the display. Every major oscilloscope manufacturer has been working on this problem. Tektronix offers a special mode on some of its products known as InstaVu (TM) [4]. This special mode allows these scopes to process up to 400,000 waveforms

per second to their display. Hewlett Packard has developed a multiple parallel processor technology [5]

in the HP 54600 series of benchtop scopes that provides a high-speed, low dead-time display in a lowcost

instrument. These instruments can plot 1,500,000 points per second to their display and they have no

dead-time at their slower sweep speeds. LeCroy has been applying the Power PC as an embedded processor

for its scopes to increase display throughput. There are other special modes being produced by other

vendors, so be sure to understand what these can do before selecting an oscilloscope. Figure 1.28 shows

the effect of display update rate on a rapidly changing waveform. An amplitude-modulated signal is

displayed with a high-speed display and with the display speed reduced by the use of hold-off.

## Triggering

The trigger of the oscilloscope has no direct effect on the scope's ability to measure a voltage except that

the trigger does enable the oscilloscope to produce a stable display of the voltage of interest. Ref. [6]

presents a thorough discussion of this subject.

## Conclusion

The mental model that oscilloscope users have created in their minds of the oscilloscope's operation can

be helpful in reducing measurement errors. If the operator's mental model is based on the following

facts, measurement errors can be minimized:

- Oscilloscopes have a frequency response that affects measurement accuracy.
- Digital scopes are more accurate than analog scopes.
- Analog scopes do not have continuous displays.
- Oscilloscope accuracy specifications always contain a percent of full-scale term.
- Measurements should be made

at the largest possible deflection in order to minimize errors. • Maximum sampling speed is available only at the scope's fastest sweep speeds. • Deeper memory depth allows faster sampling at more sweep speeds. • All digital scopes can produce aliases, some more than others. • Display dead-time is an important characteristic of digital scopes that is often not specified. • Display dead-time affects measurement accuracy because it can cause a distorted display. • The scope with the highest maximum sampling speed specification might not be the most accurate or have the lowest display dead-time. • The operator must have some knowledge of the signals being measured to be able to make the best possible measurements.

The person who has the mental model of the oscilloscope that takes these factors into account will be

able to purchase the scope that is best suited to his/her application and not spend too much money on

unnecessary performance. In addition, that person will be able to make measurements that are up to the

full accuracy capabilities of the scope.

FIGURE 1.27 Peak detection. This special mode has the effect of increasing the scope's sampling speed at time base

settings where it would be decimated. In operation, each memory location contains either the maximum or minimum

value of the waveform at that location in time. (a) A series of 300 ns wide pulses being captured at a slow sweep

speed; (b) the same setup with peak detection disabled. These narrow pulses would appear as intermittent pulses if

the scope could be seen in operation with peak detection disabled.

Selecting the Oscilloscope

There are ten points to consider when selecting an oscilloscope. This author has published a thorough

discussion of these points [7] and they are summarized as follows:

1. Analog or Digital? There are a few places where the

analog scope might be the best choice, and the reader can make an informed selection based on the information presented here.

2. How much bandwidth? This is a place where the person selecting an oscilloscope can save money by not purchasing more bandwidth than is needed. When analog oscilloscopes were the only choice, many people were forced to purchase more bandwidth than they needed because they needed to view infrequent or low repetition signals. High-bandwidth analog scopes had brighter CRTs so that they were able to display high-frequency signals at very fast time base settings. At a sweep speed of 5 ns/div, the phosphor is being energized by the electron beam for 50 ns, so the

FIGURE 1.28 Display dead-time. The time that an oscilloscope is blind to the input signal has an effect on the

scope’s ability to correctly display rapidly changing signals. (a) An amplitude-modulated signal with a high-speed

display; (b) the same signal with the dead-time increased by use of hold-off. electron beam had to be very high energy to produce a visible trace. This situation does not apply to digital scopes. Now, one needs to be concerned only with the bandwidth required to make the measurement. Figure 1.29 shows the effect of oscilloscope bandwidth on the display of a 50 MHz square wave.

The oscilloscope’s bandwidth should be >2¥ the fundamental highest frequency signal to be measured. The bandwidth of the scope’s vertical system can affect the scope’s ability to correctly display narrow pulses and to make time interval measurements. Because of the scope’s Gaussian frequency response, one can determine its ability to correctly display a transient event in terms of risetime with Equation 1.52. (1.52)

FIGURE 1.29 The effect of the scope’s bandwidth is shown in this set of waveforms. The same 50 MHz square wave

is shown as it was displayed on scopes of 500 MHz in Figure 1.28(a) all the way down to 20 MHz in Figure 1.29(e).

Notice that the 100 MHz scope produced a usable display although it was missing the high-frequency details of the

500 MHz display. The reason that the 100 MHz scope looks so

good is the fact that its bandwidth is slightly greater

than 100 MHz. This performance, which is not specified on any data sheet, is something to look for in any evaluation. t r BW= 0 35. Therefore, a 100 MHz scope will have a risetime of 3.5 ns. This means that if the scope were to have a signal at its input with zero risetime edges, it would be displayed with 3.5 ns edges. This will affect the scope's measurements in two ways. First is narrow pulses. Figure 1.30 shows the same 5 ns wide pulse being displayed on oscilloscopes of 500 MHz and 60 MHz bandwidths, and the effect of the lower bandwidth on this event that is closest to the risetime of the slower scope is apparent. The second is fast time interval measurements. A measurement of signal risetime is an example. The observed risetime on the scope's display is according to Equation 1.53. (1.53) If a 10 ns risetime were to be measured with a 100 MHz scope, one would obtain a measurement of 10.6 ns based on Equation 1.53. The scope would have made this measurement with a 6% reading error before any other factors, such as time base accuracy, are considered.

FIGURE 1.29 (continued) t t t observed signal scope = + ( ) 2 2 1 2

The scope's risetime should be at least no more than 1/5 of the shortest time interval to be measured.

For time interval measurements, this should be >1/10.

3. How many channels? Most oscilloscopes in use today are dual-channel models. In addition, there are models described as being 2+2 and four channels. This is one time where 2+2 is not equal to 4. The 2+2 models have limited features on two of their channels and cost less than 4-channel models. Most oscilloscope suppliers will hold the 4-channel description only for models with four fullfeatured channels, but the user should check the model under consideration so as to be sure if it is a 4- or 2+2 model. Either of the four channel classes is useful for applications involving the testing and development of digital-based systems where the relationship of several signals must be observed. Hewlett Packard introduced a new class of oscilloscopes that is tailored for the applications involving both analog and digital technologies, or mixed-signal systems. The mixed signal oscilloscope (MSO) [4] provides 2 scope channels and 16 logic channels so that it can display both the analog and digital operation of a mixed-signal system on its display.

4. What sampling speed? Do not simply pick the scope with the highest banner specification. One needs to ask, what is the sampling speed at the sweep speeds that my application is most likely to require? As observed in Equation 1.51 the scope's sampling speed is a function of memory depth and full-scale time base setting. If waveforms are mostly repetitive, one can save a lot of money by selecting an oscilloscope that provides equivalent time or random repetitive sampling.

5. How much memory? As previously discussed, memory depth and sampling speed are related. The memory depth required depends on the time span needed to measure and the time resolution required. The longer the time span to be captured and the finer the resolution required, the more memory one will need. High-speed waveform memory is expensive. It takes time to process a longer memory, so the display will have more dead-time in a long memory scope than a shallow memory model. All the suppliers of deep memory scopes provide a memory depth control. They provide this control so that the user can choose between a high-speed display and deep memory for the application at hand. Hewlett Packard introduced MegaZoom (TM) technology [3] in 1996; it produces a high-speed low dead-time display with deep memory all the time.

6. Triggering? All scope manufacturers are adding new triggering features to their products. These features are important because they allow for triggering on very specific events. This can be a valuable troubleshooting tool because it will let the user prove whether a suspected condition

FIGURE 1.29 (continued) exists or not. Extra triggering features add complexity to the scope's user interface; so be sure to try them out to make sure that they can be applied.

7. Trustworthy display? Three factors critically affect a scope's ability to display the unknown and complex signals that are encountered in oscilloscope applications. If the user loses confidence in the scope's ability to correctly display what is going on at its probe tip, productivity will take a real hit. These are display update rate, dead-time, and aliasing. Because all digital scopes operate on sampled data, they are subject to aliasing. An alias is a false reconstruction of the signal caused by under-sampling the original. An alias will always be displayed as a lower frequency than the actual signal. Some vendors employ proprietary techniques to minimize the

likelihood of this problem occurring. Be sure to test any scope being considered for purchase on your worse-case signal to see if it produces a correct or aliased display. Do not simply test it with a single-shot signal that will be captured at the scope's fastest sweep speed because this will fail to test the scope's ability to correctly display signals that require slower sweep speeds.

FIGURE 1.30 Bandwidth and narrow events. (a) A 5 ns wide pulse as displayed on a 500 MHz scope; (b) the same

pulse displayed on a 60 MHz scope. The 60 MHz scope has a risetime of 5.8 ns, which is longer than the pulse width.

This results in the pulse shape being incorrectly displayed and its amplitude being in error.

8. Analysis functions? Digital oscilloscopes with their embedded microprocessors have the ability to perform mathematical operations that can give additional insight into waveforms. These operations often include addition, subtraction, multiplication, integration, and differentiation. An FFT can be a powerful tool, but do not be misled into thinking it is a replacement for a spectrum analyzer. Be sure to check the implementation of these features in any scope being considered. For example, does the FFT provide a selection of window functions? Are these analysis functions implemented with a control system that only their designer could apply?

9. Computer I/O? Most of the digital scopes on the market today can be interfaced to a PC. Most of the scope manufacturers also provide some software that simplifies the task of making the scope and PC work together. Trace images can be incorporated into documents as either PCX or TIF files. Waveform data can be transferred to spreadsheet applications for additional analysis. Some scope models are supplied with a disk drive that can store either waveform data or trace images.

10. Try it out? Now one has the information to narrow oscilloscope selection to a few models based on bandwidth, sampling speed, memory depth, and budget requirements. Contact the scope vendors (Table 1.3) and ask for an evaluation unit. While the evaluation unit is in the lab, look for the following characteristics:

TABLE 1.3 Major Suppliers of Oscilloscopes and their Web Addresses

Vendor Description Web address

B&K Precision

6460 W. Cortland St.

Chicago, IL 60635 Analog and digital scopes and Metrix scopes in France http://bkprecision.com

Boonton Electronics Corp.

25 Estmans Road

P.O. Box 465

Parsippany, NJ 07054-0465 U.S. importer for Metrix analog, mixed analog, and digital scopes from France http://www.boonton.com

Fluke

P.O. Box 9090

Everett, WA 98206-9090 Hand-held, battery-powered scopes (ScopeMeter), analog scopes, and CombiScopes(R) http://www.fluke.com

Gould

Roebuck Road, Hainault,

Ilford, Essex IG6 3UE, England 200 MHz DSO products http://www.gould.co.uk

Hewlett Packard Co.

Test & Measurement

Mail Stop 51LSJ

P.O. Box 58199

Santa Clara, CA 95052-9952 A broad line of oscilloscopes and the Mixed Signal oscilloscope for technical professionals http://www.tmo.hp.com/tmo/pia search on “oscilloscopes”

LeCroy Corp.

700 Chestnut Ridge Road

Chestnut Ridge, NY 10977 Deep memory oscilloscopes for the lab http://www.lecroy.com

Tektronix Inc.

Corporate Offices

26600 SW Parkway

P.O. Box 1000

Watsonville, OR 97070-1000 The broad line oscilloscope supplier with products ranging from hand-held to highperformance lab scopes http://www.tek.com/measurement search on “oscilloscopes”

Yokogawa Corp. of America

Corporate offices

Newnan, GA

1-800-258-2552 Digital oscilloscopes for the lab http://www.yca.com • Control panel responsiveness: Does the scope respond quickly to inputs or does it have to think about it for a while? • Control panel layout: Are the various functions clearly labeled? Does the user have to refer to the manual even for simple things? • Display speed: Turn on a couple of automatic measurements and check that the display speed remains fast enough to follow changing signals. • Aliasing: Does the scope produce an alias when the time base is reduced from fast to slow sweep speeds? How does the display look for the toughest signal?

The oscilloscope is undergoing a period of rapid change. The major manufacturers of oscilloscopes

are no longer producing analog models and the digital models are evolving rapidly. There is confusion

in the oscilloscope marketplace because of the rapid pace of this change. Hopefully, this discussion will

prove valuable to the user in selecting and applying oscilloscopes in the lab in the years to come.

1. A. DeVibiss, Oscilloscopes, in C.F. Coombs, Jr. (ed.), Electronic Instrument Handbook, 2nd ed., New York, McGraw-Hill, 1995.

2. R. A. Witte, A family of instruments for testing mixed-signal circuits and systems, Hewlett Packard J., April 1996, Hewlett Packard Co., Palo Alto, CA.

3. M.S. Holcomb, S.O. Hall, W.S. Tustin, P.J. Burkart, and S.D. Roach, Design of a mixed signal oscilloscope, Hewlett Packard J., April 1996, Hewlett Packard Co., Palo Alto, CA.

4. InstaVu acquisition mode, Tektronix Measurement Products Catalog, Tektronix Inc., Beaverton, OR, 1996, 69.

5. M.S. Holcomb and D.P. Timm, A high-throughput acquisition architecture for a 100 MHz digitizing oscilloscope, Hewlett Packard J., February 1992, Hewlett Packard Co., Palo Alto,CA.

6. R.A. Witte, Elecronic Test Instruments, Theory and Applications, Englewood Cliffs, NJ, PrenticeHall, 1993.

7. J. Murphy, Ten points to ponder in picking an oscilloscope, IEEE Spectrum, 33(7), 69-77, 1996.

## 1.3 Inductive and Capacitive Voltage Measurement

Cipriano Bartoletti, Luca Podestã, and Giancarlo Sacerdoti

This chapter section addresses electrical measurements where the voltage range to be measured is very

large — from $10^{-10}$ V to $10^{7}$ V. The waveform can be continuous, periodic, or impulsive. If it is periodic,

the spectrum components can vary for different situations, and within the same electric power network

there may be subharmonic components. In impulsive voltage measurement, it is often important to get

maximum value, pulse length, etc. Capacitive and inductive voltage sensors are mainly utilized in low

frequency electric measurements.

### Capacitive Sensors

The voltage to be measured can be reduced by means of capacitive dividers (Figure 1.31). Capacitive

dividers are affected by temperature and frequency and therefore are not important, at least in Europe.

Capacitive sensors detect voltage by different methods:

1. Electrostatic force (or torque)

2. Kerr or Pockels effect

3. Josephson effect

4. Transparency through a liquid crystal device

5. Change in refractive index of the optic fiber or in light pipe

1. The relations that rule the listed capacitive voltage sensors are reported below. The force between

two electrodes is (Figure 1.32): (1.54)

where e 0 = Dielectric constant S = Area of the electrode d = Distance V 1 , V 2 = Potentials of the electrodes

The torque between electrostatic voltmeter quadrants (Figure 1.33) is given by: (1.55)

where C = Capacitance q = Angle between electrodes

To get the torque from the rate of change (derivative) of electrostatic energy vs. the angle is easy. Obtaining

the torque by mapping the electric field is difficult and requires long and complex field computing.

2. The rotation of the polarization plane of a light beam passing through a KDP crystal under the

influence of an electric field (Pockels effect) is expressed by (Figure 1.34): (1.56)

where k p = Electro-optic constant l = Length of crystal

FIGURE 1.31 Schematic arrangement of a capacitive divider.

FIGURE 1.32 Force between two electrodes with an applied voltage. F S d V V= ( ) e 0 1 2 2 T C V V= ∂ ∂ ( ) 1 2 1 2 2 q q p = ( ) k l V V 1 2

One obtains a rotation of p/2 by applying a voltage of the order of 1 kV to a KDP crystal of a few centimeters in length.

If a light beam passes through a light pipe that performs the Kerr effect, one observes a quadratic

dependence of the rotation vs. V. (1.57)

3. The Josephson effect consists of translation of a voltage into a periodical signal of a certain frequency,

carried out by a special capacitive sensor. There is an array of N layers of Josephson superconducting

junctions; the frequency of emitted signal, when a voltage V is applied, is given by: (1.58)

4. The transparency of a liquid crystal device depends on the difference of potential applied. There are

liquid crystal devices working in transmission or in reflection. A change in transparency is obtained when

a difference of potential of a few volts is applied.

FIGURE 1.33 Scheme of an electrostatic voltmeter. (a) Lateral view; (b) top view: (1), (2), (3), (4) are the static

electrodes; the moving vane is shown in transparency.

FIGURE 1.34 Scheme of an electrooptic KDP device. The parts are labeled as: (B) a light beam, (P) a polarizer,

(A) an analyzer, (K) a KDP crystal, with the voltage to be measured Vx applied to its (E) transparent electrodes. q ∫ ∫ Φ kE k V 2 2 v eV Nh = 2

5. The change in refractive index due to the presence of an electric field can be detected by: • Interferometric methods (where the velocity of light is equal to c/n) • Change in light intensity in a beam passing through an optical wave guide device like Li-Nb (Figure 1.35).

By means of method 1, many kinds of instruments (voltmeters) can be realized. Methods 2 through 5

are used in research laboratories but are not yet used in industrial measurements.

Voltage Transformers (VTs)

Voltage transformers have two different tasks: • Reduction in voltage values for meeting the range of normal measuring instruments or protection relays • Insulation of the measuring circuit from power circuits (necessary when voltage values are over 600 V)

Voltage transformers are composed of two windings — one primary and one secondary winding. The

primary winding must be connected to power circuits; the secondary to measuring or protection circuits.

Electrically, these two windings are insulated but are connected magnetically by the core.

One can define: (1.59)

as the ratio between the magnitude of primary and secondary rated voltages. (1.60)

as the ratio between the magnitudes of primary and secondary actual voltages.

Burden is the value of the apparent power (normally at cosj = 0.8) that can be provided on the

secondary circuit (instruments plus connecting cables).

Burden limits the maximum value of secondary current and then the minimum value of impedance

of the secondary circuit is:

FIGURE 1.35 Li-Nb optical wave guide device. Nominal ratio = =K V V n 1 2 n n Actual ratio = =K V V 1 2 (1.61)

where A n = VT burden

For example, if A n = 25 VA and V 2n = 100 V, one obtains: (1.62)

There are two kinds of errors:

1. Ratio error = (1.63)

2. Angle error = the phase displacement between the primary

voltage and the secondary voltage (positive if the primary voltage lags the secondary one).

Voltage transformers are subdivided into accuracy classes related to the limits in ratio and angle error

(according to CEI and IEC normative classes 0.1, 0.2, 0.5, 1, 3; see Table 1.4). To choose the voltage

transformer needed, the following technical data must be followed: • Primary and secondary voltage (rated transformation ratio). Normally, the secondary value is 100 V. • Accuracy class and rated burden in VA: e.g., cl. 0.5 and A n = 10 VA. • Rated working voltage and frequency • Insulation voltage • Voltage factor: the ratio between maximum operating voltage permitted and the rated voltage. The standard voltage factor is 1.2 V n (i.e., the actual primary voltage) for an unlimited period of time (with VT connected with phases), and is 1.9 V n for a period of 8 h for VT connected between phase and neutral. • Thermal power is the maximum burden withstood by VT (errors excluded).

For extremely high voltage values, both capacitive dividers and voltage transformers are normally used,

as shown in Figure 1.36. The capacitive impedance must compensate for the effect of the transformer’s

internal inductive impedance at the working frequency.

Other Methods

The ac voltage inductive sensors act by interaction between a magnetic field (by an electromagnet excited

by voltage to be measured) and the eddy current induced in an electroconductive disk, producing a force

or a torque. This can be achieved by the scheme shown in Figure 1.37. The weight of many parts of the TABLE 1.4 Angle and Ratio Error Limit Table Accepted by CEI-IEC Standards Percentage voltage (ratio) Phase displacement Class error (±) Minutes (±) Centiradians (±) 0.1 0.1 5 0.15 0.2 0.2 10 0.3 0.5 0.5 20 0.6 1 1 40 1.2 3 3 — — 3P 3 120 3,5 6P 6 240 7 Z V A min = 2 2 n n Z min . = = 100 0 25 400 W Ratio error = = h K K K % n

indicator can be some tens of grams. The power absorbed is on the order of a few watts. The precision is not high, but it is possible to get these sensors or instruments as they are similar to the widely produced

induction energy meters. They are quite robust and are priced between $50 and $100, but they are not

widely used. The relation between torque and voltage is quadratic: (1.64)

The proportionality factor $k_i$ depends on magnet characteristics and disk geometry.

G.E.C., Landys & Gyr, A.B.B., Schlumberger, etc. are the major companies that furnish components

and instruments measuring voltage by inductive and capacitive sensors.

## Defining Terms

CEI: Comitato Elettrotecnico Italiano.

IEC: International Electric Committee.

KDP: Potassium dihydrogen phosphate.

Li-Nb: ($LiNbO_3$) lithium niobate.

## Further Information

J. Moeller and G. Rosenberger, Instrument Transformers for HV Systems, Siemens Power Engineering III (1981) Special Issue, High-Voltage Technology.

G. Sacerdoti, O. Jappolo, and R. Paggi, Misure Elettriche, Vol. I Strumenti, Bologna, Zanichelli, 1994.

G. Zingales, Metodi e Strumenti per le Misure Elettriche, Bologna, UTET, 1976.

FIGURE 1.36 Capacitive divider and voltage transformer device for extremely high voltage.

FIGURE 1.37 Schematic inductive voltmeter. The parts are labeled as: (i) index, (d) metallic disk, (M1) and (M2)

electromagnets, (m) spring, (F1) and (F2) generated fluxes.

$$T = k_i V^2$$

## 2 2. Current Measurement


1. I. Genzer and P. Youngner, Physics, Chapter 10-6, Morristown, NJ: General Learning Corporation, 1973. A high school physics text.

2. H.E. White, Modern College Physics, Chapter 51, New York: D. Van Nostrand, 1952.

3. B.N. Taylor, Guide for the Use of the International System of Units (SI), NIST Special Publication 811, Washington, D.C., U.S. Government Printing Office, 1995, Appendix A.5.

4. A. Sommerfeld, Electrodynamics, Lectures on Theoretical Physics, Vol. III, New York: Academic Press, 1952.

5. R.C. Dorf (Ed.), The Electrical Engineering Handbook, Chapter 49, The Hall Effect, Boca Raton, FL, CRC Press, 1993.

6. K.B. Rochford, A.H. Rose, M.N. Deeter, and G.W. Day, Faraday effect current sensor with improved sensitivity-bandwidth product, Optics Lett., 19, 1903-1905, 1994.

7. S.R. Forrest (Ed.), JTEC Panel Report on Optoelectronics in Japan and the United States, Baltimore, MD: International Technology Research Institute, Loyola College, February 1996. NTIS PB96152202. Available http://itri.loyola.edu/opto/c6_s3.htm.

8. M.N. Deeter, Domain effects in Faraday effect sensors based on iron garnets, Appl. Opt., 34, 655, 1995.

9. M.N. Deeter, A.H. Rose, and G.W. Day, Faraday effect magnetic field sensors based on substituted iron garnets, in Fiber Optic and Laser Sensors VIII, R.P. DePaula and E. Udd, eds., Proc. Soc. PhotoOpt. Instrumentation. Eng. 1367, 243-248, 1990.

10. S. Cantor, NASA Tech Briefs, November 1993, p. 54.

## 3 3. Power Measurement


1. G. Zingales, Power measurements on single-phase ac circuits (Chap. VI, 6.2), in Methods and Instruments for Electrical Measurements, (in Italian), Torino: UTET, 1980.

2. G. Korànyi, Measurement of power and energy, in L. Schnell (Ed.), Technology of Electrical Measurements, Chichester: John Wiley & Sons, 1993.


FIGURE 3.39 Block diagram of an instrument measuring average power per duty cycle.

FIGURE 3.40 Block diagram of an instrument based on dc/pulse power comparison technique.


3. J. Milmann and C.C. Halkias, Integrated Electronics: Analog and Digital Circuits and Systems, New York: McGraw-Hill, 1972.

4. F.F. Mazda, Ac analogue instruments (Chap. VI, 6.3), in Electronic Instruments and Measurement Techniques, Cambridge, U.K.: Cambridge University Press, 1987.

5. P.S. Filipski, A TDM wattmeter with 0.5 MHz carrier frequency, IEEE Trans. Instrum. Meas., IM39, 15-18, 1990.

6. J.R. Carstens, Electrical Sensors and Transducers, Englewood Cliffs, NJ: Prentice-Hall, 1992.

7. Lu Zu-Liang, An error estimate for quasi-integer-period sampling and an approach for improving its accuracy, IEEE Trans. Instrum. Meas., IM-23, 337-341, 1984.

8. V. Haasz, The error analysis of digital measurements of electrical power, Measurement, 6, 1986.

9. P. Arpaia, F. Avallone, A. Baccigalupi, and C. De Capua, Real-time algorithms for active power measurements on PWM-based electric drives, IEEE Trans. Instrum. Meas., IM-45, 462-466, 1996.

10. X. Dai and R. Gretsch, Quasi-synchronous sampling algorithm and its applications, IEEE Trans. Instrum. Meas., IM-43, 204-209, 1994.

11. M. Bellanger, Digital Processing of Signals: Theory and Practice, Chichester: John Wiley & Sons, 1984.

12. M. Bertocco, C. Offelli, and D. Petri, Numerical

algorithms for power measurements, Europ. Trans. Electr. Power, ETEP 3, 91-101, 1993.

13. G. Zingales, Measurements on steady-state circuits (Chap. VI), in Methods and Instruments for Electrical Measurements (In Italian), Torino: UTET, 1980.

14. H.N. Norton, Thermometers (Chap. 19), in Handbook of Transducers, Englewood Cliffs, NJ: PrenticeHall, 1989.

15. Anonymous, Thermistor mounts and instrumentation (Chap. II), in Fundamental of RF and Microwave Power Measurements, Application Note 64-1, Hewlett Packard, 1978.

16. R.E. Pratt, Power measurements (15.1–15.16), in C.F. Coombs (Ed.), in Handbook of Electronic Instruments, New York: McGraw-Hill, 1995.

17. F. F. Mazda, High-frequency power measurements (Chap. VIII, 8.4), in Electronic Instruments and Measurement Techniques, Cambridge: Cambridge University Press, 1987.

18. Anonymous, Diode detector power sensors and instrumentation (Chap. IV), in Fundamental of RF and Microwave Power Measurements, Application Note 64-1, Hewlett Packard, 1978.

19. J.W. Gardner, Microsensors: Principles and Applications, Chichester: John Wiley & Sons, 1994.

20. H.N. Norton, Radiation pyrometers (Chap. 20), in Handbook of Transducers, Englewood Cliffs, NJ: Prentice-Hall, 1989.

21. F.F. Mazda, Pulse power measurement (Chap. VIII, 8.5), in Electronic Instruments and Measurement Techniques, Cambridge, U.K.: Cambridge University Press, 1987.

Further Information

F.K. Harris, The Measurement of Power (Chap. XI), in Electrical Measurements, Huntington, NY: R.E.

Krieger Publishing, 1975; a clear reference for line-frequency power measurements.

Anonymous, Fundamental of RF and Microwave Power Measurements, Application Note 64-1, Hewlett Packard, 1978; though not very recent, is a valid and comprehensive reference for main principles of high-frequency power

measurements.

J.J. Clarke and J.R. Stockton, Principles and theory of wattmeters operating on the basis of regularly spaced sample pairs, J. Phys. E. Sci. Instrum., 15, 645-652, 1982; gives basics of synchronous sampling for digital wattmeters.

T.S. Rathore, Theorems on power, mean and RMS values of uniformly sampled periodic signals, IEE Proc. Pt. A, 131, 598-600, 1984; provides fundamental theorems for effective synchronous sampling wattmeters.

J.K. Kolanko, Accurate measurement of power, energy, and true RMS voltage using synchronous counting, IEEE Trans. Instrum. Meas., IM-42, 752–754, 1993; provides information on the implementation of a synchronous dual-slope wattmeter.

G.N. Stenbakken, A wideband sampling wattmeter, IEEE Trans. Power App. Syst., PAS-103, 2919-2925, 1984; gives basics of asynchronous sampling-based wattmeters and criteria for computing uncertainty in time domain.

F. Avallone, C. De Capua, and C. Landi, Measurement station performance optimization for testing on high efficiency variable speed drives, Proc. IEEE IMTC/96 (Brussels, Belgium), 1098–1103, 1996; proposes an analytical model of uncertainty arising from power measurement systems working under highly distorted conditions.

F. Avallone, C. De Capua, and C. Landi, Metrological performance improvement for power measurements on variable speed drives, Measurement, 21, 1997, 17-24; shows how compute uncertainty of measuring chain components for power measurements under highly distorted conditions.

F. Avallone, C. De Capua, and C. Landi, Measurand reconstruction techniques applied to power measurements on high efficiency variable speed drives, Proc. of XIV IMEKO World Congress (Tampere, Fi), 1997; proposes a technique to improve accuracy of power measurements under highly distorted conditions.

F. Avallone, C. De Capua, and C. Landi, A digital technique based on real-time error compensation for high accuracy power measurement on variable speed drives, Proc. of IEEE IMTC/97 (Ottawa, Canada), 1997; reports about a real-time technique for error compensation of transducers working under highly distorted conditions.

J.C. Montano, A. Lopez, M. Castilla, and J. Gutierrrez, DSP-based algorithm for electric power measurement, IEE Proc. Pt.A, 140, 485–490, 1993; describes a Goertzel FFT-based algorithm to compute power under nonsinusoidal conditions.

S.L. Garverick, K. Fujino, D.T. McGrath, and R.D. Baertsch, A programmable mixed-signal ASIC for power metering, IEEE J. Solid State Circuits, 26, 2008–2015, 1991; reports about a programmable mixed analog-digital integrated circuit based on six first-order sigma-delta ADCs, a bit serial DSP, and a byte-wide static RAM for power metering.

G. Bucci, P. D'Innocenzo, and C. Landi, A modular high-speed dsp-based data acquisition apparatus for on-line quality analysis on power systems under non-sinusoidal conditions, Proc. of 9th IMEKO TC4 Int. Sym. (Budapest, Hungary), 286–289, 1996; shows the strategy of measurement system design for power measurements under non-sinusoidal conditions.

K.K. Clarke and D.T. Hess, A 1000 A/20 kV/25 kHz-500 kHz volt-ampere-wattmeter for loads with power factors from 0.001 to 1.00, IEEE Trans. Instrum. Meas., IM-45, 142-145, 1996; provides information on the implementation of an instrument to perform an accurate measurement of currents (1 A to 1000 A), voltages (100 V to 20 kV), and powers (100 W to 20 MW) over the frequency range from 25 kHz to 500 kHz.

P. Arpaia, G. Betta, A. Langella, and M. Vanacore, An Expert System for the Optimum Design of Measurement Systems, IEE Proc. (Pt. A), 142, 330–336, 1995; reports about an Artificial Intelligence tool for the automatic design of power measuring systems.

In any case, IEEE Transactions on Instrumentation and Measurement and Measurement journals provide

current research on power measurements.

1. D.F. Bullock, Methods of measuring apparent power, GE Meter, Missouri Valley Electrical Assoc., Annu. Eng. Conf., Kansas City, MO, 1996.

2. ANSI/IEEE Std 100, Dictionary of Electrical and Electronic Terms, 1992, New York: IEEE, 1993.

3. D.F. Bullock, Phase Angle Relationships, Part 1: Theory, GE Meter Arkansas Electric Meter School, 1995.

4. I. Assimov, Assimov's Biographical Encyclopedia of Science and Technology, New Rev. Ed., Garden City, NY: Doubleday & Co., 1972.

5. D.F. Bullock, private communication.

6. D.F. Bullock and D.D. Elmore, MIND UR PS & QS, GE Meter EEI-AEIC Meter and Service Committee Meeting, Dallas, TX, 1994.

## Further Information

H.L. Curtis and F.B. Silsbee, Definitions of power and related quantities, AIEE Summer Conf., 1935. TABLE 4.3 Manufacturers of Power Factor Measuring Harmonic Analyzers Amprobe Instruments Cutler Hammer Inc. GE Meter 630 Merrick Road Westinghouse & Cutler-Hammer Products 130 Main Street Lynbrook, NY 11563 Five Parkway Center Somersworth, NH 03878 Tel: (516) 593-5600 Pittsburgh, PA 15220 Tel: (603) 749-8477 Tel: (412) 937-6100 BMI Reliable Power Meters, Inc. 3250 Jay Street Dranetz Technologies, Inc. 400 Blossom Hill Road Santa Clara, CA 95054 1000 New Durham Road Los Gatos, CA 95032 Tel: (408) 970-3700 Edison, NJ 08818-4019 Tel: (408) 358-5100 Tel: (800) DRANTEC Cooper Power Systems Division Square D Power Logic 11131 Adams Road Fluke Corporation 330 Weakley Road P.O. Box 100 P.O. Box 9090 Smyrna, TN 37167-9969 Franksville, WI 53126-0100 Everett, WA 98206 Tel: (615) 459-8552 Tel: (414) 835-2921 Tel: (800) 44FLUKE

# 5 5. Phase Measurement


1. D. Gabor, The theory of communication, J. Inst. Elec. Eng., 93(III), 429-457, 1946.

2. A.V. Oppenheim and R.W. Schafer, Discrete-Time Signal Processing, Engelwood-Cliffs, NJ: PrenticeHall, 1989.

3. K.C. Pohlmann (Ed.), Advanced Digital Audio, Carmel, IN: Howard Sams and Co., 1991.

4. D. Rife and R. Boorstyn, Single tone parameter estimation from discrete-time observations, IEEE Trans. Inf. Theory, 20, 591-598, 1974.

5. T. Abotzoglou, A fast maximum likelihood algorithm for estimating the frequency of a sinusoid based on Newton's algorithm, IEEE Trans. Acoust., Speech Signal Process., 33, 77-89, 1985.

6. D. McMahon and R. Barrett, ML estimation of the fundamental frequency of a harmonic series, Proc. of ISSPA 87, Brisbane, Australia, 1987, 333-336.

7. A. Nehorai and B. Porat, Adaptive comb filtering for harmonic signal enhancement, IEEE Trans. Acoust., Speech Signal Process., 34, 1124-1138, 1986.

8. L. White, An iterative method for exact maximum likelihood estimation of the parameters of a harmonic series, IEEE Trans. Automat. Control, 38, 367-370 1993.

9. C. Kelly and S. Gupta, Discrete-time demodulation of continuous time signals, IEEE Trans. Inf. Theory, 18, 488-493, 1972.

10. B.D.O. Anderson and J.B. Moore, Optimal Filtering, Englewood Cliffs, NJ: Prentice-Hall, 1979.

11. P. Parker and B. Anderson, Frequency tracking of periodic noisy signals, Signal Processing, 20(2), 127-152, 1990.

12. R.E. Best, Phase-Locked Loops; Theory, Design and Applications, 2nd ed., New York: McGraw-Hill, 1993.

13. L. White, Estimation of the instantaneous frequency of a noisy signal, in B. Boashash (ed.), TimeFrequency Signal Analysis, Methods and Applications, Melbourne, Australia: Longman-Cheshire; New York: Halsted Press, 1992.

## Further Information

A.D. Helfrick and W.D. Cooper, Modern Electronic Instrumentation and Measurement Techniques, Englewood Cliffs, NJ: Prentice-Hall, 1990.

McGraw-Hill Encyclopedia of Science and Technology, 8th ed., New York: McGraw-Hill, 1997.

H. Taub and D.L. Schilling, Principles of Communication Systems, 2nd ed., New York: McGraw-Hill, 1986.

J.D. Lenk, Handbook of Oscilloscopes: Theory and Application, Englewood Cliffs, NJ: Prentice-Hall, 1982.

# 7 7. Electrical Conductivity and Resistivity


1. P. Drude, Zur elektronentheorie der metalle, Annalen der Physik, 1, 566–613, 1900; 3, 369–402, 1900. See Ref. [2] for modern discussions of the Drude model and electrical conductivity.

2. N.W. Ashcroft and N.D. Mermin, Solid State Physics, Philadelphia, PA: Saunders College, 1976; C. Kittel, Introduction to Solid State Physics, 7th ed., New York: John Wiley & Sons, 1996.

3. L.B. Valdes, Resistivity measurements on germanium for transistors, Proc. I.R.E., 42, 420–427, 1954.

4. L.J. van der Pauw, A method of measuring specific resistivity and Hall effect of discs of arbitrary shape, Philips Res. Rep., 13, 1–9, 1958.

5. R. Chwang, B.J. Smith, and C.R. Crowell, Contact size effects on the van der Pauw method for resistivity and Hall coefficient measurement, Solid-State Electron., 17, 1217–1227, 1974. r r r = + ( ) a b 2 f R R R R R R R R f > < > < > < + ( ) [ ] ( ) ∫ ( ) + ( ) Ï Ì Ó ˛ ý þ È Î Í ù û ú > < ln ln 4 1 1 4 1 1


Further Information


H.H. Wieder, Laboratory Notes on Electrical and Galvanomagnetic Measurements, New York: Elsevier, 1979.

L.I. Maissel, Electrical properties of metallic thin films, 13-1 to 13-33, in L.I. Maissel and R. Glang (Eds.), Handbook of Thin Film Technology, San Francisco: McGraw-Hill, 1970.

D.C. Look, Electrical Characterization of GaAs Materials and Devices, New York: John Wiley & Sons, 1989.

D.C. Look, Bulk and contact electrical properties by the magneto-transmission-line method: application to GaAs, Solid-State Electron., 30, 615–618, 1987.

# 8 8. Charge Measurement

1. E. Cohen and B. Taylor, The 1986 Adjustment of the Fundamental Physical Constants, Rev. Modern Phys., 59(4), 1121-1148, 1987.

2. F. Harris, Electrical Measurements, New York: John Wiley & Sons, 1952.

3. W. Michels, Electrical Measurements and Their Applications, Princeton, NJ: D. Van Nostrand, 1969.

4. A.J. Diefenderfer, Principles of Electronic Instrumentation, 3rd ed., Philadelphia: Saunders College Publishing, 1994.

5. Keithley Instruments, Inc., Low Level Measurements, 4th ed., Cleveland: Keithley Instruments, Inc., 1993.

6. K. Kawamura, S. Sakamoto, and F. Noto, Design and Development of New Electrostatic Voltmeter Using Strain Gauge, IEEE Trans. Ind. Appl., 25, 563-568, 1989.

7. C. Hsu and R. Muller, Micromechanical Electrostatic Voltmeter, Transducers '91. Int. Conf. SolidState Sensors Actuators, 1991, 659-662.

8. Trek Incorporated, 3922 Salt Works Rd., P.O. Box 728, Medina, NY 14103.

9. M.M. Horenstein, Measuring surface charge with a noncontacting voltmeter, Proc. IEEE Industry Appl. Soc. Annu. Meeting, Toronto, Ontario, Canada IAS'93, 3, 1811-1816, 1993.

10. P. Gunther, Determination of charge density and charge centroid location in electrets with semiconducting substrates, IEEE Trans. Electrical Insul., 27, 698-701, 1992.

11. B. MacDonald and B. Fallone, Surface-charge distributions on radiation-charged electrets. 7th Int. Symp. Electrets (ISE 7). Berlin, Germany, 1991, 798-803.

12. V. Radeka, Signal, noise and resolution in position-sensitive detectors, 20th Nuclear Science Symp. 5th Nuclear Power Syst. Symp., San Francisco, CA, IEEE Trans. Nucl. Sci., 21, 51-64, 1974.

13. N. Kutsuwada, T. Shohdohji, N. Okada, H. Izawa, T.

Sugai, Y. Nakamura, and T. Murata, Measurement of electric charge of electrophotographic toner, J. Imaging Sci. Technol., 37(5), 480-484, 1993.

14. M. Mehlin and R.M. Hess, Electrical charge measurement of toner particles using the q/d meter, J. Imaging Sci. Technol., 36, 142-150, 1992.

15. A.F. Hebard, G.S. LaRue, J.D. Phillips, and C.R. Fisel, Search for Fractional Charge in Near Zero: New Frontiers of Physics, New York: W. H. Freeman and Company, 1988, 511.

16. S. Buchman, T. Quinn, G.M. Keiser, D. Gill, and T.J. Sumner, Charge measurement and control for the Gravity Probe B gyroscopes, Rev. Sci. Instrum., 66, 120-129, 1995.

17. M.N. Horenstein, Peak sampled vibrating-reed for the measurement of electric fields in the presence of space charge, Rev. Sci. Instrum., 54, 591-593, 1983.

18. K.-C. Lin and T.G. Wang, Noncontact charge measurement, Rev. Sci. Instrum., 63, 2040-2043, 1992.

19. K. Hidaka, Progress in Japan of space charge field measurement in gaseous dielectrics using a Pockels sensor, IEEE Electrical Insul. Mag., 12(1), 17-28, 1996.

20. R. Kesavamoorthy, T. Sakuntala, and A.K. Arora, In situ measurement of charge on polystyrene particles in colloidal suspension, Meas. Sci. Technol., 1(5), 440-445, 1990.

[1] A.R. Von Hippel, Dielectric Materials and Applications, The M.I.T. Press, Cambridge, MA, 1961.

[2] N.E. Hill, W.E. Vaughan, A.H. Price, and M. Davies, Dielectric Properties and Molecular Behaviour, Van Nostrand Reinhold Co., London, 1969.

[3] H. Altschuler, Dielectric constant, in Handbook of Microwave Measurements, Vol. II, M. Sucher and J. Fox, Eds., Polytechnic Press, Brooklyn, NY, 1963.

[4] R. Chatterjee, Advanced Microwave Engineering, Ellis Horwood Ltd., Chichester, 1988.

[5] D.K. Misra, Permittivity measurement of modified infinite samples by a directional coupler and a sliding load, IEEE Trans. Microwave Theory Tech., 29(1), 65–67, 1981.

[6] D.K. Ghodgaonkar, V.V. Varadan, and V.K. Varadan, A free-space method for measurement of dielectric constants and loss tangents at microwave frequencies, IEEE Trans. Instrum. Meas., 38(3), 789–793, 1989.

[7] A.P. Gregory, R.N. Clarke, T.E. Hodgetts, and G.T. Symm, RF and microwave dielectric measurements upon layered materials using a reflectometric coaxial sensor, NPL Report DES 125, UK, March 1993.

[8] D. Misra, On the measurement of the complex permittivity of materials by an open-ended coaxial probe, IEEE Microwave Guided Wave Lett., 5(5), 161–163, 1995.

[9] S. Gabriel, R.W. Lau, and C. Gabriel, The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz, Phys. Med. Biol., 41, 2251–2269, 1996.

[10] M.F. Kabir, K.B. Khalid, W.M. Daud, and S. Aziz, Dielectric properties of rubber wood at microwave frequencies measured with an open-ended coaxial line, Wood Fiber Sci., 29(4), 319–324, 1997.

[11] Y. Xu, R.G. Bosisio, A. Bonincontro, F. Pedone, and G.F. Mariutti, On the measurement of microwave permittivity of biological samples using needle-type coaxial probes, IEEE Trans. Instrum. Meas., IM-42(4), 822–827, 1993.

[12] C.L. Pournaropoulos and D.K. Misra, The coaxial aperture electromagnetic sensor and its application in material characterization, Meas. Sci. Technol., 8, 1191-1202, 1997.

[13] H.C.F. Martens, Reedijk, and H.B. Brom, Measurement of the complex dielectric constant down to helium temperatures. I. Reflection method from 1 MHz to 20 GHz using an open-ended coaxial line, Rev. Scientific Instrum., 71(2), 473–477, 2000.

[14] S.D. Zandron, C. Pournaropoulos, and D.K. Misra, Complex permittivity measurement of materials by the open-ended coaxial probe technique, J. Wave-Mater. Interact., 5 & 6(4), 329–342, 1991.

# 11 11. Electric Field Strength1

1. J.A. Stratton, Electromagnetic Theory, New York: McGraw-Hill, 1941.

2. ANSI/IEEE Std. 1227-1990, IEEE Guide for the Measurement of dc Electric-Field Strength and Ion Related Quantities.

3. P.E. Secker and J.N. Chubb, Instrumentation for electrostatic measurements, J. Electrostatics, 16, 1–19, 1984.

4. R.E. Vosteen, Dc electrostatic voltmeters and fieldmeters, Conference Record of Ninth Annual Meeting of the IEEE Industrial Applications Society, October 1974.

5. P.J.L. Wildman, A device for measuring electric field in the presence of ionisation, J. Atmos. Terr. Phys., 27, 416–423, 1965.

6. M. Misakian, Generation and measurement of dc electric fields with space charge, J. Appl. Phys., 52, 3135–3144, 1981.

7. G.V. Keller and F.C. Frischknecht, Electrical Methods in Geophysical Prospecting, Oxford, U.K.: Pergamon Press, 1966.

8. A.A. Kaufman and G.V. Keller, The Magnetotelluric Sounding Method, Amsterdam: Elsevier, 1981.

9. IEEE Std. 291-1991, IEEE Standard Methods for Measuring Electromagnetic Field Strength of Sinusoidal Continuous Waves, 30 Hz to 30 GHz.

10. E.C. Jordan and K.G. Balmain, Electromagnetic Waves and Radiating Systems, 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 1968.

11. ANSI/IEEE Std. 644-1987, IEEE Standard Procedures for Measurement of Power Frequency Electric and Magnetic Fields from Ac Power Lines.

12. IEEE Std. 1308-1994, IEEE Recommended Practice for Instrumentation: Specifications for Magnetic Flux Density and Electric Field Strength Meters — 10 Hz to 3 kHz.

13. C.J. Miller, The measurements of electric fields in live line working, IEEE Trans. Power Apparatus Sys., PAS-16, 493–498, 1967.

14. E.V. Jull, Aperture Antennas and Diffraction Theory, Stevenage, U.K.: Peter Peregrinus, 1981.

15. M. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions, Nat. Bur. Stand. (U.S.), Spec. Pub. AMS 55, 1968.

16. D.A. Hill, M. Kanda, E.B. Larsen, G.H. Koepke, and R.D. Orr, Generating standard reference electromagnetic fields in the NIST anechoic chamber, 0.2 to 40 GHz, Natl. Inst. Stand. Technol. Tech. Note 1335, 1990.

17. M. Kanda, An electromagnetic near-field sensor for simultaneous electric and magnetic-field measurements, IEEE Trans. Electromag. Compat., EMC-26, 102-110, 1984.

18. M. Kanda and D.A. Hill, A three-loop method for determining the radiation characteristics of an electrically small source, IEEE Trans. Electromag. Compat., 34, 1-3, 1992.

19. T.T. Wu, Theory of the thin circular antenna, J. Math. Phys., 3, 1301-1304, 1962.

20. I. Sreenivasiah, D.C. Chang, and M.T. Ma, Emission characteristics of electrically small radiating sources from tests inside a TEM cell, IEEE Trans. Electromag. Compat., EMC-23, 113-121, 1981.

21. D.R. Novotny, K.D. Masterson, and M. Kanda, An optically linked three-loop antenna system for determining the radiation characteristics of an electrically small source, IEEE Int. EMC Symp., 1993, 300-305.

22. M. Kanda, Standard probes for electromagnetic field measurements, IEEE Trans. Antennas Propagat., 41, 1349-1364, 1993.

23. M. Kanda and L.D. Driver, An isotropic electric-field probe with tapered resistive dipoles for broadband use, 100 kHz to 18 GHz, IEEE Trans. Microwave Theory Techniques, MTT-35, 124-130, 1987.

1. K.S. Lion, Instrumentation in Scientific Research. Electrical Input Transducers, New York: McGrawHill, 1959.

2. H.R. Everett, Sensors for Mobile Robots: Theory and Application, Wellesley, MA: A.K. Peters, 1995.

3. J.E. Lenz, A review of magnetic sensors, Proc. of IEEE, 78, 973–989, 1990.

4. M.A. Payne, SI and Gaussian cgs units, conversions and equations for use in geomagnetism, Phys. of Earth and Planitary Interiors, 26, P10–P16, 1981.

5. F.W. Grover, Induction Calculations. Working Formulas and Tables, New York: Dover Publications, 1973.

6. R.M. Bozorth, Ferromagnetism, New York: D. Van Nostrand, 1951.

7. R.M. Bozorth and D.M. Chapin, Demagnetization factors of rods, J. Appl. Phys., 13, 320–326, 1942.

8. S. A. Macintyre, A portable low noise low current three-axis search coil magnetometer, IEEE Trans. Magnetics, MAG-16, 761–763, 1980.

9. J.P. Hauser, A 20-Hz to 200-kHz magnetic flux probe for EMI surveys, IEEE Trans. Electromagnetic Compatibility, 32, 67–69, 1990.

10. F. Primdahl, The fluxgate magnetometer, J. Phys. E: Sci. Instrum., 1, 242–253, 1979.

11. C.J. Pellerin and M.H. Acuna, A miniature two-axis fluxgate magnetometer, NASA Technical Note, TN D-5325, NASA, 1970.

12. S.V. Marshall, A gamma-level portable ring-core magnetometer, IEEE Trans. Magnetics, MAG-7, 183–185, 1971.

13. W.A. Geyger, Nonlinear-Magnetic Control Devices, New York: McGraw-Hill, 1964.

14. M. Acuna, C. Scearce, J. Seek, and J. Schelfiele, The MAGSAT vector magnetometer — a precise fluxgate magnetometer for the measurement of the geomagnetic field, NASA Technical Report.

15. D. Cohen, Measurements of the magnetic field produced by the human heart, brain and lung, IEEE Trans. Magnetics, MAG-11, 694–700, 1975.

16. C.M. Falco and I.K. Schuller, SQUIDs and their sensitivity for geophysical applications, SQUID Applications to Geophysics, The Society of Exploration Geophysicists, 13–18, 1981.

17. J. Clark, SQUIDs, Sci. Am., 46–53, August 1994.

18. T.H. Casselman and S.A. Hanka, Calculation of the performance of a magnetoresistive permalloy magnetic field sensor, IEEE Trans. Magnetics, MAG-16, 461–464, 1980.

19. W. Kwaitkawski and S. Tumanski, The permalloy magnetoresistive sensors-properties and applications, J. Phys. E: Sci. Instrum., 19, 502–515, 1986.

20. Permalloy Magnetic Sensors, Honeywell Technical Note.

21. U. Dibbern and A. Petersen, The magnetoresistor sensor — a sensitive device for detecting magnetic field variations, Electronic Components and Applications, 5(3), 148–153, 1983.

22. S. Tumanski and M.M. Stabrowski, Optimization of the performance of a thin film magnetoresistive sensor, IEEE Trans. Magnetics, MAG-20, 963–965, 1984.

23. L.W. Parson and Z.M. Wiatr, Rubidium vapor magnetometer, J. Sci. Instrum., 39, 292–299, 1962.

24. W.H. Farthing and W.C. Folz, Rubidium vapor magnetometer for near Earth orbiting spacecraft, Rev. Sci. Instrum., 38, 1023–1030, 1967.

25. F. Hartmann, Resonance magnetometers, IEEE Trans. Magnetics, MAG-8, 66–75, 1972.

26. F. Wellstood, C. Heiden, and J. Clark, Integrated DC SQUID magnetometer with high slew rate, Rev. Sci. Instrum., 66, 952–957, 1984.

# 13 13. Permeability and Hysteresis Measurement


1. B.D. Cullity, Introduction to Magnetic Materials, Reading, MA: Addison-Wesley, 1972.

2. D.R. Lide (Ed.), CRC Handbook of Chemistry and Physics, Boca Raton, FL: CRC Press, 1992-3.

3. Anonymous, 1995 Annual Book of ASTM Standards, Philadelphia, PA: ASTM, 3.04, 1995.

4. C. Kittel, Introduction to Solid State Physics, 5th ed., New York: John Wiley & Sons, 1976.

5. Page 11 of Ref. [3].

6. R. Thomas and G. Morgan, Proc. Eleventh Ann. Conf. on Properties and Applications of Magnetic Materials, Chicago, 1992.

7. F. Brailsford, Magnetic Materials, 3rd ed., London: Metuen and New York: Wiley, 1960.

1. R.C. Dorf, Introduction to Electric Circuits, New York: John Wiley & Sons, 1989.

2. J.P. Bentley, Principles of Measurement Systems, 2nd ed., Harlow: Longman Group U.K., New York: John Wiley & Sons, 1988.

3. A.D. Helfrick and W.D. Cooper, Modern Electronic Instrumentation and Measurement Techniques, Englewood Cliffs, NJ: Prentice-Hall, 1990.

4. L.D. Jones and A.F. Chin, Electronic Instruments and Measurements, 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 1991.

5. M.U. Reissland, Electrical Measurement, New York: John Wiley & Sons, 1989.

6. P.H. Sydenham (Ed.), Handbook of Measurement Science, Vol. 1, Theoretical Fundamentals, New York: John Wiley & Sons, 1982.

TABLE 14.1 Basic Features of Selected Types of LCR Meters Manufacturer, Model Measurement Range Basic (Designation) of Inductance Frequency Accuracy Price Leader LCR 740 0.1 mH–1100 H int. 1 kHz 0.5% $545 (LCR bridge) ext. 50 Hz–40 kHz Electro Scientific 200 mH–200 H 1 kHz (3.5 digit) $995 Industries 253 (Digital impedance meter) Stanford RS 0.1 nH–100 kH 100 Hz–10 kHz 0.2% $1295 SR 715 ($1425) (LCR meter) Wayne Kerr 4250 0.01 nH–10 kH 120 Hz–100 kHz 0.1% $3500 General Radio 1689 0.00001 mH–99.999 H 12 Hz–100 kHz 0.02% $4120 (Precision LCR meter) Hewlett-Packard 4284A 0.01 nH–99.9999 kH 20 Hz–1 MHz 0.05% $9500 (Precision LCR meter) L C R C R = = 2 1 2 2 2 w w r r r ,

7. P.H. Sydenham (Ed.), Handbook of Measurement Science, Vol. 2, Practical Fundamentals, New York: John Wiley & Sons, 1983.

8. R.L. Bonebreak, Practical Techniques of Electronic Circuit Design, 2nd ed., New York: John Wiley & Sons, 1987.

9. S. Hashimoto and T. Tamamura, An automatic wide-range digital LCR meter, Hewlett-Packard J., 8, 9–15, 1976.

10. T. Wakasugi, T. Kyo, and T. Tamamura, New multi-frequency LCZ meters offer higher-speed impedance measurements, Hewlett-Packard J., 7, 32–38, 1983.

11. T. Yonekura, High-frequency impedance analyzer, Hewlett-Packard J., 5, 67–74, 1994.

12. M.A. Atmanand, V.J. Kumar, and V.G.K. Murti, A novel method of measurement of L and C, IEEE Trans. Instrum. Meas., 4, 898–903, 1995.

13. M.A. Atmanand, V.J. Kumar, and V.G.K. Murti, A microcontroller- based quasi-balanced bridge for the measurement of L, C and R, IEEE Trans. Instrum. Meas., 3, 757–761, 1996.

14. J. Gajda and M. Szyper, Electromagnetic transient states in the micro resistance measurements, Systems Analysis Modeling Simulation, Amsterdam: Overseas Publishers Association, 22, 47–52, 1996.

15. Information from Internet on LCR meters.

# 15 15. Immittance Measurement

1. B.M. Oliver and J.M. Cage, Electronic Measurements and Instrumentation, New York: McGraw-Hill, 1971.

2. H.W. Bode, Network Analysis and Feedback Amplifier Design, Princeton: Van Nostrand, 1959. TABLE 15.2 Companies Producing Immittance Measurement Equipment Agilent Headquarters Rohde & Schwarz, Inc. 395 Page Mill Rd. 7150-K Riverwood Drive P.O. Box 10395 Columbia, MD 21046 Palo Alto, CA 94303 Tel: (410) 910-7800 Tel: (877) 4-Agilent www.rsd.de/www/dev_center.nsf/USA www.agilent.com SRS Stanford Research Systems Anritsu Co. 1290-D Reamwood Ave. 1155 East Collins Blvd. Sunnyvale, CA 94089 Richardson, TX 75081 Tel: (408) 744-9040 Tel: (800) ANRITSU (267-4878) www.thinksrs.com www.global.anritsu.com TTi (Thurlby Thandar Instruments Ltd.) Fluke Corporation Glebe Road 6929 Seaway Boulevard Huntingdon P.O. Box 9090 Cambs. PE29 7DR Everett, WA 98206 U.K. Tel: (800) 44-FLUKE Tel: +44-1480-412451 www.fluke.com www.tti-test.com Keithley Instruments, Inc. Voltech Instruments, Inc. 28775 Aurora Road 11637 Kelly Road Cleveland, OH 44139 Suite 306 Tel: (800) 552-1115 Fort Myers, FL 33908-2544 www.keithley.com Tel: (239) 437-0494 www.voltech.com QuadTech, Inc. 5 Clock Tower Place Suite 210 East Maynard, MA 01754 Tel: (800) 253-1230 www.quadtechinc.com

3. A.T. de Hoop, Handbook of Radiation and Scattering of Waves, London: Academic Press, 1995.

4. R.E. Collin, Foundations for Microwave Engineering, New York: McGraw-Hill, 1992.

5. P.I. Somlo and J.D. Hunter, Microwave Impedance Measurement, London: Peter Peregrinus, 1985.

6. R.L. Thomas, A Practical Introduction to Impedance Matching, Dedham, MA: Artech House, 1976.

7. M. Honda, The Impedance Measurement Handbook, Yokogawa: Hewlett-Packard, 1989.

8. Anonymous, 4800A Vector Impedance Meter Operating and Service Manual, Rockaway: HewlettPackard, 1967.

9. C.R. Paul, S.A. Nasar, and L.E. Unnewehr, Introduction to Electrical Engineering, New York: McGraw-Hill, 1992.

10. T.S. Laverghetta, Modern Microwave Measurements and Techniques, Norwood, MA: Artech House, 1988.

11. A.E. Bailey (Ed.), Microwave Measurement, London: Peter Peregrinus, 1985.

12. T.S. Laverghetta, Handbook of Microwave Testing, Dedham, MA: Artech House, 1981.

13. E.K. Miller (Ed.), Time-Domain Measurements in Electromagnetics, New York: Van Nostrand Reinhold, 1986.

14. Anonymous, TDR Fundamentals for Use with HP 54120T Digitizing Oscilloscope and TDR, Application notes AN 62, Palo Alto, CA: Hewlett-Packard, 1988.

Further Information

L.S. Bobrow, Fundamentals of Electrical Engineering, New York: Oxford University Press, 1996.

W.H. Roadstrum and D.H. Wolaver, Electrical Engineering for all Engineers, New York: John Wiley & Sons, 1994.

A.S. Morris, Principles of Measurement and Instrumentation, London: Prentice Hall International (U.K.), 1993.

G.H. Bryant, Principles of Microwave Measurement, London: Peter Peregrinus, 1988.

Anonymous, Low Level Measurements, Cleveland, OH: Keithley Instruments, 1984.

# 16 16. Q Factor Measurement

1. Reference Data for Engineers: Radio Electronics, Computer and Communications, 8th ed., Sams, Englewood Cliffs, NJ: Prentice-Hall Computer Publishing, 1993.

2. A. Helfrick and W. Cooper, Modern Instrumentation and Measurement Techniques, Englewood Cliffs, NJ: Prentice-Hall, 1990.

3, D. Kajfez, Q factor, thesis, Vector Fields, 1994.

1. P.R. Gray and R.G. Meyer, Analysis and Design of Analog Integrated Circuits, New York: John Wiley & Sons, 1993.

2. P. Horowitz and W. Hill, The Art of Electronics, 2nd ed., New York: Cambridge Press, 1983.

3. W.M. Leach, Jr., Fundamentals of low-noise analog circuit design, Proc. IEEE, 82(10), 1515–1538, 1994.

4. C.D. Motchenbacher and J.A. Connelly, Low Noise Electronic System Design, New York: Wiley, 1993.

5. Y. Netzer, The design of low-noise amplifiers, Proc. IEEE, 69, 728-741, 1981.

6. H.W. Ott, Noise Reduction Techniques in Electronic Systems, 2nd ed., New York: Wiley, 1988.

7. A. Van der Ziel, Noise in solid-state devices and lasers, Proc. IEEE, 58, 1178-1206, 1970.

8. H.A. Haus et al., Representation of noise in linear twoports, Proc. IRE, 48, 69–74, 1960.

9. R.E. Burgess, Electronic fluctuations in semiconductors, Br. J. Appl. Phys., 6, 185–190, 1955.

10. H. Fukui, The noise performance of microwave transistors, IEEE Trans. Electron Devices, ED-13, 329–341, 1966.

11. H. Rothe and W. Dahlke, Theory of noisy fourpoles, Proc. IRE, 44, 811–818, 1956.

12. H.A. Haus et al., IRE standards on methods of measuring noise in linear twoports, 1959, Proc. IRE, 48, 60–68, 1960.

13. A.L. Lance, Introduction to Microwave Theory and Measurements, New York: McGraw-Hill, 1964.

## Further Information

H. Fukui, Low-Noise Microwave Transistors & Amplifiers, New York: IEEE Press, 1981.

M.S. Gupta, Ed., Electrical Noise: Fundamentals & Sources, New York: IEEE Press, 1977.

A. Van der Ziel, Noise: Sources, Characterization, Measurements, Englewood Cliffs, NJ: Prentice-Hall, 1970.

J.C. Bertails, Low frequency noise considerations for MOS amplifier design, IEEE J. Solid-State Circuits, SC-14, 773–776, 1979.

C.A. Liechti, Microwave field-effect transistors — 1976, IEEE Trans. Microwave Theory and Tech., MTT-24, 279–300, 1976.

M. Steyaert, Z.Y. Chang, and W. Sansen, Low-noise monolithic amplifier design: bipolar vs. CMOS, Analog Integrated Circuits and Signal Processing, 1(1), 9–19, 1991.

# 19 19. Microwave Measurement


1. A. Fantom, Radio Frequency and Microwave Power Measurements, London: Peter Peregrinus, 1990.

2. A.L. Lance, Microwave Measurements in Handbook of Microwave Optical Components, Vol. 1, K. Chang (Ed.), New York: John Wiley & Sons, 1989.

3. J. Minck, Fundamentals of RF and microwave power measurements, Hewlett-Packard Application Note 64-1A, 1997.

4. G.H. Bryant, Principles of Microwave Measurements, London: Peter Peregrinus, 1988.

5. Microwave Powermate, Marconi Instruments Ltd., 1989.

6. D. Janik, H. Wolf, and R. Schneider, High-Tc edge bolometer for detecting guided millimeter waves, IEEE Trans. Appl. Superconduct., 3, 2148–2151, 1993.

7. P.L. Richards, Bolometers for infrared and millimeter waves, J. Appl. Phys., 76(1), 1–24, 1994.

8. D.M. Rowe, Handbook of Thermoelectrics, Boca Raton, FL: CRC Press, 1995.

9. P. Kopystynski, E. Obermayer, H. Delfs, W. Hohenester, and A. Löser, Silicon power sensor from dc to microwave, in Micro Systems Technologies 90, Berlin: Springer, 1990, 605–610.

10. G.C.M. Meijer and A.W. Herwaarden, Thermal Sources, Bristol: IOP Ltd., 1994.

11. T. Närhi, Nonlinearity characterisation of microwave detectors for radiometer applications, Electron. Lett., 32, 224–225, 1996.

12. F. Wiedmann, B. Huyart, E. Bergeault, and L. Jallet, New structure for a six-port reflectometer in monolithic microwave integrated-circuit technology, IEEE Trans. Instrument. Measure., 46(2), 527–530, 1997.

13. A. Dehé, V. Krozer, B. Chen, and H. L. Hartnagel, High-sensitivity microwave power sensor for GaAs-MMIC implementation, Electron. Lett., 32(23), 2149–2150, 1996.

14. Circuits Multi Projects, Information CMP — 42. Grenoble, France, December 1996.

15. Hints for making better spectrum analyzer measurements, Hewlett-Packard Application Note, No. 1286-1, 1997.

16. C. Brown, Spectrum Analysis Basics, Hewlett-Packard 1997 Back to Basics Seminar, 1997. TABLE 19.6 Companies Supplying Network Analyzers for S-Parameter Measurements Frequency Company min max Method Cal. Methods Hewlett-Packard HP 8510 45 MHz 110 GHz Heterodyne SOLT, TLR, LRL, LRM, TRM Wiltron 45 MHz 110 GHz Heterodyne SOLT, TLR, LRL, LRM, TRM Rhode & Schwarz 10 Hz 4 GHz Heterodyne SOLT, TLR, LTL, LRM, TRM AB Millimeterique 2 GHz 800 GHz Heterodyne TLR, LRL, proprietary

17. A. Kiiss, Microwave instrumentation and measurements, in Handbook of Microwave Technology, T.K. Ishii (Ed.), San Diego: Academic Press, 1995, Vol. 2, 562.

18. H. Brand, Spectrum analyzers: precision test instruments, Microwave J., 37(3), 98, 1994.

19. Spectrum analysis: Spectrum analysis basics, Hewlett-Packard Application Note 150, November 1989.

20. T.K. Ishii, Spectrum Analysis: Amplitude and Frequency Modulation, Application Note 150-1, January 1989.

21. M. Sucher and J. Fox, Handbook of Microwave Measurements, Vol. II, Polytechnic Press of the Polytechnic Institute of Brooklyn, 1963, chap. VIII.

22. E.L. Ginzton, Microwave Measurements, New York: McGraw-Hill, 1957, chap. 7 and 8.

23. F.E. Terman and J.M. Pettit, Electronic Measurements, New York: McGraw-Hill, 1952, chapter 4–15.

24. C.A. Lee and G.C. Dalman, Microwave Devices, Circuits and Their Interaction, New York: John Wiley & Sons, 1994.

25. G.D. Vendelin, A.M. Pavio, and U.L. Rohde, Microwave Circuit Design Using Linear and Nonlinear Techniques, New York: John Wiley & Sons, 1990.

26. J. Dobrowolski, Microwave Circuit Analysis with Wave Variables, Norwood, MA: Artech House, 1991.

27. G. Gonzalez, Microwave Transistor Amplifiers, Englewood Cliffs, NJ: Prentice-Hall, 1984.

28. D. Weiner and G. Naditch, A scattering variable approach to the Volterro analysis of nonlinear systems, IEEE Trans. Microwave Theory & Technique, MTT-24(7), 422–433, 1976.

29. D.A. Frickey, Conversion between S, Z, Y, H, ABSD, and T parameters which are valid for complex source and load impedances, IEEE Trans. Microwave Theory & Technique, MTT-42, 205-211, 1994.

30. T.K. Ishii, Handbook of Microwave Technology, San Diego, CA: Academic Press, 1995.

31. R.A. Hackborn, An automatic network analyzer system, Microwave J., 45–52, 1968.

32. S. Rehnmark, On the calibration process of automatic network analyzer systems, IEEE Trans. Microwave Theory & Technique, MTT-22(4), 457–458, 1974.

33. J. Fitzpatrick, Error models of vector systems measurements, Microwave J., 21(5), 63–66, 1978.

34. D. Rytting, An analysis of vector measurement accuracy enhancement techniques, RF & Microwave Measurement Symp. and Exhibition, Hewlett Packard, 1982.

35. N.R. Franzen and R.A. Speciale, A new procedure for system calibration and error removal in automated S-parameter measurements, 5th European Microwave Conf., 1975, 69–73.

36. R.A. Soares and C.A. Hoer, A unified mathematical approach to two-port calibration techniques and some applications, IEEE Trans. Microwave Theory & Techniques, MTT-37(11), 1669–1674, 1989.

37. R.A. Soares, GaAs MESFET Circuit Design, Norwood, MA: Artech House, 1988.

38. Cascade Microtech, Microwave Wafer Probe Calibration Standards: HP8510 Network Analyzer Input, Cascade Microtech Instruction Manual, 1990.

39. Understanding the fundamental principles of vector network analysis, Hewlett-Packard Application Note 1287-1, 1997.

40. B. Donecker, Determining the measurement accuracy of the HP8510 microwave network analyzer, RF & Microwave

Measurement Symp. and Exhibition, Hewlett-Packard, March 1985.

41. Exploring the architectures of network analyzers, Hewlett-Packard Application Note 1287-2, 1997.

42. Applying error correction to network analyzer measurements, Hewlett-Packard Application Note 1287-3, 1997.

43. D. Ballo, Network analyzer basics, Hewlett-Packard 1997 Back to Basics Seminar, 1997.

44. H.J. Eul and B. Schiek, Thru-Match-Reflect: one result of a rigorous theory for de-embedding and network analyzer calibration, 18th European Microwave Conf., 1988, 909–914.

45. G.F. Engen and C.A. Hoer, Thru-Reflect-Line: an improved technique for calibrating the dual 6-port automatic network analyzer, IEEE Trans. Microwave Theory & Technique, MTT-27(12), 993–998, 1979. II -1 II

Signal

Processing 20 Amplifiers and Signal

# 20 20. Amplifiers and Signal Conditioners

1. R. Pallás-Areny and J.G. Webster, Sensors and Signal Conditioning, 2nd ed., New York: John Wiley & Sons, 2001.

2. J. Graeme, Photodiode Amplifiers, Op Amp Solutions, New York: McGraw-Hill, 1996.

3. C. Kitchin and L. Counts, Instrumentation Amplifier Application Guide, 2nd ed., Application Note, Norwood, MA: Analog Devices, 1992.

4. C.D. Motchenbacher and J.A. Connelly, Low-Noise Electronic System Design, New York: John Wiley & Sons, 1993.

5. S. Franco, Design with Operational Amplifiers and Analog Integrated Circuits, 2nd ed., New York: McGraw-Hill, 1998.

6. R. Pallás-Areny and J.G. Webster, Common mode rejection ratio in differential amplifiers, IEEE Trans. Instrum. Meas., 40, 669–676, 1991.

7. R. Pallás-Areny and O. Casas, A novel differential synchronous demodulator for ac signals, IEEE Trans. Instrum. Meas., 45, 413–416, 1996.

8. M.L. Meade, Lock-in Amplifiers: Principles and Applications, London: Peter Peregrinus, 1984.

9. W.G. Jung, IC Op Amp Cookbook, 3rd ed., Indianapolis, IN: Howard W. Sams, 1986.

10. Y.J. Wong and W.E. Ott, Function Circuits Design and Application, New York: McGraw-Hill, 1976.

11. A.J. Peyton and V. Walsh, Analog Electronics with Op Amps, Cambridge, U.K.: Cambridge University Press, 1993.

12. D.H. Sheingold, Ed., Multiplier Application Guide, Norwood, MA: Analog Devices, 1978.

## Further Information

B.W.G. Newby, Electronic Signal Conditioning, Oxford, U.K.: Butterworth-Heinemann, 1994, is a book for those in the first year of an engineering degree. It covers analog and digital techniques at beginners’ level, proposes simple exercises, and provides clear explanations supported by a minimum of equations.

P. Horowitz and W. Hill, The Art of Electronics, 2nd ed., Cambridge, U.K.: Cambridge University Press, 1989. This is a highly recommended book for anyone interested in building electronic circuits without worrying about internal details for active components.

M.N. Horenstein, Microelectronic Circuits and Devices, 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 1996, is an introductory electronics textbook for electrical or computer engineering students. It provides many examples and proposes many more problems, for some of which solutions are offered.

J. Dostál, Operational Amplifiers, 2nd ed., Oxford, U.K.: Butterworth-Heinemann, 1993, provides a good combination of theory and practical design ideas. It includes complete tables which summarize errors and equivalent circuits for many op amp applications.

Manufacturers' data books provide a wealth of information, albeit nonuniformly. Application notes for special components should be consulted before undertaking any serious project. In addition, application notes provide handy solutions to difficult problems and often inspire good designs. Most manufacturers offer such literature free of charge. The following have shown to be particularly useful and easy to obtain: 1993 Applications Reference Manual, Analog Devices; 1994 IC Applications Handbook, Burr-Brown; 1990 Linear Applications Handbook and 1993 Linear Applications Handbook, Vol. II, Linear Technology; 1994 Linear Application Handbook, National Semiconductor; Linear and Interface Circuit Applications, Vols. 1, 2, and 3, Texas Instruments.

R. Pallás-Areny and J.G. Webster, Analog Signal Processing, New York: John Wiley & Sons, 1999, offers a design-oriented approach to processing instrumentation signals using standard analog integrated circuits, that relies on signal classification, analog domain conversions, error analysis, interference rejection and noise reduction, and highlights differential circuits.

# 21 21. Modulation


1. A.B. Carlson, Communication Systems, 3rd ed., New York: McGraw-Hill, 1986.

2. H. Taub and D.L. Schilling, Principles of Communication Systems, New York: McGraw-Hill, 1971.

3. M. Schwartz, Information Transmission, Modulation, and Noise, 4th ed., New York: McGraw-Hill, 1990.

4. W. Tomasi, Advanced Electronic Communications Systems, 2nd ed., Englewood Cliffs, NJ: PrenticeHall, 1992.

5. D.J. Nowicki, Voltage measurement and signal demodulation, in J. G. Webster, Ed., Electrical Impedance Tomography, Bristol, UK: Adam Hilger, 1990.

6. M.L. Meade, Lock-in Amplifiers: Principles and Applications, London: Peregrinus, 1984.

7. J.G. Webster, Ed., Electrical Impedance Tomography, Bristol, UK: Adam Hilger, 1990.

8. D.L. Hershberger, Build a synchronous detector for AM radio, Pop. Electron., 20 (4), 61, 66–71, 1982.

9. G.A. Breed, Receiver basics—part 2. Fundamental receiver architectures, RF Design, 17 (3), 84, 86, 88–89, 1994.

10. Anonymous, AD630 balanced modulator/demodulator, in Analog Devices 1992 Special Linear Reference Manual, pp. 2.35–2.41, 1992.

11. J.G. Webster, Ed., Medical Instrumentation: Application and Design., 3rd ed., New York: John Wiley & Sons, 1998.

12. D.M. Beams. Project TUNA — the development of a LabVIEW virtual instrument as a class project in a junior-level electronics course, Comput. Educ. J. 11, 1, 28–33, 2001. Additional information available at

13. W.H. Beyer, CRC Standard Mathematical Tables, 28th ed., Boca Raton, FL: CRC Press, 1987.

14. J. Sherwin and T. Regan, FM remote speaker system, National Semiconductor Application Note AN-146, in National Semiconductor Corp. 1991 Linear Applications Databook, 1991.

15. J.J. DeFrance, Communication Electronics Circuits, 2nd ed., San Francisco, CA: Rinehart Press, 1972.

16. D.G. Sokol, R.B. Whitaker, J.J. Lord, and D.M. Beams, Combine data center, U.S. Patent No. 4,376,298, 1983.

17. K.P. Cohen, D. Panescu, J.H. Booske, J.G. Webster, and W.J. Tompkins, Design of an inductive plethysmograph for ventilation measurement, Physiol. Meas., 15, 217–229, 1994.

18. W.J. Bachman, Capacitive-type seed sensor for a planter monitor, U.S. Patent No. 4,782,282, 1988.

19. W.J. Bachman, Private communication, 1995.

1. S. Mitra and J. Kaiser, Handbook for Digital Signal Processing, New York: John Wiley & Sons, 1993.

2. E. Christian and E. Eisenmann, Filter Design Tables and Graphs, New York: John Wiley & Sons, 1966.

3. A. Antoniou, Digital Filter: Analysis and Design, New York: McGraw-Hill, 1979.

4. R. Saal, Handbuch zum Filterentwurf [Handbook of filter design], Frankfurt: Allgemeine ElektricitätsGesellschaft AEG-Telefunken, 1979.

5. G. Temes and S. Mitra, Modern Filter Theory and Design, New York: John Wiley & Sons, 1973.

6. E. Cunningham, Digital Filtering, New York: John Wiley & Sons, 1995.

FIGURE 22.17 An adaptive filter uses an adaptive algorithm to change the performance of a digital filter in response

to defined conditions.

7. P.P. Vaidyanathan, Multirate Systems and Filter Banks, Englewood Cliffs, NJ: Prentice-Hall, 1993.

8. M. Cerna and R. Jamal, The design of digital filters using graphical programming techniques, Proc. MessComp, 232–238, 1995.

9. LabVIEW Digital Filter Design Toolkit, Austin, TX: National Instruments, 1996.

10. R.A. Roberts and C.T. Multis, Digital Signal Processing, Reading, MA: Addison-Wesley, 1987.

11. A.H. Gray, Jr. and J.D. Markel, Digital lattice and ladder filter synthesis, IEEE Trans. Acoust. Speech Signal Process., ASSP-23: 268–277, 1975.

12. DSP Committee, IEEE ASSP, Eds., Programs for Digital Signal Processing, New York: IEEE Press, 1979.

13. F. Taylor, Digital Filter Design Handbook, New York: Marcel Dekker, 1983.

14. A. Fettweis, Wave digital filters: theory and practice,

Proc. IEEE, 74, 270-327, 1986.

15. W.K. Chen, The Circuits and Filters Handbook, Boca Raton, FL: CRC Press, 1995.

16. M.L. Honig and D. Messerschmitt, Adaptive Filters — Structures, Algorithms, and Applications, Boston: Kluwer Academic Publishers, 1984.

17. M. Bellanger, Digital Processing of Signals — Theory and Practice, New York: John Wiley & Sons, 1988.

18. R.C. Dorf, Ed., The Electrical Engineering Handbook, Boca Raton, FL: CRC Press, 1993.

19. A. Zverev, Handbook of Filter Synthesis, New York: John Wiley & Sons, 1967.

20. C. Lindquist, Active Network Design, Long Beach, CA: Steward & Sons, 1977.

# 23 23. Spectrum Analysis and Correlation


1. B.P. Lathi, Communication Systems, New York: John Wiley & Sons, 1968.

2. R.N. Bracewell, The Fourier Transform and Its Applications, Englewood Cliffs, NJ: Prentice-Hall, 1988.

3. E. Kreyszig, Advanced Engineering Mathematics, New York: John Wiley & Sons, 1962.

4. P.A. Lynn and W. Fuerst, Introductory Digital Signal Processing with Computer Applications, Chichester: John Wiley & Sons, 1989.

5. J.G. Proakis and D.G. Manolakis, Digital Signal Processing: Principles, Algorithms, and Applications, 2nd ed., New York: Macmillan, 1992.

6. J.S. Bendat and A.G. Piersol, Random Data: Analysis and Measurement Procedures, 2nd ed., New York: John Wiley & Sons, 1986.

7. A.V. Oppenheim and R.W. Schafer, Discrete-Time Signal Processing, Englewood Cliffs, NJ: PrenticeHall, 1989.

8. G.D. Bergland, A guided tour of the fast Fourier transform, IEEE Spectrum, 6: 41–52, 1969.

9. M.B. Priestley, Spectral Analysis and Time Series, London: Academic Press, 1981.

10. J.W. Cooley and J.W. Tukey, An algorithm for the machine computation of complex Fourier series, Math. Computation, 19: 297–301, 1965.

11. L.R. Rabiner and B. Gold, Theory and Application of Digital Signal Processing, Englewood Cliffs, NJ: Prentice-Hall, 1975.

12. E.O. Brigham, The Fast Fourier Transform and Its Applications, 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 1988.

13. Digital Signal Processing Committee, Programs for Digital Signal Processing, New York: IEEE Press, 1979.

14. R.C. Singleton, An algorithm for computing the mixed radix fast Fourier transform, IEEE Trans. Audio Electroacoust., 17: 93–103, 1969.

15. H.V. Sorensen, D.L. Jones, M.T. Heideman, and C.S. Burrus, Real-valued fast Fourier transform algorithms, IEEE Trans. Acoust. Speech Signal Proc., 35: 849–863, 1987.

16. P.R. Uniyal, Transforming real-valued sequences: fast Fourier versus fast Hartley transform algorithms, IEEE Trans. Signal Proc., 41: 3249–3254, 1994.

17. G. Brunn, z-Transform DFT filters and FFT, IEEE Trans. Acoust. Speech Signal Proc., 26: 56–63, 1978.

18. S.M. Kay, Modern Spectral Estimation, Englewood Cliffs, NJ: Prentice-Hall, 1988.

19. F.J. Harris, On the use of windows for harmonic analysis with the discrete Fourier transform, Proc. IEEE, 66: 51–83, 1978.

20. J.M. Tribolet, A new phase unwrapping algorithm, IEEE Trans. Acoust. Speech Signal Proc., 25: 170–177, 1977.

21. P.D. Welch, The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, IEEE Trans. Audio Electroacoust., 15: 70–73, 1967.

22. R.B. Blackman and J.W. Tukey, The Measurement of Power Spectra, New York: Dover Publications, 1958.

23. T.G. Stockham, Jr., High-speed convolution and correlation, 1966 Spring Joint Computer Conference, AFIPS Proc., 28: 229–233, 1966.

24. R. Braham, Math & visualization: new tools, new frontiers, IEEE Spectrum, 32(11): 19–36, 1995.

25. W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, Numerical Recipes: The Art of Scientific Computing, Cambridge, UK: Cambridge University Press, 1986.

26. D.M. Monro, Complex discrete fast Fourier transform, Appl. Stat., 24: 153–160, 1975.

27. S.L. Marple, Jr., Digital Spectral Analysis with Applications, Englewood Cliffs, NJ: Prentice-Hall, 1987.

28. G.D. Bergland, Fast Fourier transform hardware implementations — an overview, IEEE Trans. Audio Electroacoust., 17: 104–108, 1969.

29. E. Bidet, D. Castelain, C. Joanblanq, and P. Senn, A fast single-chip implementation of 8192complex point FFT, IEEE J. Solid-State Circuits, 30: 300-305, 1995.

30. S.M. Kay and S.L. Marple, Jr., Spectrum analysis — a modern perspective, Proc. IEEE, 69: 1380–1419, 1981.

31. J. Leuridan, H.V. Van der Auweraer, and H. Vold, The analysis of nonstationary dynamic signals, Sound Vibration, 28: 14–26, 1994.

32. L. Cohen, Time-frequency distributions — a review, Proc. IEEE, 77: 941–981, 1989.

33. I. Daubechies, Orthonormal bases of compactly supported wavelets, Commun. Pure Appl. Math., 41: 909–996, 1988.

34. A. Aldroubi and M. Unser, Eds., Wavelets in Medicine and Biology, Boca Raton, FL: CRC Press, 1996.

35. A. Bruce, D. Donoho, and H.Y. Gao, Wavelet analysis, IEEE Spectrum, 33: 26–35, 1996.

## 23.2 RF/Microwave Spectrum Analysis [1]

A. Ambrosini, C. Bortolotti, N. D’Amico, G. Grueff, S. Mariotti,

S. Montebugnoli, A. Orfei, and G. Tomassetti

A signal is usually defined by a time-varying function carrying some sort of information. Such a function

most often represents a time-changing electric or magnetic field, whose propagation can be in free space

or in dielectric materials constrained by conductors (waveguides, coaxial cables, etc.). A signal is said to

be periodic if it repeats itself exactly after a given time T called the period. The inverse of the period T,

measured in seconds, is the frequency f measured in hertz (Hz).

1 All figures have been reproduced courtesy of Hewlett Packard, Rohde Schwarz, Hameg, Tektronix companies,

and IEEE Microwave Measurements.

A periodic signal can always be represented in terms of a sum of several (possibly infinite) sinusoidal

signals, with suitable amplitude and phase, and having frequencies that are integer multiples of the signal

frequency. Assuming an electric signal, the square of the amplitudes of such sinusoidal signals represent

the power in each sinusoid, and is said to be the power spectrum of the signal. These concepts can be

generalized to a nonperiodic signal; in this case, its representation (spectrum) will include a continuous

interval of frequencies, instead of a discrete distribution of integer multiples of the fundamental frequency.

The representation of a signal in terms of its sinusoidal components is called Fourier analysis. The

(complex) function describing the distribution of amplitudes and phases of the sinusoids composing a

signal is called its Fourier transform (FT). The Fourier analysis can be readily generalized to functions

of two or more variables; for instance, the FT of a function of two (spatial) variables is the starting point

of many techniques of image processing. A time-dependent electrical signal can be analyzed directly as

a function of time with an oscilloscope which is said to operate in the time domain. The time evolution

of the signal is then displayed and evaluated on the vertical and horizontal scales of the screen.

The spectrum analyzer is said to operate in the frequency domain because it allows one to measure the

harmonic content of an electric signal, that is, the power of each of its spectral components. In this case

the vertical and horizontal scales read powers and frequencies. The two domains are mathematically well

defined and, through the FT algorithm, it is not too difficult to switch from one response to the other.

Their graphical, easily perceivable representation is shown in Figure 23.6 where the two responses are shown lying on orthogonal planes. It is trivial to say that the easiest way to make a Fourier analysis of a time-dependent signal is to have it displayed on a spectrum analyzer. Many physical processes produce (electric) signals whose nature is not deterministic, but rather stochastic, or random (noise). Such signals can also be analyzed in terms of FT, although in a statistical sense only.

A time signal is said to be band-limited if its FT is nonzero only in a finite interval of frequencies, say $(F_{max} - F_{min}) = B$. Usually, this is the case and an average frequency $F_0$ can be defined. Although the definition is somewhat arbitrary, a (band-limited) signal is referred to as RF (radio frequency) if $F_0$ is in the range 100 kHz to 1 GHz and as a microwave signal in the range 1 to 1000 GHz. The distinction is not fundamental theoretically, but it has very strong practical implications in instrumentation and spectral measuring techniques. A band-limited signal can be described further as narrowband, if $B/F_0 \ll 1$, or wideband otherwise.

The first step in performing a spectral analysis of a narrowband signal is generally the so-called heterodyne downconversion: it consists in the mixing ("beating") of the signal with a pure sinusoidal signal of frequency $F_L$, called local oscillator (LO). In principle, mixing two signals of frequency $F_0$ and $F_L$ in any nonlinear device will result in a signal output containing the original frequencies as well as the difference $(F_0 - F_L)$ and the sum $(F_0 + F_L)$

frequencies, and all their harmonic (multiple) frequencies. In

the practical case, a purely quadratic mixer is used, with an LO frequency $F_L < F_0$; the output will include

the frequencies $(F_0 - F_L)$, $2F_L$, $2F_0$, and $(F_0 + F_L)$, and the first term (called the intermediate frequency

FIGURE 23.6 How the same signal can be displayed.

or IF) will be easily separated from the others, which have a much higher frequency. The bandwidth of

the IF signal will be the same as the original bandwidth B; however, to preserve the original information

fully in the IF signal, stringent limits must be imposed on the LO signal, because any deviation from a

pure sinusoidal law will show up in the IF signal as added phase and amplitude noise, corrupting the

original spectral content. The process of downconverting a (band-limited) signal is generally necessary

to perform spectral analysis in the very high frequency (microwave) region, to convert the signal to a

frequency range more easily handled technically.

When the heterodyne process is applied to a wideband signal (or whenever $F_L > F_{min}$) "negative"

frequencies will appear in the IF signal. This process is called double sideband mixing, because a given

IF bandwidth B (i.e., $(F_L + B/2)$ will include two separate bands of the original signal, centered at $F_L +$

IF ("upper" sideband) and $F_L -$ IF ("lower" sideband). This form of mixing is obviously undesirable in

spectrum analysis, and input filters are generally necessary to split a wideband signal in several narrow

band signals before downconversion. Alternatively, special mixers can be used that can deliver the upper

and lower sidebands to separate IF channels. A band-limited signal in the frequency interval ($F_{max} - F_{min}$) = B is said to be converted to baseband when the LO is placed at $F_L = F_{min}$, so that the band is converted to the interval (B−0). No further lowering of frequency is then possible, unless the signal is split into separate frequency bands by means of filters.

After downconversion, the techniques employed to perform power spectrum analysis vary considerably depending on the frequencies involved. At lower frequencies, it is possible to employ analog-to-digital converters (ADC) to get a discrete numerical representation of the analog signal, and the spectral analysis is then performed numerically, either by direct computation of the FT (generally via the fast Fourier transform, FFT, algorithm) or by computation of the signal autocorrelation function, which is directly related to the square modulus of the FT via the Wiener–Khinchin theorem. Considering that the ADC must sample the signal at least at the Nyquist rate (i.e., at twice the highest frequency present) and with adequate digital resolution, this process is feasible and practical only for frequencies (bandwidths) less than a few megahertz. Also, the possibility of a real-time analysis with high spectral resolution may be limited by the availability of very fast digital electronics and special-purpose computers. The digital approach is the only one that can provide extremely high spectral resolution, up to several hundred thousand channels. For high frequencies, several analog techniques are employed.

## A Practical Approach to Spectrum Analysis [1]

Spectrum analysis is normally done in order to verify the harmonic content of oscillators, transmitters, frequency multipliers, etc. or the spurious components of amplifiers and mixer. Other specialized appli

cations are possible, such as the monitoring of radio frequency interference (RFI), electromagnetic

interference (EMI), and electromagnetic compatibility (EMC). These applications, as a rule, require an

antenna connection and a low-noise, external amplifier. Which are then the specifications to look for in

a good spectrum analyzer? We would suggest:

1. It should display selectable, very wide bands of the EM radio spectrum with power and frequency readable with good accuracy.

2. Its selectivity should range, in discrete steps, from few hertz to megahertz so that sidebands of a selected signal can be spotted and shown with the necessary details.

3. It should possess a very wide dynamic range, so that signals differing in amplitude six to eight orders of magnitude can be observed at the same time on the display.

4. Its sensitivity must be compatible with the measurements to be taken. As already mentioned, specialized applications may require external wideband, low-noise amplifiers and an antenna connection.

5. Stability and reliability are major requests but they are met most of the time.

Occasionally a battery-operated option for portable field applications may be necessary. A block diagram

of a commercial spectrum analyzer is shown in Figure 23.7.

Referring to Figure 23.7 we can say that we are confronted with a radio-receiver-like superhet with a

wideband input circuit. The horizontal scale of the instrument is driven by a ramp generator which is

also applied to the voltage-controlled LO [2].

A problem arises when dealing with a broadband mixing configuration like the one shown above, namely, avoiding receiving the image band.

The problem is successfully tackled here by upconverting the input band to a high-valued IF. An easily

designed input low-pass filter, not shown in the block diagram for simplicity, will now provide the

necessary rejection of the unwanted image band.

Nowadays, with the introduction of YIG bandpass filter preselectors, tunable over very wide input

bands, upconversion is not always necessary. Traces of unwanted signals may, however, show up on the

display although at very low level (less than –80 dBc) on good analyzers.

A block diagram of a commercial spectrum analyzer exploiting both the mentioned principles is shown

in Figure 23.8. This instrument includes a very important feature which greatly improves its performance:

the LO frequency is no longer coming from a free-running source but rather from a synthesized unit

referenced to a very stable quartz oscillator. The improved quality of the LO both in terms of its own

noise and frequency stability, optimizes several specifications of the instrument, such as frequency

determining accuracy, finer resolution on display, and reduced noise in general.

FIGURE 23.7 Block diagram of a commercial spectrum analyzer.

FIGURE 23.8 Standard block diagram of a modern spectrum analyzer.

Further, a stable LO generates stable harmonics which can then be used to widen the input-selected

bands up to the millimeter region. As already stated, this option requires external devices, e.g., a mixer amplifier as shown in Figure 23.9a and b.

The power reference on the screen is the top horizontal line of the reticle. Due to the very wide dynamic

range foreseen, the use of a log scale (e.g., 10 dB/square) seems appropriate. Conventionally, 1 mW is

taken as the zero reference level: accordingly, dBm are used throughout.

The noise power level present on the display without an input signal connected (noise floor) is due

to the input random noise multiplied by the IF amplifier gain. Such a noise is always present and varies

with input frequency, IF selectivity, and analyzer sensitivity (in terms of noise figure).

The “on display dynamic range” of the analyzer is the difference between the maximum compression

free level of the input signal and the noise floor. As a guideline, the dynamic range of a good instrument

could be of the order of 70 to 90 dB.

An input attenuator, always available on the front panel, allows one to apply more power to the analyzer

while avoiding saturation and nonlinear readings. The only drawback is the obvious sensitivity loss. One

should not expect a spectrum analyzer to give absolute power level readings to be better than a couple

of dB.

For the accurate measurement of power levels, the suggestion is to use a power meter. An erratic signal

pattern on display and a fancy level indication may be caused by the wrong setting of the “scan time”

knob. It must be realized that high-resolution observation of a wide input band requires the proper

scanning time. An incorrect parameter setting yields wrong readings but usually an optical alarm is automatically switched on to warn the operator.

The knowledge of the noise floor level allows a good valuation of the noise temperature, T n (and

therefore of the sensitivity), of the analyzer, a useful parameter on many occasions. The relations involved

are as follows.

The Nyquist relation states that

where P = noise floor power level read on the display (W) k = Boltzmann constant = 1.38 ¥ 10 –23 (J/K) B = passband of the selected IF (Hz)

therefore,

Usually engineers prefer to quote the noise figure of receivers. By definition we can write

FIGURE 23.9 Encreasing the input bandwidth characteristics. P kT B= n T P kB n = ( )

where N = noise factor T 0 = 290 K F (noise figure) = 10 log N

A typical F for a good spectrum analyzer is of the order of 30 dB.

It must be said, however, that the “ultimate sensitivity” of the spectrum analyzer will depend not only

on its noise figure but also on the setting of other parameters like the video filter, the IF bandwidth, the

insertion of averaging functions, the scan speed, the detector used, etc.

As a rough estimate a noise floor level of –130/–140 dBm is very frequently met by a good instrument.

Another criterion to select a spectrum analyzer is a good “IMD dynamic range,” that is, the tendency

to create spurious signals by intermodulation due to saturation.

This figure is generally quoted by the manufacturers, but it is also easily checked by the operator by

injecting two equal amplitude sinusoidal signals at the input socket of the analyzer. The frequency

separation between the two should be at least a couple of “resolution bandwidths,” i.e., the selected IF

bandwidth. As the input levels increase, spurious lines appear at the sum and difference frequencies and

spacing of the input signals.

The range in decibels between the nonoverloaded input signals on display and the barely noticeable

spurious lines is known as the “spurious free dynamic range,” shown graphically in Figure 23.10a, where

the third-order “intercept point” is also graphically determined. If input power is increased, higher-order

spurious signals appear, as shown in Figure 23.10b. The input connector of most spectrum analyzers is

of the 50 W coaxial type. Past instruments invariably used N-type connectors because of their good

mechanical and electrical behavior up to quite a few gigahertz. Today SMA or K connectors are preferred.

External millimeter wave amplifiers and converters use waveguide input terminations. As is discussed

in the next section, multipurpose analyzers are available where power meter, frequency counter, tracking

generator, etc. can all be housed in the same cabinet. The economic and practical convenience of these

units must be weighed on a case-by-case basis.

Finally, we mention that spectrum analyzers are available equipped with AM and FM detectors to

facilitate their use in the RFI monitoring applications.

What Is the Right Spectrum Analyzer for My Purpose?

Several manufacturers offer a large number of spectrum analyzer models; the choice may be made on

the basis of application field (i.e., CATV, mobile telephony, service, surveillance, R&D, etc.), performance

(resolution bandwidth, frequency range, accuracy, battery operation etc.), or cost.

In addition, it is important to know that most spectrum analyzers need some accessories generally

not furnished as a standard: for example, a connectorized, coaxial, microwave cable is always required;

a directional coupler, or power divider, or handheld sampler antenna may be very useful to pick up the

signals; and a personal computer is useful to collect, store, reduce, and analyze the data.

There are four main families of RF and microwave spectrum analyzers.

FIGURE 23.10 (a) Spurious free dynamic range. (b) Higher-order spurious. Noise Level Spurious Free Dynamic Range P out IP 3 P in 11 1 1 13 1 211-12 N T T= ( ) + 0 0 1

Family 1

The bench instruments are top performance, but also large, heavy, and the most expensive class, intended

for metrology, certification, factory reference, and for radio surveillance done by government and military

institutions.

The frequency ranges span from a few tens of hertz up to RF (i.e., 2.9 GHz), up to microwave region

(i.e., 26.5 GHz), or up to near millimeter wavelength (i.e., 40 GHz). This class of instruments includes

lower noise figures, approximately 20 dB, and may be decreased down to 10 to 15 dB with an integrated

preamplifier. The synthesized local oscillator has a good phase noise (typically 10 dB better than other

synthesized spectrum analyzers) for precise, accurate, and stable measurement. Also this class of instru

ments, by sharing the display unit, can be integrated with plug-in instruments like a power meter (for

more accurate power measurements) or a tracking generator (for network analysis and mixer testing).

The interface to a computer (and a printer) such IEEE-488 or RS-232 is standard; it allows remote

control and data readings; this class of spectrum analyzer often has a powerful microprocessor, RAM,

and disks for storing data and performing statistical and mathematical analysis.

The best known families are the Hewlett-Packard series, 71xxxx [3] and the Rhode & Schwarz series

FSxx. [4]. Indicative prices are between $50,000 and $90,000.

Family 2

Less expensive bench instruments, the workhorse class of spectrum analyzers, portable and lightweight,

are associated with a synthesized local oscillator, that includes a frequency range from a few kilohertz up

to RF region (i.e., 2.9 GHz), microwave region (i.e., 26.5 GHz), or near millimeter wavelengths (i.e.,

40 to 50 GHz). A typical noise figure of 30 dB is good enough to ensure most measurements. A large

number of filters down to few hertz of resolution are offered; digital filters are preferable to analog ones,

because they give a faster refresh rate of the trace on the display. This kind of spectrum analyzer nearly

always has the capability to extend the frequency range up to millimeter and submillimeter wavelengths

with an external mixer. One of the most important features for a spectrum analyzer in this class is the

quality of the local oscillator; it should be synthesized

(PLL) to achieve stability, precision, accuracy, and

low phase noise. Demodulation is also an important feature to listen to AM, FM on the loudspeaker and

to display TV pictures or complex modulations onto the screen, which is often required by people working

on surveillance, TV, and mobile telephone. The interface to a computer such as IEEE-488 or RS232 is

standard in a large number of spectrum analyzers, and allows the remote control and data reading,

storing, and manipulation.

This kind of instrument may integrate a tracking generator, a frequency counter, and other instruments

that can transform the spectrum analyzer into a compact, full-featured RF and microwave laboratory.

The most popular families are the Hewlett-Packard series 856xx [3, 5], Rhode & Schwarz series

FSExxx [4], Anritsu series MS26x3 [6], IFR mod. AN930 [7], and Marconi Instruments series 239x [9].

The Tektronix production should be taken in account. Prices typically span from $30,000 to $60,000.

Family 3

The entry level, a more economical class of spectrum analyzer, is intended for field use or for one specific

application. If your need is mainly EMI/EMC, CATV, mobile telephone, or surveillance, perhaps you do

not need the extreme stability of a synthesized local oscillator, and a frequency range up to 2 GHz may

be enough; however, if you need some special functions such as “quasi-peak detector” or “occupied

bandwidth measurement,” two functions that are a combination of a mathematical treatment with some

legislative aspects, these are easily measured with a spectrum analyzer including those functions. As the normatives can change, the capability to easily upgrade the measurement software is important; some

models come with a plug-in memory card, some others with 3.5≤ disks.

A large number of spectrum analyzer models are tailored to meet the specific needs of a customer.

This is the case with the HP series 859x [3], Tektronix series 271x [10], IFR series A-xxxx [8], Anritsu

MS2651 [6], and Advantest series U4x4x [4]. Costs typically are around $10, 000 to $20,000.

Family 4

The most economical class of spectrum analyzer, with prices around $2,000 to $6000, includes instru

ments that perform only the basic functions with a limited frequency range and filter availability and

without digital capability. They are intended for service, for general-purpose measurements (i.e., IP 3 ,

harmonic distortion) or for precertification in EMI/EMC measurements. One of the most popular series

is the Hameg series HM50xx [11].

In this class are some special spectrum analyzers that come on a personal computer (PC) board. Such

spectrum analyzers, generally cheap (typically $3,000 to $5,000), with frequency range up to 2 GHz, may

include PLL local oscillators, tracking generators, and other advanced characteristics. The input is through

a coaxial connector on the board, the output and the control is done by a virtual instrument running

on the PC. One model is made by DKD Instruments [12].

Other unusual RF spectrum analyzers working in conjunction with a PC and worth noting are the

instruments for EMI/EMC measurements and reduction in power lines and power cords. For this type

of instrument, the core is not the hardware but the software that performs the measurement according

to international standards and may guide the engineer to meet the required compatibility. An example

is given by Seaward Electronic Sceptre [13].

## Advanced Applications

New technological approaches and the use of spectrum analysis concepts in radioastronomy constitute

some advanced spectrum analysis applications. Autocorrelators, with a typical frequency resolution of

~5/25 kHz, have been extensively used in radioastronomy. Their performance is well documented; the

autocorrelation function is computed online and recorded. Later the FFT of the function is computed

off line in order to get the power spectrum. Recently, the Tektronix 3054 Fourier Analyzer, based on a

bank of programmable filters, was introduced as an alternative approach. The state of the art in integrated

digital signal processors (DSPs) allows an alternative approach to spectrum analysis. By paralleling several

of these DSPs, one is able to compute online the FFT directly on a very wide input bandwidth (several

tens of megahertz).

By using this technique, high time and frequency resolution can be achieved.

A system based on the Sharp LH9124-LH9320 chip set is described in Figure 23.11. It is based on

VME boards: one or two 10-bit, 40-MS/s ADCs and two boards in charge to compute the FFT of the

incoming streams of data, in real time [14]. A following block computes the power and averages on the board up to 64 K spectra before storing the result on disk or tape. The FFT boards are powered by one of the fastest state-of-the-art DSPs (Sharp LH9124). The overall system is controlled by an embedded FORCE 3 Sparcstation. The LH 9124 DSP works with 24+24 bits (with 6 exponent bit) in block floating point. The system architecture allows expansion of the input bandwidth and the number of channels by paralleling more DSP boards. All the computing core is housed in a VME crate and is able to produce single-sided spectra from 1024 frequency bins to 131072 bins at an input bandwidth of 12 MHz without losing data or with 56% of time efficiency at 20 MHz. Single- or double-channel operation mode is provided. In a single-channel mode the main features of the system are reported as

Input bandwidth 0.5–20 MHz

Time efficiency 100% at 12 MHz (56% at 20 MHz)

FFT size 1K, 2K, 256K (points)

Avgs out format <256 averages Æ integer 24 bits >256 averages Æ float 32 bits

Windows Hanning, Hamming, Kaiser Bessel

This spectrometer was developed (1993) as a cost-effective system for both the NASA-SETI (Search for Extraterrestrial Intelligence) program [15,16] and for radioastronomical spectroscopy [17] at the CNR Institute of Radio Astronomy of Bologna. The digital spectrometer was first used to investigate the effects of the Jupiter/SL9 comet impacts (July 1994) [18,19]. In this application, the high time resolution of the spectrometer (a 16K points FFT every 1.3 ms) was exploited to compensate for the fast planet rotational velocity Doppler shift.

The system has been successfully used at the 32 m dish radiotelescope near Bologna in many line

observations with unique results. Note that the use of such a high time and resolution system in radio

astronomy may help to observe the molecular line in a very precise and unusual way. The whole pattern

(a couple of megahertz wide) of a NH 3 molecule line coming from the sky was obtained in flash mode

with a frequency resolution high enough to distinguish the different components. The same machine

can be used for high-time-resolution observations of pulsar and millisecond pulsar. In those cases, the

possibility of performing the FFT of the RF signal online allows coherent dedispersion of the pulses. This

new technological approach in computing the FFT may be successfully addressed to many different fields,

such as image processing, medical diagnostic systems, radio surveillance, etc.

1. A.E. Bailey, Ed., Microwave Measurements, 2nd ed., Peter Peregrins on behalf of IEEE, London, 1989.

2. Hewlett-Packard, A.N 63 and A.N 243, Hewlett-Packard Company.

3. Hewlett-Packard, Test & Measurement Catalog 1997, pp. 224–257, Hewlett-Packard Company.

4. Rohde & Schwarz, Catalog 96/97 Test & Measurement Products pp. 109–145, Rohde & Schwarz Company.

5. Hewlett-Packard, HP 8560E, HP8561E and HP8563E Portable Spectrum Analizers Technical Data, Hewlett-Packard Company, pp. 5091–3274E.

6. Anritsu-Wiltron, Electronic Measuring Instruments, pp. 67–100, Anritsu-Wiltron Company.

7. IFR, The AN 930 Microwave Spectrum Analyzer, IFR System, Inc.

8. IFR, The A-8000 & A-7500 Spectrum Analyzer, IFR System, Inc.

9. Marconi Instruments, Test and Measurement 96/97 Edition Instrument System, pp. 4-1–4-5, Marconi Instruments Company.

10. Tektronix, Measurement Products Catalog 1997/1998, pp. 145–171, Tektronix Company.

11. Hameg Instruments, Catalog 1994, Hameg GmbH.

12. DKD Instruments, Model 1800 Data Sheet, DKD Instruments.

13. C. Galli, Analizzatori di spettro: rassegna commerciale, Sel. Elettron., 15, 1996 [in Italian].

FIGURE 23.11 Block diagram of spectrum analyzer for radioastronomy.

14. S. Montebugnoli et al., “A new 6 MHz 128000 channels spectrum analyzer,” IAA-95-IAA.9.1.10. 46th International Astronautical Congress, October, Oslo, Norway, 1995.

15. S. Montebugnoli, C. Bortolotti, S. Buttaccio, A. Cattani, N. D’Amico, G. Grueff, A. Maccaferri, G. Maccaferri, A. Orfei, M. Roma, G. Tuccari, and M. Tugnoli, “Upgrade of the mini spectrum analyzer” in 5th Proceedings of the International Conference on Bioastronomy, Capri, Vol. I, July 1996.

16. A. Orfei, S. Montebugnoli, C. Miani, J. Monari, and G. Tuccari, “The SETI facilities at the Medicina/ Noto sites,” (preprint) IAA-96-IAA.9.1.11, 47th International Astronautical Congress, October, Beijing, China, 1996.

17. S. Montebugnoli, C. Bortolotti, S. Buttaccio, A. Cattani, N. D’Amico, G. Grueff, A. Maccaferri, G. Maccaferri, A. Orfei, M. Roma, G. Tuccari, and M. Tugnoli, “A new high-resolution digital spectrometer for radioastronomy applications,” Rev. Sci. Instrum. 67, 365–370, 1996.

18. C.B. Cosmovici et al., “Detection of the 22 GHz line of water during and after the SL-9/Jupiter event,” European SL-9/Jupiter Workshop–February, ESO Headquarters, Munchen, Germany, 1995.

19. C.B. Cosmovici, S. Montebugnoli, A. Orfei, S.

Pogrebenko, and P. Colom, “First evidence of planetary water MASER emission induced by the comet/Jupiter catastrophic impact,” Planetary Space Sci., 44, 735-773, 1996.

1. Baker, D., Kohler, D.C., Fleckenstein, W.O., Roden, C.E., and Sabia, R., Eds., Physical Design of Electronic Systems, Vol. 4, Prentice-Hall, Englewood Cliffs, NJ, 1972.

2. Jain, A.K. and Mao, J., Artificial neural networks: a tutorial, IEEE Comput., pp. 31–44, 1996.

3. Kaufmann, A., The Science of Decision Making, Weidenfeld and Nicholson, London, 1968.

4. Mauris, G., Benoit, E., and Foulloy, L., Fuzzy sensors for the fusion of information, in Proc. 13th IMEKO Conference, Turin, pp. 1009–1014, 1994.

5. Nicholas, J., in Handbook of Measurement and Control, Vol. 3, Sydenham, P.H. and Thorn, R., Eds., John Wiley and Sons, Chichester, U.K., 1992.

6. Sizemore, N.L., Test techniques for knowledge-based systems, ITEA J., XI (2), 34–43, 1990.

7. Tang, K.S., Man, K.F., Kwong, S., and He, Q., Genetic algorithms and their applications, IEEE Signal Processing Mag., 13(6), 22–37, 1996.

8. Venmuri, Artificial Neural Networks: Theoretical Concepts, IEEE Computer Society Press, Los Angeles, CA, 1988.

# 25 25. Analog-to-Digital Converters


1. C.E. Shannon, Communication in the presence of noise, Proc. IRE, 37(1): 10–21, 1949.

2. H. Nyquist, Certain topics in telegraph transmission theory, AIEE Transactions, 617–644, April, 1928.

3. A.B. Williams and F.J. Taylor, Electronic Filter Design Handbook: LC, Active, and Digital Filters, 2nd ed., New York: McGraw-Hill, 1988.

4. S.P. Lipshitz, R.A. Wannamaker, and J. Vanderkooy, Quantization and dither: a theoretical survey, J. Audio Eng. Soc., 40, 355–375, 1992.

5. H.W. Ott, Noise Reduction Techniques in Electronic Systems, 2nd ed., New York: John Wiley & Sons, 1988.

6. R. Morrison, Grounding and Shielding Techniques in Instrumentation, 3rd ed., New York: John Wiley & Sons, 1986.

7. J. Williams, Designs for High Performance Voltage-to-Frequency Converters, Application Note 14, Linear Technology Corporation, March 1986.


Further Information


M. Demler, High-Speed Analog-to-Digital Conversion, San Diego, CA: Academic Press, 1991.

B.M. Gordon, The Analogic Data-Conversion Systems Digest, Wakefield, MA: Analogic Corporation, 1981.

D.H. Sheingold, Ed., Analog-Digital Conversion Handbook, Englewood Cliffs, NJ: Prentice-Hall, 1986.

Crystal Semiconductor, Delta Sigma A/D Conversion Technique Overview, Application Note AN10, Austin, TX, 1989.

TABLE 25.4 Companies That Manufacture Plug-In Data Acquisition Boards Analogic Corporation Keithley Metrabyte 360 Audubon Road 440 Myles Standish Blvd. Wakefield, MA 01880 Taunton MA 02780 (508) 977-3000 (508) 880-3000 http://www.metrabyte.com ComputerBoards, Inc. 125 High Street National Instruments Corporation Mansfield, MA 02048 6504 Bridge Point Parkway (508) 261-1123 Austin, TX 78730 (512) 794-0100 Data Translation, Inc. http://www.natinst.com 100 Locke Drive Marlboro, MA 01752-1192 (508) 481-3700 http://www.datx.com

# 26 26. Computers


1. Micrology pbt, Inc., The VMEbus Specification Manual: Revision C1, PRINTEX, 1985.

2. S. Heath, VMEbus: A Practical Companion, Boston, MA: Butterworth Heinemann, 1993.

3. J. Di Giacomo, Ed., Digital Bus Handbook, New York: McGraw-Hill, 1990.

4. T. Shanley and D. Anderson, ISA System Architecture, 3rd ed., Reading, MA: Addison-Wesley, 1995.

5. T. Shanley and D. Anderson, EISA System Architecture, 2nd ed., Reading, MA: Addison-Wesley, 1995.

6. T. Shanley and D. Anderson, PCMCIA System Architecture, 2nd ed., Reading, MA: Addison-Wesley, 1995.

7. M.D. Seyer, RS232 Made Easy, Englewood Cliffs, NJ: Prentice-Hall, 1984.

8. Intel Corporation, Intel Microsystems Component Handbook, Vol. 2.

9. IEEE Std 488.1-1987, IEEE Standard Digital Interface for Programmable Instrumentation.

10. IEEE Std 488.2-1987, IEEE Standard Codes, Formats, Protocols and Common Commands.

# 27 27. Telemetry


1. B.P. Dagarin, R.K. Taenaka, and E.J. Stofel, Galileo probe battery system, IEEE Aerospace Electron. Syst. Mag., 11 (6): 6–13, 1996.

2. M.W. Pollack, Communications-based signaling: advanced capability for mainline railroads, IEEE Aerospace Electron. Syst. Mag., 11 (11): 13–18, 1996.

3. M.C. Shults, R.K. Rhodes, S.J. Updike, B.J. Gilligan, and W.N. Reining, A telemetry-instrumentation system for monitoring multiple subcutaneously implanted glucose sensors, IEEE Trans. Biomed. Eng., 41: 937–942, 1994.

4. M. Rezazadeh and N.E. Evans, Multichannel physiological monitor plus simultaneous full-duplex speech channel using a dial-up telephone line, IEEE Trans. Biomed. Eng., 37: 428–432, 1990.

5. Y.S. Trisno, P. Hsieh, and D. Wobschall, Optical pulses powered signal telemetry system for sensor network application, IEEE Trans. Instrum. Meas., 39: 225–229, 1990.

6. R. Pallás-Areny and J.G. Webster, Sensors and Signal Conditioning, 2nd ed., New York: John Wiley & Sons, 2001, 352–379.

7. Part 90: Private land mobile radio services. Subpart J: Non voice and other specialized operations (Sec. 90.238: Telemetry), Code of Federal Regulations, Title 47, Telecommunication. Chapter I: Federal Communications Commission.

8. Part 15: Radio frequency devices, Code of Federal Regulations, Title 47, Telecommunication. Chapter I: Federal Communications Commission.

9. H. Taub and D. L. Schilling, Principles of Communication Systems, 2nd ed., New York: McGrawHill, 1986.

10. IRIG, Telemetry Standard IRIG 106-96, Range Commander Council, U.S. Army White Sands Missile Range, 1996.

11. K. M. Uglow, Noise and bandwidth in FM/FM radio telemetry, IRE Trans. Telemetry Remote Control, 19–22, 1957.

12. J.L. Maury and J. Styles, Development of optimum frame synchronization codes for Goddard space flight center PCM

telemetry standards, Proc. Natl. Telemetering Conf., 1964.

Further Information

E. H. Higman, Pneumatic instrumentation, in B.E. Noltingk, Ed., Instrument Technology: Instrumentation Systems, Vol. 4, London: Butterworths, 1987.

C.H. Hoeppner, Telemetry, in R.C. Dorf, Ed., The Electrical Engineering Handbook, Boca Raton, FL: CRC Press, 1993.

R.S. Mackay, Bio-Medical Telemetry, 2nd ed., Piscataway, NJ: IEEE Press, 1993.

Telemetry Group, Test Methods for Telemetry Systems and Subsystems. Document 118-97, Range Commanders Council, U.S. Army White Sands Missile Range, 1997.

Telemetry Group, Telemetry Applications Handbook Document 119-88, Range Commanders Council, U.S. Army White Sands Missile Range, 1988.

# 28 28. Sensor Networks and Communication


1. ISO/IEC 7498-1:1994 Information Technology — Open Systems Interconnection — Basic Reference Model: The Basic Model, International Organization for Standardization (ISO), 1, rue de Varembé, Case postale 56, CH-1211 Genève 20, Switzerland, [online]. Available http://www.iso.ch/.

2. 8802-3: 1996 (ISO/IEC) [ANSI/IEEE Std 802.3, 1996 Edition] Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements — Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications, Institute of Electrical and Electronics Engineers, 445 Hoes Lane, Piscataway, NJ 08855-1331, [online]. Available http://www.ieee.org/.

3. ANSI/EIA/TIA-232-E-91, Electronic Industries Association, 2500 Wilson Boulevard, Arlington, VA 22201-3834, [online]. Available http://www.eia.org/.

4. Distributed, Intelligent I/O for Industrial Control and Data Acquisition… The SERIPLEX Control Bus, Bulletin No. 8310PD9501R4/97, Seriplex Technical Organization, P.O. Box 27446, Raleigh, NC 27611-7446, [online]. Available http://www.seriplex.org/.

5. AS-Interface U.S.A., 5010 East Shea Blvd., Suite C-226, Scottsdale, AZ 85254, [online]. Available http://www.as-interface.com/.

6. Interbus-S Protocol Structure, Data Sheet 0005C, Phoenix Contact, P.O. Box 4100, Harrisburg, PA 17111, [online]. Available http://www.ibsclub.com/.

7. CAN Specification, Version 2.0, 1991, Robert Bosch GmbH, Postfach 50, D-7000 Stuttgart 1, Germany.

8. ISO 11898:1993 Road vehicles—Interchange of digital information—Controller area network (CAN) for high-speed communication, International Organization for Standardization (ISO), 1, rue de Varembé, Case postale 56, CH-1211 Genève 20, Switzerland, [online]. Available http://www.iso.ch/.

9. CiA Draft Standard 301 (Version 3.0), CANopen Communication Profile for Industrial Systems, CiA Headquarters, Am Weichselgarten 26, D-91058 Erlangen, Germany, [online]. Available http:// www.can-cia.de/.

10. SAE J 1939 — Recommended Practice for Serial Control and Communications Vehicle Network, Society of Automotive Engineers, 400 Commonwealth Drive, Warrendale, PA 15096, [online]. Available http://www.sae.org/.

11. DeviceNet Specification v1.4, Open DeviceNet Vendors Association, 8222 Wiles Road, Suite 287, Coral Springs, FL 33067, [online]. Available http://www.odva.org/.

12. Smart Distributed System Application Layer Protocol Specification v2.0, 1996, Honeywell MICRO SWITCH Division, 11 West Spring Street, Freeport, IL 61032, [online]. Available http://www.sensing. honeywell.com/sds/.

13. HART Communication Foundation, 9390 Research Boulevard, Suite I-350, Austin, TX 78759 [online]. Available http://www.ccsi.com/hart/hcfmain.html.

14. EN 50 170 — Volume 2, CENELEC Central Secretariat, 35, rue de Stassart, B-1050 Brussels, Belgium.

15. PROFIBUS Trade Organization U.S.A., 5010 East Shea Blvd., Suite C-226, Scottsdale, AZ 852544683, [online]. Available http://www.profibus.com/.

16. ISA, the International Society for Measurement & Control, P.O. Box 12277, Research Triangle Park, NC 27709, [online]. Available http://www.isa.org/.

17. IEC 61158-2(1993-12), Fieldbus standard for use in industrial control systems — Part 2: Physical layer specification and service definition, International Electrotechnical Commission, 3, rue de Varembé, PO Box 131, 1211 Geneva 20, Switzerland, [online]. Available http://www.iec.ch/.

18. WorldFIP Headquarters, 2, rue de Bone, 92160 Antony, France, [online]. Available http:// www.world-fip.org/.

19. 005-0017-01 Rev C, LonTalk Protocol, Echelon Corporation, 4015 Miranda Avenue, Palo Alto, CA 94304, [online]. Available http://www.echelon.com/.

20. P1451.2, Draft Standard for a Smart Transducer Interface for Sensors and Actuators — Transducer to Microprocessor Communication Protocols and Transducer Electronic Data Sheet (TEDS) Formats, Institute of Electrical and Electronics Engineers, 445 Hoes Lane Piscataway, NJ 08855-1331, [online]. Available

http://www.ieee.org/.

## Further Information

P.Z. Peebles, Jr., Digital Communications Systems, Englewood Cliffs, NJ: Prentice-Hall, 1987, provides a good general text on communication.

B. Svacina, Understanding Device Level Buses, Minneapolis, MN: TURCK Inc. (3000 Campus Dr., Minneapolis, MN 55441), 1996, is an in-depth study of the subject of communication networks for industrial field devices.

J.D. Gibson, Ed., The Communications Handbook, Boca Raton, FL: CRC Press, 1997, includes recent material on communication techniques.

# 29 29. Electromagnetic Compatibility

1. E.C. Jordan and K.G. Balmain, Electromagnetic Waves and Radiating Systems, 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 1968, 540–547.

2. J.P. Mills, Electromagnetic Interference Reduction in Electronic Systems, Englewood Cliffs, NJ: Prentice-Hall, 1993, 232–233.

3. M.L. Crawford, Generation of standard EM fields using TEM transmission cells, IEEE Trans. Electromagn. Compat., EMC-16, 189–195, 1974.

4. M.L. Crawford and J.L. Workman, Predicting free-space radiated emissions from electronic equipment using TEM cell and open-field site measurements, IEEE Int. Symp. Electromagn. Compat., Baltimore, MD, Oct 7–9, IEEE, Piscataway, NJ, 1980, 80–85.

5. Part 15 — Radio Frequency Devices, in Code of Federal Regulations, Volume 47: Telecommunications, Washington, D.C.: U.S. Government Printing Office, 1991, subpt. A-B.

6. Methods of measurement of radio-noise emissions from low-voltage electrical and electronic equipment (ANSI C63.4-1988), American National Standards Institute (ANSI) Standards, New York: The Institute of Electrical and Electronics Engineers, 1989.

Further Information

Guide for construction of open area test sites for performing radiated emission measurements (ANSI C63.7-1988), American National Standards Institute (ANSI) Standards, New York: The Institute of Electrical and Electronics Engineers, 1988.

H.W. Ott, Noise Reduction Techniques in Electronic Systems, 2nd ed., New York: John Wiley & Sons, 1988.

W.H. Hayt, Engineering Electromagnetics, 5th ed., New York: McGraw-Hill, 1989.

J.D. Kraus, Electromagnetics, 4th ed., New York, NY: McGraw-Hill, 1992.

S. Ramo, J.R. Whinnery, and T. Van Duzer, Fields and Waves in Communication Electronics, 2nd ed., New York: John Wiley & Sons, 1984. III -1 III

Displays 30 Human Factors in

1. S. Sherr, Electronic Displays, 2nd ed. New York: John Wiley & Sons, 1993.

2. M.S. Sanders and E.J. McCormick, Human Factors in Engineering and Design, 6th ed. New York: McGraw-Hill, 1987.

3. E. Grandjean, Design of VDT workstations, in G. Salvendy, Ed., Handbook of Human Factors, New York: John Wiley & Sons, 1987.

4. D.G. Payne, V.A. Lang, and J.M. Blackwell, Mixed vs. pure display format in integration and nonintegration visual display monitoring tasks, Human Factors, 37(3), 507–527, 1995.

5. N.H. Lehrer, The challenge of the cathode-ray tube, in L.E. Tannas, Jr., Ed., Flat-Panel Displays and CRTs, New York: Van Nostrand Reinhold, 1985.

6. M. Hartney, ARPA display program and the national flat-panel display initiative, in Proc. 2nd Int. Workshop on Active Matrix Liquid Crystal Displays, IEEE, 8–15, 1995.

7. T. Morita, An overview of active matrix LCDs in business and technology, in Proc. 2nd Int. Workshop on Active Matrix Liquid Crystal Displays, IEEE, 1–7, 1995.

8. I-W. Wu, High-definition displays and technology trends in TFT-LCDs, J. Soc. Inf. Display, 2(1), 1–14, 1991.

9. L. Weber, Plasma displays, in L.E. Tannas, Jr., Ed., Flat-Panel Displays and CRTs, New York: Van Nostrand Reinhold, 1985.

10. Y. Amano, A new hybrid ac–dc plasma display panel, J. Soc. Inf. Display, 2(1), 57–58, 1994.

Further Information

E. Grandjean, Ergonomics in Computerized Offices, London: Taylor & Francis, 1987, is an excellent allaround treatise on the principles of effective VDT selection and use. Summarizes a wide range of research literature. If this volume is difficult to obtain, a chapter by the same author is also included in the Handbook of Human Factors (Reference 2).

M.G. Helander, Design of visual displays, in G. Salvendy, Ed., Handbook of Human Factors, New York: John Wiley & Sons, 1987, is an excellent and concise review of major human factors principles for display design and use. Includes a critical review of the foundation literature in this area.

S. Sherr, Electronic Displays, 2nd ed., New York: John Wiley & Sons, 1993, offers clear presentations of all important display technologies, together with a good summary of performance measurement methods for display systems. Well illustrated with a variety of commercial products.

L.E. Tannas, Jr., Ed., Flat-Panel Displays and CRTs, New York: Van Nostrand Reinhold, 1985, provides a thorough, yet highly readable examination of the physical principles behind essentially every major display technology. Although the technology capabilities have become dated since publication, this is well worth review.

# 31 31. Cathode Ray Tube Displays


1. G. Aboud, Cathode Ray Tubes, 1997, 2nd ed., San Jose, CA, Stanford Resources, 1997.

2. G. Aboud, Cathode Ray Tubes, 1997, Internet excerpts, available http://www.stanfordresources.com/ sr/crt/crt.html, Stanford Resources, February 1998.

3. G. Shires, Ferdinand Braun and the Cathode Ray Tube, Sci. Am., 230 (3): 92–101, March 1974.

4. N.H. Lehrer, The challenge of the cathode-ray tube, in L.E. Tannas, Jr., Ed., Flat Panel Displays and CRTs, New York: Van Nostrand Reinhold, 1985.

5. P. Keller, The Cathode-Ray Tube, Technology, History, and Applications, New York: Palisades Press, 1991.

6. D.C. Ketchum, CRTs: the continuing evolution, Society for Information Display International Symposium, Conference Seminar M-3, 1996.

7. L.R. Falce, CRT dispenser cathodes using molybdenum rhenium emitter surfaces, Society for Information Display International Symposium Digest of Technical Papers, 23: 331–333, 1992.

8. J.H. Lee, J.I. Jang, B.D. Ko, G.Y. Jung, W.H. Kim, K. Takechi, and H. Nakanishi, Dispenser cathodes for HDTV, Society for Information Display International Symposium Digest of Technical Papers, 27: 445–448, 1996.

9. T. Nakadaira, T. Kodama, Y. Hara, and M. Santoku, Temperature and cutoff stabilization of impregnated cathodes, Society for Information Display International Symposium Digest of Technical Papers, 27: 811–814, 1996.

10. W. Kohl, Materials Technology for Electron Tubes, New York, Reinhold Publishing, 1951.

11. S. Sugawara, J. Kimiya, E. Kamohara, and K. Fukuda, A new dynamic-focus electron gun for color CRTs with tri-quadrupole electron lens, Society for Information Display International Symposium Digest of Technical Papers, 26: 103–106, 1995.

12. J. Kimiya, S. Sugawara, T. Hasegawa, and H. Mori, A 22.5 mm neck color CRT electron gun with simplified dynamically activated quadrupole lens, Society for

Information Display International Symposium Digest of Technical Papers, 27: 795–798, 1996.

13. D. Imabayashi, M. Santoku, and J. Karasawa, New pre-focus system structure for the trinitron gun, Society for Information Display International Symposium Digest of Technical Papers, 27: 807–810, 1996.

14. K. Kato, T. Sase, K. Sasaki, and M. Chiba, A high-resolution CRT monitor using built-in ultrasonic motors for focus adjustment, Society for Information Display International Symposium Digest of Technical Papers, 27: 63–66, 1996.

15. S. Sherr, Electronic Displays, 2nd ed., New York: John Wiley, 1993.

16. N. Azzi and O. Masson, Design of an NIS pin/coma-free 108° self-converging yoke for CRTs with super-flat faceplates, Society for Information Display International Symposium Digest of Technical Papers, 26: 183–186, 1995.

17. J.F. Fisher and R.G. Clapp, Waveforms and spectra of composite video signals, in K. Benson and J. Whitaker, Television Engineering Handbook, Featuring HDTV Systems, New York: McGraw-Hill Reinhold, 1992.

18. D. Pritchard, Standards and recommended practices, in K. Benson and J. Whitaker, Television Engineering Handbook, Featuring HDTV Systems, New York: McGraw-Hill Reinhold, 1992.

19. A. Vecht, Phosphors for color emissive displays, Society for Information Display International Symposium Conference Seminar Notes F-2, 1995.

20. Optical Characteristics of Cathode Ray Tube Screens, EIA publication TEP116-C, Feb., 1993.

21. G. Wyszecki and W.S. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd ed., New York: John Wiley & Sons, 1982.

22. A. Robertson and J. Fisher, Color vision, representation, and reproduction, in K. Benson and J. Whitaker, Television Engineering Handbook, Featuring HDTV Systems, New York: McGraw-Hill Reinhold, 1992.

23. M. Maeda, Trinitron technology: current status and future trends, Society for Information Display

International Symposium Digest of Technical Papers, 27: 867–870, 1996.

24. C. Sherman, Field sequential color takes another step, Inf. Display, 11 (3): 12–15, March, 1995.

25. L. Ozawa, Helmet mounted 0.5 in. crt for SVGA images, Society for Information Display International Symposium Digest of Technical Papers, 26: 95–98, 1995.

26. C. Infante, CRT display measurements and quality, Society for Information Display International Symposium Conference Seminar Notes M-3, 1995.

27. J. Whitaker, Electronic Displays, Technology, Design, and Applications, New York: McGraw-Hill, 1994.

28. P. Keller, Electronic Display Measurement, Concepts, Techniques, and Instrumentation, New York: John Wiley & Sons, 1997.

Further Information

L. Ozawa, Cathodoluminescence: Theory and Applications, New York: Kodansha, 1990.

V.K. Zworykin and G.A. Morton, Television: The Electronics of Image Transmission in Color and Monochrome, New York: John Wiley & Sons, 1954.

B. Wandell, The foundations of color measurement and color perception, Society for Information Display International Symposium, Conference Seminar M-1, 1993. A nice brief introduction to color science (31 pages).

Electronic Industries Association (EIA), 2500 Wilson Blvd., Arlington, VA 22201 (Internet: www.eia.org). The Electronic Industries Association maintains a collection of over 1000 current engineering publications and standards. The EIA is an excellent source for information on CRT engineering, standards, phosphors, safety, market information, and electronics in general.

The Society for Information Display (SID), 1526 Brookhollow Dr., Suite 82, Santa Ana, CA 92705-5421 (Internet: www.display.org). The Society for Information Display is a good source of engineering research and development information on CRTs and information display technology in general.

## Internet Resources

The following is a brief list of places to begin looking on the World Wide Web for information on CRTs

and displays, standards, metrics, and current research. Also many of the manufacturers listed in Table 31.3

maintain Web sites with useful information.

The Society for Information Display www.display.org

The Society of Motion Picture and Television Engineers www.smpte.org

The Institute of Electrical and Electronics Engineers www.ieee.org

The Electronic Industries Association www.eia.org

National Information Display Laboratory www.nta.org

The International Society for Optical Engineering www.spie.org

The Optical Society of America www.osa.org

Electronics & Electrical Engineering Laboratory www.eeel.nist.gov

National Institute of Standards and Technology (NIST) www.nist.gov

The Federal Communications Commission www.fcc.gov

# 32 32. Liquid Crystal Displays


1. P.J. Collings, Liquid Crystals: Nature's Delicate Phase Matter, Princeton University Press, Princeton, NJ, 1990.

2. L.M. Blinov and V.G. Chigrinov, Electro-optical Effects in Liquid Crystal Materials, Springer, New York, 1996.

3. P.G. De Gennes and J. Prost, The Physics of Liquid Crystals, 2nd ed., Clarendon Press, Oxford, 1993.

4. B. Bahadur, Ed., Liquid Crystals: Applications and Uses, World Scientific, Singapore, 1990.

5. E. Kaneko, Liquid Crystal TV Displays: Principles and Applications of Liquid Crystal Displays, D. Reidel Publishing, Boston, 1987.

6. T. Tsukada, TFT/LCD: Liquid Crystal Display Addressed by Thin Film Transistors, Gordon and Breach, Canada, 1996.

7. E. Lueder, Fundamentals of passive and active addressed liquid crystal displays, Short Course S-1, SID Conference, San Diego, CA, May 12, 1996.

8. P.W. Ross, L.K.M. Chan, and P.W. Surguy, Ferroelectric LCD: simplicity and versatility, SID '94 Digest, 147–150, 1996.

9. J.W. Doane, D.K. Yang, and Z. Yaniv, Front-lit flat panel display for polymer stabilized cholesteric textures, Proc. 12th International Display Conference, Japan Display '92, Oct. 12–14, 73–76, 1992.

10. J.C. Rowell, W.D. Phillips, L.R. Melby, and M. Panar, NMR studies of some liquid crystal systems, J. Chem. Phys., 43, 3442–3454, 1965.

11. L. Pohl and E. Merck, Physical properties of liquid crystals, in Liquid Crystals: Applications and Uses, B. Bahadur, Ed., World Scientific, Singapore, 1990, 139–170.

12. P.R. Kishore, N.V.S. Rao, P.B.K. Sarma, T.F.S. Raj, M.N. Avadhanlu, and C.R.K. Murty, Field and frequency effects in nematic mixtures of negative and positive dielectric anisotropy, Mol. Cryst. Liq. Cryst., 45, 3/4, 231–241, 1978.

13. D. Lippens, J.P. Parneix, and A. Chapoton, Study of 4-heptyl 4Φ-cyanobiphenyl using the analysis of its

dielectric properties, J. Phys., 38, 1465, 1977.

14. F.C. Frank, On the theory of liquid crystals, Discuss. Faraday Soc., 25, 19, 1958.

15. W.H. de Jeu, Physical Properties of Liquid Crystalline Materials, Gordon and Breach, New York, 1980.

16. J.J. Wysocki, A. Adams, and W. Haas, Electric-field-induced phase change in cholesteric liquid crystals, Phys. Rev. Lett., 20, 1024, 1968.

17. G.H. Heilmeier, L.A. Zanoni, and L.A. Barton, Dynamic scattering: a new electo-optic effect in certain classes of nematic liquid crystals, Proc. IEEE, 56, 1162, 1968.

18. G.H. Heilmeier and L.A. Zanoni, Guest-host interaction in nematic liquid cyrstal — a new electrooptic effect, Appl. Lett., 13, 91, 1968.

19. M. Schadt and W. Helfrich, Voltage-dependent optical activity of a twisted nematic liquid crystal, Appl. Phys. Lett., 18, 127, 1971.

20. M.F. Schiekel and K. Fahrenschon, Deformation of nematic liquid crystal with vertical orientation in electric fields, Appl. Phys. Lett., 19, 391, 1971.

21. M. Hareng, G. Assouline, and E. Leiba, Liquid crystal matrix display by electrically controlled birefringence, Proc. IEEE, 60, 913, 1972.

22. C.H. Gooch and H.A. Tarry, The optical properties of twisted nematic liquid crystal structures with twist angles £ = 90°, J. Phys. D Appl. Phys., 8, 1575–1584, 1975.

23. E. Jakeman and E.P. Raynes, Electro-optical response times in liquid crystals, Phys. Lett., 39A, 69–70, 972.

24. P.M. Alt and P. Pleshko, Scanning limitations of liquid-crystal displays, IEEE Trans. Electron Dev., ED-21, 146, 1974.

25. T.J. Scheffer and J. Nerring, A new highly multiplexable liquid crystal display, Appl. Phys. Lett., 45, 1021, 1984.

26. T.J. Scheffer, Super Twisted Nematic (STN) LCDs, SID ‘95 Seminar Notes, Vol. I, M-2, 1995.

27. M. Oh-e, M. Ohta, S. Aratani, and K. Kondo, Principles and characteristics of electro-optical behavior of in-plane switching mode, in Proc. 15th International Display Research Conf., Asia Display ’95, 577–580, 1995.

28. T. Scheffer and B. Clifton, Active addressing method for high-contrast video-rate STN displays, SID’92 Digest, 228, 1992.

29. H. Muraji et al., A 9.4-in. color VGA F-STN display with fast response time and high contrast ratio by using MLS method, SID ’94 Dig., 61, 1994.

30. Higashi et al., A 1.8-in poly-Si TFT-LCD for HDTV projectors with 5V fully integrated driver, SID ’95 Dig., 81, 1995.

31. Nano et al., Characterization of sticking effects in TFT-LCD, SID ’90 Dig., 404, 1990.

# 33 33. Plasma-Driven Flat Panel Displays


1. D.L. Bitzer and H.G. Slottow, The plasma display panel — a digitally addressable display with inherent memory, 1966 Fall Joint Computer Conf., Washington, D.C., AFIPS Conf. Proc., 29, 541, 1966.

2. H.G. Slottow, Plasma displays, IEEE Trans. Electron. Devices, 23, 760–772, 1976. TABLE 33.4 Advanced Techniques for Noncontact Profiling/Chemical Characterization of Thin Films Physical Principle b-Backscatter X-Ray Microfluorescence Maximum thickness/step height A few mm 50 mm Vertical resolution About 0.1 nm 0.1 mm Spatial resolution Depends on pinhole size 50 mm–2 mm Spot diameter/FOV Depends on pinhole size 100 mm Scan rate <1 min/measurement 1 spot/min Maximum sample size No restriction 24 cm ¥ 21.5 cm Maximum sample weight No restriction 4.5 kg Elemental detection N.A. <10–300 ppm

3. T.N. Criscimagna and P. Pleshko, Ac plasma display. chap. 3, in Topics in Applied Physics, Vol. 40, Display Devices, Springer-Verlag, Berlin, 91–150, 1980.

4. L.F. Weber, Plasma displays, chap. 10, in Flat Panel Displays and CRTs, L.E. Tannas, Jr., Ed., Van Nostrand Reinhold, New York, 1985.

5. J. Nolan, Gas discharge display panel, 1969 International Electron Devices Meeting, Washington, D.C., 1969.

6. H.J. Hoehn and R.A. Martel, A 60 line per inch plasma display panel, IEEE Trans. Electron Devices, 18, 659–663, 1971.

7. O. Sahni and W.P. Jones, Spatial distribution of wall charge density in ac plasma display panels, IEEE Trans. Electron Devices, 25, 223–226, 1979.

8. O. Sahni and C. Lanza, Origin of the bistable voltage margin in the ac plasma display panel, IEEE Trans. Electron. Devices, 24, 853–859, 1977.

9. C. Lanza and O. Sahni, Numerical calculation of the characteristics of an isolated ac gas discharge display panel cell, IBM J. Res. Dev., 22, 641–646, 1978.

10. O. Sahni and M.O. Aboelfotoh, The pressure dependence of the bistable voltage margin of an ac plasma panel cell, Proc. SID, 22, 212–218, 1981.

11. O. Sahni and C. Lanza, Importance of the secondary electron emission coefficient on E/Po for Paschen breakdown curves in ac plasma panels, J. Appl. Phys., 47, 1337, 1976.

12. O. Sahni and C. Lanza, Influence of the secondary electron emission coefficient of argon on Paschen breakdown curves in ac plasma panels for neon + 0.1% argon mixture, J. Appl. Phys., 47, 5107, 1976.

13. C. Lanza, Analysis of an ac gas display panel, IBM J. Res. Dev., 18, 232–243, 1974.

14. O. Sahni, C. Lanza, and W.E. Howard, One-dimensional numerical simulation of ac discharges in a high-pressure mixture of Ne + 0.1% Ar confined to a narrow gap between insulated metal electrodes, J. Appl. Phys., 49, 2365, 1978.

15. O. Sahni and C. Lanza, Failure of Paschen's scaling law for Ne –0.1% Ar mixtures at high pressures, J. Appl. Phys., 52, 196, 1981.

16. H. Wakabayshi, Display market projections report from Namora Research Institute, paper presented at Imaging 2001: The U.S. Display Consortium Business Conference, January 28, 1997, San Jose, CA.

17. K. Werner, Plasma hits the ground running, Inf. Display, 12(12), 30–34, 1996.

18. Nikkei Materials & Technology, [135], November, 1993, pp. 22–23.

19. H. Murakami and R. Toyonaga, A pulse discharge panel display for producing a color TV picture with high luminance and luminous efficiency, IEEE Trans. Electron Devices, 29, 988, 1982.

20. M. Ushirozawa, Y. Motoyama, T. Sakai, K. Wani, and K. Takahashi, Color DC-PDP with an improved resistor design in each cell, paper 33.2, pp. 719–722, SID '94 Symp. Dig., 1994.

21. S. Mikoshiba, Color plasma displays: where are we now?, Inf. Display, 10(10), 21, 1994.

22. Y. Takano et al., A 40-in, DC-PDP with new pulse memory drive scheme, SID '97 Dig., 1997.

23. P.S. Friedman, Are plasma display panels a low-cost

technology?, Inf. Display, 11(10), 22–25, 1995.

24. Inf. Display, 10(7 & 8), 28, 1994.

25. A. von Engle, Ionized Gases, AIP Press, New York, 1994.

26. S. Brown, Basic Data of Plasma Physics — The Fundamental Data on Electrical Discharges in Gas, American Institute of Physics Press, New York, 1993.

27. F.F. Chen, Introduction to Plasma Physics, Plenum Press, New York, 1977.

28. D. Miller, J. Ogle, R. Cola, B. Caras, and T. Maloney, An improved performance self-scan I panel design, Proc. SID, 22, 159–163, 1981.

29. M.H. Miller and R.A. Roig, Transition probabilities of XeI and XeII, Phys. Rev. A, 8, 480, July 1973.

30. Atomic Data and Nuclear Data Tables, Vol. 21, No. 6, Academic Press, New York, 1978.

31. R.S. Van Dyck, C.E. Johnson, and H.A. Shugart, Lifetime lower limits for the 3 P 0 and 3 P 2 metastable states of neon, argon and krypton, Phys. Rev. A, 5, 991–993, 1972.

32. A.A. Kruithof and F.M. Penning, The Townsend ionization coefficient and some elementary processes in neon with small admixtures of argon, Physica, 4(6), 450, 1937.

33. F.L. Jones and W.R. Galloway, The sparking potential of mercury vapor, Proc. Phys. Soc. Lond., 50, 207–212, 1938.

34. R.T. McGrath, R. Veerasingam, J.A. Hunter, P.D. Rockett, and R.B. Campbell, Measurements and simulations of VUV emissions from plasma flat panel display pixel microdischarges, IEEE Trans. on Plasma Sci., 26(5), 1998.

35. P. Hines, Spectroscopic and electrical measurements on opposed electrode ac plasma display pixels, Honors engineering senior thesis, The Pennsylvania State University, University Park, 1997.

36. R. Veerasingam, R.B. Campbell, and R.T. McGrath, One-dimensional single and multipulse simulations of the ON/OFF voltages and bistable margin for He, Xe, and He/Xe filled plasma display pixels, IEEE Trans. Plasma Sci., 24(6), 1399–1410, 1996.

37. R. Veerasingam, R.B. Campbell, and R.T. McGrath, Two-dimensional simulations of plasma flow and charge spreading in ac plasma displays, IEEE Trans. Plasma Sci., 24, 1411-1421, 1996.

38. M. Scott, Los Alamos National Laboratory, Private communication, 1997.

39. H. Uchike, K. Miura, N. Nakayama, T. Shinoda, and Y. Fukushima, Secondary electron emission characteristics of dielectric materials in ac-operated plasma display panels, IEEE Trans. Electron. Devices, 23, 1211–1217, 1976.

40. N.J. Chou and O. Sahni, Comments on “Secondary Emission Characteristics of Dielectric Materials in ac Operated Plasma Display Panels,” IEEE Trans. Electron. Devices, 25, 60–62, 1978.

41. T. Shinoda, H. Uchike, and S. Andoh, Low-voltage operated ac plasma-display panels, IEEE Trans. Electron. Devices, 26, 1163–1167, 1979.

42. M. Aboelfotoh and J.A. Lorenzen, Influence of secondary-electron emission from MgO surfaces on voltage-breakdown curves in Penning mixtures for insulated-electrode discharges, J. Appl. Phys., 48, 4754–4759, 1977.

43. T. Nenadovic, B. Perraillon, Z. Bogdanov, Z. Djordjevic, and M. Milic, Sputtering and surface topography of oxides, Nucl. Instrum. Methods Phys. Res., B48, 538–543, 1990.

44. Schone, D. Walsh, R.T. McGrath and J.H. Burkhart, Microbeam Rutherford backscattering measurements of flat panel display thin film erosion, Nuclear Instru. Methods Phys. Rev.-B, 130(1–4), 543–550, 1998.

45. CRC Handbook of Chemistry and Physics, 48th ed., R.C. Weast and S. M. Selby, Eds., The Chemical Rubber Co., Cleveland, OH, 1967.

46. S. Matsumoto, Electronic Displays Devices, John Wiley & Sons, New York, 1990.

47. G.E. Holz, The primed gas discharge cell — a cost and capability improvement for gas discharge matrix displays, Proc. SID, 13(2), 1972.

48. T. Yamamoto, T. Kuriyama, M. Seki, T. Katoh, H.

Murakami, K. Shimada, and H. Ishiga, A 40 in. diagonal HDTV plasma display, SID' 93 Symp. Dig., 165–168, 1993.

49. J. Deschamps and H. Doyeus, Plasma displays, Phys. World, June 1977.

50. N. Yocum, R.S. Meltzer, K.W. Jang, and M. Grimm, New green phosphors for plasma displays, J. SID, 4/3, 169–172, 1996.

51. H. Fujii, H. Tanabe, H. Ishiga, M. Harayama, and M. Oka, A sandblasting process for fabrication of color PDP phosphor screens, paper 37.5, 728–731, SID '92 Symp. Dig., 1992.

52. M.C. Castex, Experimental determination of the lowest excited Xe 2 molecular states from VUV absorption spectra, J. Chem. Phys., 74, 759–771, 1981.

53. M.R. Flannery, K.J. McCann, and N.W. Winter, Cross sections for electron impact ionization of metastable rare gas excimers (He 2 *, Kr 2 *, Xe 2 *), J. Phys. B Mol. Phys., 14, 3789–3796, 1981.

54. K. Tachibana, Spatio-temperal measurement of excited Xe(1s 4 ) atoms in a discharge cell of a plasma display panel by laser induced spectroscopic microscopy, Appl. Phys. Lett., 65, 935–937, 1994.

55. T. Holstein, Imprisonment of resonance radiation in gases, Phys. Rev., 72, 1212, 1947.

56. De Husk and S.E. Schnatterly, Quantum efficiency and linearity of 16 phosphors in the soft x-ray regime, J. Opt. Soc. Am. B, 9, 660–663, 1992.

57. J. Koike, T. Kojima, and R. Toyonaga, New tricolor phosphors for gas discharge display, J. Electrochem. Soc. Solid-State Sci. Tech., 1008, June 1979.

58. P.N. Yocum, Future requirements of display phosphors from an historical perspective, J. SID, 4/3, 149, 1996.

59. L.E. Tannas, Jr., Ed., Flat Panel Displays and CRTs, Van Hostrand Reinhold, New York, 1985.

60. S.G. McLean and W.W. Duley, VUV absorption in thin MgO films, J. Phys. Chem. Solids, 45, 223, 1984.

61. R. Kuntz, CyberOptics, Private communication,

1-800-746-6315.

62. Zygo, Private communication, 1-860-347-8506, www.zygo.com.

63. A. Machura, Leica, Private communication, 49 (0) 6441 29-2316, Andreas.Machura@lmw.leica.com.

64. B.E. Mills, D.A. Buchenauer, A.E. Pontau, and M. Ulrickson, Characterization of deposition and erosion on the TFTR bumper limiter and wall, J. Nucl. Mater., 162–164, 343–349, 1989.

1. Y.A. Ono, Electroluminescent Displays, Singapore: World Scientific, 1995.

2. T. Inoguchi, M. Takeda, Y. Kakihara, Y. Nakata, and M. Yoshida, “Stable high-brightness thin-film electroluminescent panels,” Dig. 1974 SID International Symposium, 84, 1974.

3. C.N. King, R.E. Coovert, and W.A. Barrow, “Full-color 320 ¥ 240 TFEL display panel,” Eurodisplay ’87, 14, 1987.

4. W.A. Barrow, R.E. Coovert, C.N. King, and M.J. Ziuchkovski, “Matrix-addressed full-color TFEL display,” Dig. 1988 SID International Symposium, 284, 1988.

5. R. Tueta and M. Braguier, “Fabrication and characterization of indium tin oxide thin films for electroluminescent applications,” Thin Solid Films, 80: 143, 1981.

6. J.D. Davidson, J.F. Wager, and I. Khormaei, “Electrical characterization and SPICE modeling of ZnS:Mn ACTFEL devices,” Dig. 1991 SID International Symposium, 77, 1991.

7. A.A. Douglas and J.F. Wager, “ACTFEL device response to systematically varied pulse waveforms,” in Electroluminescence — Proceedings of the Sixth International Workshop on Electroluminescence, V.P. Singh and J.C. McClure, Eds., El Paso, TX: Cinco Puntos Press, 1992.

8. D.H. Smith, “Modeling AC thin-film electroluminescent devices,” J. Lumin., 23: 209, 1981.

9. P.M. Alt, “Thin-film electroluminescent displays: device characteristics and performance,” Proc. SID, 25: 123, 1984.

10. Y.A. Ono, H. Kawakami, M. Fuyama, and K. Onisawa, “Transferred charge in the active layer and EL device characteristics of TFEL cells,” Jpn. J. Appl. Phys., 26: 1482, 1987.

11. E. Bringuier, “Charge transfer in ZnS-type electroluminescence,” J. Appl. Phys., 66: 1314, 1989.

12. A. Abu-Dayah, S. Kobayashi, and J. F. Wager, “Internal charge-phosphor field characteristics of alternating-current thin-film electroluminescent devices,”

Appl. Phys. Lett., 62: 744, 1993.

13. A. Mikami, K. Terada, K. Okibayashi, K. Tanaka, M. Yoshida, and S. Nakajima, “Aging characteristics of ZnS:Mn electroluminescent films grown by a chemical vapor deposition technique,” J. Appl. Phys., 72: 773, 1992.

# 35 35. Light-Emitting Diode Displays


M.A. Karim, Electro-Optical Devices and Systems, Boston: PWS-Kent Publishing, 1990.

M.A. Karim (Ed.), Electro-Optical Displays, New York: Marcel Dekker, 1992.

L.E. Tannas, Jr. (Ed.), Flat-Panel Displays and CRTs, New York: Van Nostrand Reinhold, 1985.

T. Uchida, Multicolored Liquid Crystal Displays, Opt. Eng., 23, 247–252, 1984.

J. Wilson and J.F.B. Hawkes, Optoelectronics: An Introduction, Englewood Cliffs, NJ: Prentice-Hall International, 1985.

1. H. Vermariën, Recorders, graphic, in J.G. Webster, Ed., Encyclopedia of Medical Devices and Instrumentation, New York: John Wiley & Sons, 1988.

2. D.A. Bell, Electronic Instrumentation and Measurements, 2nd ed., Englewood Cliffs, NJ: PrenticeHall, 1994.

3. A. Miller, O.S. Talle, and C.D. Mee, Recorders, in B.M. Oliver and J.M. Cage, Eds., Electronic Measurementsand Instrumentation, New York: McGraw-Hill, 427–479, 1975.

4. R.J. Smith, Circuits, Devices and Systems: A First Course in Electrical Engineering, New York: John Wiley & Sons, 1976.

5. W.H. Olson, Basic concepts in instrumentation, in J.G. Webster, Ed., Medical Instrumentation: Application and Design, 3rd ed., New York: John Wiley & Sons, 1998.

## 36.2 Data Acquisition Systems

Edward McConnell

The fundamental task of a data acquisition system is the measurement or generation of real-world

physical signals. Before a physical signal can be measured by a computer-based system, a sensor or

transducer is used to convert the physical signal into an electrical signal, such as voltage or current. Often

only a plug-in data acquisition (DAQ) board is considered the data acquisition system; however, a board

is only one of the components in the system. A complete DAQ system consists of sensors, signal condi

tioning, interface hardware, and software. Unlike stand-alone instruments, signals often cannot be directly

connected to the DAQ board. The signals may need to be conditioned by some signal-conditioning

accessory before they are converted to digital information by the plug-in DAQ board. Software controls

the data acquisition system — acquiring the raw data, analyzing the data, and presenting the results. The

components are shown in Figure 36.1.

Signals

Signals are physical events whose magnitude or time variation contains information. DAQ systems

measure various aspects of a signal in order to monitor and control the physical events. Users of DAQ

FIGURE 36.1 Components of a DAQ system.

systems need to know the relation of the signal to the physical event and what information is available

in the signal. Generally, information is conveyed by a signal through one or more of the following signal

parameters: state, rate, level, shape, or frequency content. The physical characteristics of the measured

signals and the related information help determine the design of a DAQ system.

All signals are, fundamentally, analog, time-varying signals. For the purpose of discussing the methods

of signal measurement using a plug-in DAQ board, a given signal should be classified as one of five signal

types. Because the method of signal measurement is determined by the way the signal conveys the needed

information, a classification based on this criterion is useful in understanding the fundamental building

blocks of a DAQ system.

As shown in Figure 36.2, any signal can generally be classified as analog or digital. A digital, or binary,

signal has only two possible discrete levels of interest — a high (on) level and a low (off) level. The two

digital signal types are on–off signals and pulse train signals. An analog signal, on the other hand, contains

information in the continuous variation of the signal with time. Analog signals are described in the time

or frequency domains depending upon the information of interest. A dc type signal is a low-frequency

signal, and if the phase information of a signal is presented with the frequency information, then there

is no difference between the time or frequency domain representations. The category to which a signal

belongs depends on the characteristic of the signal to be measured. The five types of signals can be closely

paralleled with the five basic types of signal information — state, rate, level, shape, and frequency content.

Basic understanding of the signal representing the physical event being measured and controlled assists

in the selection of the appropriate DAQ system.

Plug-In DAQ Boards

The fundamental component of a DAQ system is the plug-in DAQ board. These boards plug directly into

a slot in a PC and are available with analog, digital, and timing inputs and outputs (I/O). The most versatile

of the plug-in DAQ boards is the multifunction I/O board. As the name implies, this board typically

contains various combinations of analog-to-digital converters (ADCs), digital-to-analog converters

FIGURE 36.2 Classes of signals.

(DACs), digital I/O lines, and counters/timers. ADCs and DACs measure and generate analog voltage

signals, respectively. The digital I/O lines sense and control digital signals. Counters/timers measure pulse

rates, widths, delays, and generate timing signals. These many features make the multifunction DAQ

board useful for a wide range of applications.

Multifunction boards are commonly used to measure analog signals. This is done by the ADC, which

converts the analog voltage level into a digital number that the computer can interpret. The analog

multiplexer (MUX), the instrumentation amplifier, the sample-and-hold (S/H) circuitry, and the ADC

compose the analog input section of a multifunction board (see Figure 36.3).

Typically, multifunction DAQ boards have one ADC. Multiplexing is a common technique for mea

suring multiple channels (generally 16 single-ended or 8 differential) with a single ADC. The analog

MUX switches between channels and passes the signal to the instrumentation amplifier and the S/H

circuitry. The MUX architecture is the most common approach taken with plug-in DAQ boards. While

plug-in boards typically include up to only 16 single-ended or 8 differential inputs, the number of analog

input channels can be further expanded with external MUX accessories.

Instrumentation amplifiers typically provide a differential input and selectable gain by jumpers or

software. The differential input rejects small common-mode voltages. The gain is often software pro

grammable. In addition, many DAQ boards also include the capability to change the amplifier gain while

scanning channels at high rates. Therefore, one can easily monitor signals with different ranges of

amplitudes. The output of the amplifier is sampled, or held at a constant voltage, by the S/H device at

measurement time so that voltage does not change during digitization.

The ADC transforms the analog signal into a digital value which is ultimately sent to computer memory.

There are several important parameters of A/D conversion. The fundamental parameter of an ADC is

the number of bits. The number of bits of an A/D determines the range of values for the binary output

of the ADC conversion. For example, many ADCs are 12-bit, so a voltage within the input range of the

ADC will produce a binary value that has one of $2^{12}$ = 4096 different values. The more bits that an ADC

has, the higher the resolution of the measurement. The resolution determines the smallest amount of

change that can be detected by the ADC. Resolution is expressed as the number of digits of a voltmeter

or dynamic range in decibels, rather than with bits. Table 36.6 shows the relation among bits, number

of digits, and dynamic range in decibels.

The resolution of the A/D conversion is also determined by the input range of the ADC and the gain.

DAQ boards usually include an instrumentation amplifier that amplifies the analog signal by a gain factor

prior to the conversion. This gain amplifies low-level signals so that more accurate measurements can

be made.

Together, the input range of the ADC, the gain, and the number of bits of the board determine the

minimum resolution of the measurement. For example, suppose a low-level ±30 mV signal is acquired using

a 12-bit ADC that has a ±5 V input range. If the system includes an amplifier with a gain of 100, the resulting

resolution of the measurement will be range/(gain * $2^{bits}$) = resolution, or 10 V/(100 * $2^{12}$) = 0.0244 mV.

FIGURE 36.3 Analog input section of a plug-in DAQ board. Note: FIFO = first-in first-out buffer, S/H = sample

and-hold, Inst. Amp = instrumentation amplifier, and Mux =

analog multiplexer.

Finally, an important parameter of digitization is the rate at which A/D conversions are made, referred

to as the sampling rate. The A/D system must be able to sample the input signal fast enough to measure

the important waveform attributes accurately. In order to meet this criterion, the ADC must be able to

convert the analog signal to digital form quickly enough.

When scanning multiple channels with a multiplexing DAQ system, other factors can affect the

throughput of the system. Specifically, the instrumentation amplifier must be able to settle to the needed

accuracy before the A/D conversion occurs. With multiplexed signals, multiple signals are being switched

into one instrumentation amplifier. Most amplifiers, especially when amplifying the signals with larger

gains, will not be able to settle to the full accuracy of the ADC when scanning channels at high rates. To

avoid this situation, consult the specified settling times of the DAQ board for the gains and sampling

rates required by the application.

Types of ADCs

Different DAQ boards use different types of ADCs to digitize the signal. The most popular type of ADC

on plug-in DAQ boards is the successive approximation ADC, because it offers high speed and high

resolution at a modest cost.

Subranging (also called half-flash) ADCs offer very high speed conversion with sampling speeds up

to several million samples per second.

The state-of-the-art technology in ADCs is sigma-delta modulating ADCs. These ADCs sample at high rates, are able to achieve high resolution, and offer the best linearity of all ADCs.

Integrating and flash ADCs are mature technologies still used on DAQ boards today. Integrating ADCs

are able to digitize with high resolution but must sacrifice sampling speed to obtain it. Flash ADCs are

able to achieve the highest sampling rate (gigahertz) but are available only with low resolution. The

different types of ADCs are summarized in Table 36.7. TABLE 36.6 Relation Among Bits, Number of Digits, and Dynamic Range (dB) Bits Digits dB 20 6.0 120 16 4.5 96 12 3.5 72 8 2.5 48 TABLE 36.7 Types of ADCs Type of ADC Advantages Features Successive approximation High resolution 1.25 MS/s sampling rate High speed 12-bit resolution Easily multiplexed 200 kS/s sampling rate 16-bit resolution Subranging Higher speed 1 MHz sampling rate 12-bit resolution Sigma–delta High resolution 48 kHz sampling rate Excellent linearity 16-bit resolution Built-in antialiasing State-of-the-art technology Integrated High resolution 15 kHz sampling rate Good noise rejection Mature technology Flash Highest speed 125 MHz sampling rate Mature technology

Analog Input Architecture

With a typical DAQ board, the multiplexer switches among analog input channels. The analog signal on

the channel selected by the multiplexer then passes to the programmable gain instrumentation amplifier

(PGIA), which amplifies the signal. After the signal is amplified, the sample and hold (S/H) keeps the

analog signal constant so that the ADC can determine the digital representation of the analog signal. A

good DAQ board will then place the digital signal in a first-in first-out (FIFO) buffer, so that no data

will be lost if the sample cannot transfer immediately over the PC I/O channel to computer memory.

Having a FIFO becomes especially important when the board is run under operating systems that have large interrupt latencies, such as Microsoft Windows.

## Basic Analog Specifications

Almost every DAQ board data sheet specifies the number of channels, the maximum sampling rate, the resolution, and the input range and gain.

The number of channels, which is determined by the multiplexer, is usually specified in two forms — differential and single ended. Differential inputs are inputs that have different reference points for each channel, none of which is grounded by the board. Differential inputs are the best way to connect signals to the DAQ board because they provide the best noise immunity.

Single-ended inputs are inputs that are referenced to a common ground point. Because single-ended inputs are referenced to a common ground, they are not as good as differential inputs for rejecting noise. They do have a larger number of channels, however. Single-ended inputs are used when the input signals are high level (greater than 1 V), the leads from the signal source to the analog input hardware are short (less than 5 m), and all input signals share a common reference.

Some boards have pseudodifferential inputs which have all inputs referenced to the same common — like single-ended inputs — but the common is not referenced to ground. These boards have the benefit of a large number of input channels, like single-ended inputs, and the ability to remove some common mode noise, especially if the common-mode noise is consistent across all channels. Differential inputs are still preferable to pseudodifferential, however, because differential is more immune to magnetic noise.

Sampling rate determines how fast the analog signal is converted to a digital signal. When measuring ac signals, sample at least two times faster than the highest frequency of the input signal. Even when measuring dc signals, oversample and average the data to increase the accuracy of the signal by reducing the effects of noise.

If the physical event consists of multiple dc-class signals, a DAQ board with interval scanning should be used. With interval scanning, all channels are scanned at one sample interval (usually the fastest rate of the board), with a second interval (usually slow) determining the time before repeating the scan. Interval scanning gives the effects of simultaneously sampling for slowly varying signals without requiring the additional cost of input circuitry for true simultaneous sampling.

Resolution is the number of bits that are used to represent the analog signal. The higher the resolution, the higher the number of divisions the input range is broken into, and therefore the smaller the possible detectable voltage. Unfortunately, some DAQ specifications are misleading when they specify the resolution associated with the DAQ board. Many DAQ board specifications state the resolution of the ADC without stating the linearities and noise, and therefore do not give the information needed to determine the resolution of the entire board. Resolution of the ADC, combined with the settling time, integral nonlinearity (INL), differential nonlinearity (DNL), and noise will give an understanding of the accuracy of the board.

Input range and gain determine the level of signal that should be connected to the board. Usually, the

range and gain are specified separately, so the two must be combined to determine the actual signal input

range as sig nal inpu t range = r ange/g ain

For example, a board using an input range of ±10 V with a gain of 2 will have a signal input range of

±5 V. The closer the signal input range is to the range of the signal, the more accurate the readings from

the DAQ board will be. If the signals have different input ranges, use a DAQ board with the feature of

different gains per channel.

Data Acquisition Software

The software is often the most critical component of the DAQ system. Users of DAQ systems usually

program the hardware in one of two ways — through register programming or through high-level device

drivers.

Board Register-Level Programming

The first option is not to use vendor-supplied software and program the DAQ board at the hardware

level. DAQ boards are typically register based; that is, they include a number of digital registers that

control the operation of the board. The developer may use any standard programming language, such

as C, C++, or Visual BASIC, to write a series of binary codes to the DAQ board to control its operation.

Although this method affords the highest level of flexibility, it is also the most difficult and time-consuming,

especially for the inexperienced programmer. The programmer must know the details of programming all hardware, including the board, the PC interrupt controller, the DMA controller, and PC memory.

## Driver Software

Driver software typically consists of a library of function calls usable from a standard programming

language. These function calls provide a high-level interface to control the standard functions of the

plug-in board. For example, a function called SCAN_OP may configure, initiate, and complete a multiple

channel scanning DAQ operation of a predetermined number of points. The function call would include

parameters to indicate the channels to be scanned, the amplifier gains to be used, the sampling rate, and

the total number of data points to be collected. The driver responds to this one function call by pro

gramming the plug-in board, the DMA controller, the interrupt controller, and CPU to scan the channels

as requested.

## What Is Digital Sampling?

Every DAQ system has the task of gathering information about analog signals. To do this, the system

captures a series of instantaneous “snapshots” or samples of the signal at definite time intervals. Each

sample contains information about the signal at a specific instant. Knowing the exact time of each

conversion and the value of the sample, one can reconstruct, analyze, and display the digitized waveform.

## Real-Time Sampling Techniques

In real-time sampling, the DAQ board digitizes consecutive samples along the signal (Figure 36.4).

According to the Nyquist sampling theorem, the ADC must sample at least twice the rate of the maximum frequency component in that signal to prevent aliasing. Aliasing is a false lower-frequency component

that appears in sampled data acquired at too low a sampling rate. The frequency at one half the sampling

frequency is referred to as the Nyquist frequency. Theoretically, it is possible to recover information about

those signals with frequencies at or below the Nyquist frequency. Frequencies above the Nyquist frequency

will alias to appear between dc and the Nyquist frequency.

For example, assume the sampling frequency, $f_s$, is 100 Hz. Also assume the input signal to be sampled

contains the following frequencies — 25, 70, 160, and 510 Hz. Figure 36.5 shows a spectral representation

of the input signal.

The mathematics of sampling theory show us that a sampled signal is shifted in the frequency domain

by an amount equal to integer multiples of the sampling frequency, $f_s$. Figure 36.6 shows the spectral

content of the input signal after sampling. Frequencies below 50 Hz, the Nyquist frequency ($f_s/2$), appear

correctly. However, frequencies above the Nyquist appear as aliases below the Nyquist frequency. For

example, F1 appears correctly; however, F2, F3, and F4 have aliases at 30, 40, and 10 Hz, respectively.

The resulting frequency of aliased signals can be calculated with the following formula:

FIGURE 36.4 Consecutive discrete samples recreate the input signal.

FIGURE 36.5 Spectral of signal with multiple frequencies.

FIGURE 36.6 Spectral of signal with multiple frequencies after sampling at $f_s$ = 100 Hz. Apparent Alias Freq. ABS Closest Integer Multiple of Sampling Freq. Input Freq. ( ) = ( )

For the example of Figures 36.5 and 36.6:

## Preventing Aliasing

Aliasing can be prevented by using filters on the front end of the DAQ system. These antialiasing filters

are set to cut off any frequencies above the Nyquist frequency (half the sampling rate). The perfect filter

would reject all frequencies above the Nyquist; however, because perfect filters exist only in textbooks,

one must compromise between sampling rate and selecting filters. In many applications, one- or two

pole passive filters are satisfactory. The rule of thumb is to oversample (5 to 10 times) and use these

antialiasing filters when frequency information is crucial.

Alternatively, active antialiasing filters with programmable cutoff frequencies and very sharp attenu

ation of frequencies above the cutoff can be used. Because these filters exhibit a very steep roll-off, the

DAQ system can sample at two to three times the filter cutoff frequency. Figure 36.7 shows a transfer

function of a high-quality antialiasing filter.

The computer uses digital values to recreate or to analyze the waveform. Because the signal could be

anything between each sample, the DAQ board may be unaware of any changes in the signal between

samples. There are several sampling methods optimized for the different classes of data; they include

software polling, external sampling, continuous scanning, multirate scanning, simultaneous sampling,

interval scanning, and seamless changing of the sample rate.

## Software Polling

A software loop polls a timing signal and starts the A/D conversion via a software command when the edge of the timing signal is detected. The timing signal may originate from the internal clock of the

computer or from a clock on the DAQ board. Software polling is useful in simple, low-speed applications,

such as temperature measurements.

The software loop must be fast enough to detect the timing signal and trigger a conversion. Otherwise,

a window of uncertainty, also known as jitter, will exist between two successive samples. Within the

window of uncertainty, the input waveform could change enough to reduce the accuracy of the ADC

drastically.

FIGURE 36.7 Magnitude portion of transfer function of an antialiasing filter. Alias F Hz Alias F Hz Alias F Hz 2 100 70 30 3 2 100 160 40 4 5 100 510 10 = = = ( ) = = ( ) =

Suppose a 100-Hz, 10-V full-scale sine wave is digitized (Figure 36.8). If the polling loop takes 5 ms

to detect the timing signal and to trigger a conversion, then the voltage of the input sine wave will change

as much as 31 mV, [DV = 10 sin (2p ¥ 100 ¥ 5 ¥ 10 -6 )]. For a 12-bit ADC operating over an input range

of 10 V and a gain of 1, one least significant bit (LSB) of error represents 2.44 mV:

But because the voltage error due to jitter is 31 mV, the accuracy error is 13 LSB.

This represents uncertainty in the last 4 bits of a 12-bit ADC. Thus, the effective accuracy of the system

is no longer 12 bits but rather 8 bits.

External Sampling

Some DAQ applications must perform a conversion based on another physical event that triggers the

data conversion. The event could be a pulse from an optical encoder measuring the rotation of a cylinder.

A sample would be taken every time the encoder generates a pulse corresponding to n degrees of rotation.

External triggering is advantageous when trying to measure signals whose occurrence is relative to another

physical phenomenon.

Continuous Scanning

When a DAQ board acquires data, several components on the board convert the analog signal to a digital

value. These components include the analog MUX, the instrumentation amplifier, the S/H circuitry, and

the ADC. When acquiring data from several input channels, the analog MUX connects each signal to

the ADC at a constant rate. This method, known as continuous scanning, is significantly less expensive

than having a separate amplifier and ADC for each input channel.

Continuous scanning is advantageous because it eliminates jitter and is easy to implement. However,

it is not possible to sample multiple channels simultaneously. Because the MUX switches between

channels, a time skew occurs between any two successive channel samples. Continuous scanning is

appropriate for applications where the time relationship between each sampled point is unimportant or

where the skew is relatively negligible compared with the speed of the channel scan.

If samples from two signals are used to generate a third value, then continuous scanning can lead to

significant errors if the time skew is large. In Figure 36.9, two channels are continuously sampled and

added together to produce a third value. Because the two sine waves are 90° out-of-phase, the sum of the signals should always be zero. But because of the skew time between the samples, an erroneous

sawtooth signal results.

FIGURE 36.8 Jitter reduces the effective accuracy of the DAQ board. Input range gain V mV ¥ Ê Ë Á ˆ ¯ = ¥ Ê Ë Á ˆ ¯ = 2 10 1 2 2 44 12n . 31 2 44 mV mV. Ê Ë Á ˆ ¯

## Multirate Scanning

Multirate scanning, a method that scans multiple channels at different scan rates, is a special case of continuous scanning. Applications that digitize multiple signals with a variety of frequencies use multirate scanning to minimize the amount of buffer space needed to store the sampled signals. Channel-independent ADCs are used to implement hardware multirate scanning; however, this method is extremely expensive. Instead of multiple ADCs, only one ADC is used. A channel/gain configuration register stores the scan rate per channel and software divides down the scan clock based on the per-channel scan rate.

Software-controlled multirate scanning works by sampling each input channel at a rate that is a fraction of the specified scan rate.

Suppose the system scans channels 0 through 3 at 10 kS/s, channel 4 at 5 kS/s, and channels 5 through 7 at 1 kS/s. A base scan rate of 10 kS/s should be used. Channels 0 through 3 are acquired at the base scan rate. Software and hardware divide the base scan rate by 2 to sample channel 4 at 5 kS/s, and by 10 to sample channels 5 through 7 at 1 kS/s.

## Simultaneous Sampling

For applications where the time relationship between the input signals is important, such as phase analysis of ac signals, simultaneous sampling must be used. DAQ boards capable of simultaneous sampling typically use independent instrumentation amplifiers and S/H circuitry for each input channel, along with an analog MUX, which routes the input signals to the ADC for conversion (as shown in Figure 36.10).

To demonstrate the need for a simultaneous-sampling DAQ board, consider a system consisting of four 50 kHz input signals sampled at 200 kS/s. If the DAQ board uses continuous scanning, the skew between each channel is 5 ms (1S/200 kS/s) which represents a 270˚ [(15 ms/20 ms) ¥ 360˚] shift in phase between the first channel and fourth channel. Alternatively, with a simultaneous-sampling board with a maximum 5 ns interchannel time offset, the phase shift is only 0.09˚ [(5 ms/20 ms) ¥ 360˚]. This phenomenon is illustrated in Figure 36.11.

## Interval Scanning

For low-frequency signals, interval scanning creates the effect of simultaneous sampling, yet maintains the cost benefits of a continuous-scanning system. This method scans the input channels at one rate and uses a second rate to control when the next scan begins. If the input channels are scanned at the fastest

FIGURE 36.9 If the channel skew is large compared with the signal, then erroneous conclusions may result.

rate of the ADC, the effect of simultaneously sampling the channels is created. Interval scanning is appropriate for slow-moving signals, such as temperature and pressure. Interval scanning results in a jitter-free sample rate and minimal skew time between channel samples. For example, consider a DAQ system with ten temperature signals. By using interval scanning, a DAQ board can be set up to scan all

channels with an interchannel delay of 5 ms, then repeat the scan every second. This method creates the

effect of simultaneously sampling ten channels at 1 S/s, as shown in Figure 36.12.

To illustrate the difference between continuous and interval scanning, consider an application that

monitors the torque and RPMs of an automobile engine and computes the engine horsepower. Two

signals, proportional to torque and RPM, are easily sampled by a DAQ board at a rate of 1000 S/s. The

values are multiplied together to determine the horsepower as a function of time.

FIGURE 36.10 Block diagram of DAQ components used to sample multiple channels simultaneously.

FIGURE 36.11 Comparison of continuous scanning and simultaneous sampling.

FIGURE 36.12 Interval scanning — all ten channels are scanned within 45 ms; this is insignificant relative to the

overall acquisition rate of 1 S/s.

A continuously scanning DAQ board must sample at an aggregate rate of 2000 S/s. The time between

which the torque signal is sampled and the RPM signal is sampled will always be 0.5 ms (1/2000). If

either signal changes within 0.5 ms, then the calculated horsepower is incorrect. But using interval

scanning at a rate of 1000 S/s, the DAQ board samples the torque signal every 1 ms, and the RPM signal

is sampled as quickly as possible after the torque is sampled. If a 5-ms interchannel delay exists between

the torque and RPM samples, then the time skew is reduced by 99% [(0.5 ms – 5 ms)/0.5 ms], and the chance of an incorrect calculation is reduced.

## Factors Influencing the Accuracy of Measurements

How does one determine if a plug-in DAQ will deliver the required measurement results? With a sophisticated measuring device like a plug-in DAQ board, significantly different accuracies can be obtained depending on the type of board used. For example, one can purchase DAQ products on the market today with 16-bit ADCs and get less than 12 bits of useful data, or one can purchase a product with a 16-bit ADC and actually get 16 bits of useful data. This difference in accuracies causes confusion in the PC industry where everyone is used to switching out PCs, video cards, printers, and so on, and experiencing similar results between equipment.

The most important thing to do is to scrutinize more specifications than the resolution of the ADC that is used on the DAQ board. For dc-class measurements, one should at least consider the settling time of the instrumentation amplifier, DNL, relative accuracy, INL, and noise. If the manufacturer of the board under consideration does not supply these specifications in the data sheets, ask the vendor to provide them or run tests to determine these specifications.

## Defining Terms

Alias: A false lower frequency component that appears in sampled data acquired at too low a sampling rate.

Asynchronous: (1) Hardware — A property of an event that occurs at an arbitrary time, without synchronization to a reference clock. (2) Software — A property of a function that begins an operation and returns prior to the completion or termination of the operation.

Conversion time: The time required, in an analog input or output system, from the moment a channel is interrogated (such as with a read instruction) to the moment that accurate data are available.

DAQ (data acquisition): (1) Collecting and measuring electric signals from sensors, transducers, and test probes or fixtures and inputting them to a computer for processing: (2) Collecting and measuring the same kinds of electric signals with ADC and/or DIO boards plugged into a PC, and possibly generating control signals with DAC and/or DIO boards in the same PC.

DNL (differential nonlinearity): A measure in LSB of the worst-case deviation of code widths from their ideal value of 1 LSB.

INL (integral nonlinearity): A measure in LSB of the worst-case deviation from the ideal A/D or D/A transfer characteristic of the analog I/O circuitry.

Nyquist sampling theorem: A law of sampling theory stating that if a continuous bandwidth-limited signal contains no frequency components higher than half the frequency at which it is sampled, then the original signal can be recovered without distortion.

Relative accuracy: A measure in LSB of the accuracy of an ADC. It includes all nonlinearity and quantization errors. It does not include offset and gain errors of the circuitry feeding the ADC.

Further Information

House, R., “Understanding Important DA Specifications,” Sensors, 10(10), June 1993.

House, R., “Understanding Inaccuracies Due to Settling Time, Linearity, and Noise,” National Instruments European User Symposium Proceedings, November 10–11, 1994, pp. 11–12.

McConnell, E., “PC-Based Data Acquisition Users Face Numerous Challenges,” ECN, August 1994.

McConnell, E., “Choosing a Data-Acquisition Method,” Electronic Design, 43(6), 147, 1995.

McConnell, E. and Jernigan, D., “Data Acquisition,” in The Electronics Handbook, J.C. Whitaker (Ed.), Boca Raton, FL: CRC Press, 1996, 1795–1822.

Potter, D. and Razdan, A., “Fundamentals of PC-Based Data Acquisition,” Sensors, 11( 2), 12–20, February 1994.

Potter, D., “Sensor to PC — Avoiding Some Common Pitfalls,” Sensors Expo Proceedings, September 20, 1994.

Potter, D., “Signal Conditioners Expand DAQ System Capabilities,” I&CS, 25–33, August 1995.

Johnson, G. W., LabVIEW Graphical Programming, New York: McGraw-Hill, 1994.

McConnell, E., “New Achievements in Counter/Timer Data Acquisition Technology,” MessComp 1994 Proceedings, September 13–15, 1994, 492–498.

McConnell, E., “Equivalent Time Sampling Extends DA Performance,” Sensors Data Acquisition, Special Issue, June, 13, 1995.

## 36.3 Magnetic and Optical Recorders

Yufeng Li

The heart of recording technology is for the process of information storage and retrieval. In addition to

its obvious importance in different branches of science and engineering, it has become indispensable to

our daily life. When we make a bank transaction, reserve an airplane ticket, use a credit card, watch a

movie from a video tape, or listen to music from a CD, we are using the technology of recording. The

general requirements for recording are information integrity, fast access, and low cost. Among the

different techniques, the most popularly used ones are magnetic and optical recording.

Typical recording equipment consists of a read/write head, a medium, a coding/decoding system, a

data access system, and some auxiliary mechanical and electronic components. The head and medium

are for data storage and retrieval purposes, and the coding/decoding system is for data error correction.

The data access system changes the relative position between the head and the medium, usually with a

servo mechanism for data track following and a spinning mechanism for on-track moving. While the

data access system and the auxiliary components are important to recording equipment, they are not

considered essential in this chapter to the understanding of recording technology, and will not be covered.

Interested readers are referred to Reference 1.

## Magnetic Recording

At present, magnetic recording technology dominates the recording industry. It is used in the forms of

hard disk, floppy disk, removable disk, and tape with either digital or analog mode. In its simplest form,

it consists of a magnetic head and a magnetic medium, as shown in Figure 36.13. The head is made of

a piece of magnetic material in a ring shape (core), with a small gap facing the medium and a coil away

from the medium. The head records (writes) and reproduces (reads) information, while the medium

stores the information. The recording process is based on the phenomenon that an electric current i

generates a magnetic flux f as described by Ampere's law. The flux f leaks out of the head core at the

gap, and magnetizes the magnetic medium which moves from left to right with a velocity V under the

head gap. Depending on the direction of the electric current i, the medium is magnetized with magne

tization M pointing either left or right. This pattern of magnetization is retained in the memory of the

medium even after the head moves away.

Two types of head may be used for reproducing. One, termed the inductive head, senses magnetic flux change rate, and the other, named the magnetoresistive (MR) head, senses the magnetic flux. When an inductive head is used, the reproducing process is just the reverse of the recording process. The flux coming out of the magnetized medium surface is picked up by the head core. Because the medium magnetization under the head gap changes its magnitude and direction as the medium moves, an electric voltage is generated in the coil. This process is governed by Faraday's law. Figure 36.13b schematically shows the digital recording/reproducing process. First, all user data are encoded into a binary format — a serial of 1s and 0s. Then a write current i is sent to the coil. This current changes its direction whenever a 1 is being written. Correspondingly, a change of magnetization, termed a transition, is recorded in the medium for each 1 in the encoded data. During the reproducing process, the electric voltage induced in the head coil reaches a peak whenever there is a transition in the medium. A pulse detector generates a pulse for each transition. These pulses are decoded to yield the user data.

The minimum distance between two transitions in the medium is the flux change length B, and the distance between two adjacent signal tracks is the track pitch W, which is wider than the signal track width w. The flux change length can be directly converted into bit length with the proper code information. The reciprocal of the bit length is called linear density, and the reciprocal of the track pitch is termed track density. The information storage areal density in the medium is the product of the linear density and the track density. This areal density roughly determines how much information a user can store in

a unit surface area of storage medium, and is a figure of merit for a recording technique. Much effort

has been expended to increase the areal density. For example, it has been increased 50 times during the

FIGURE 36.13 Conceptual diagrams illustrating the magnetic recording principle (a), and recording/reproducing

process (b).

last decade in hard disk drives, and is expected to continue increasing 60% per year in the foreseeable

future. At present, state-of-the-art hard disk products feature areal densities of more than 7 Mbits/mm 2

(B < 0.1 mm and W < 1.5 mm). This gives a total storage capacity of up to 6 Gbytes for a disk of 95 mm

diameter.

## Magnetism and Hysteresis Loop

Magnetism is the result of uncompensated electron spin motions in an atom. Only transition elements

exhibit this property, and nearly all practical interest in magnetism centers on the first transition group

of elements (Mn, Cr, Fe, Ni, and Co) and their alloys. The strength of magnetism is represented by

magnetization M, and is related to magnetic field H and magnetic flux density B by (36.1)

where m 0 is the permeability of vacuum. Since M is a property of a magnetic material, it does not exist

outside the magnetic material. H represents the strength acting on a magnetic material from a magnetic

field which is generated either by a magnetic material or by an electric current. B is the flux density which

determines the induced electric voltage in a coil. The ratio of B with and without a magnetic material is

the relative permeability m of that magnetic material.

When a magnetic field H is applied to a piece of demagnetized magnetic material, the magnetization

M starts increasing with H from zero. The rate of increase gradually slows down and M asymptotically

approaches a value $M_s$ at high H. If H is reduced to zero, then M is reduced to a lower value $M_r$. Continuous

reduction of H to a very high negative value will magnetize the material to $-M_s$. In order to bring the

material to demagnetized state, a positive field $H_c$ is required. Further increase in the H field will bring

the trace of M to a closed loop. This loop is the major hysteresis loop, as shown in Figure 36.14. The

hysteresis loop shows that a magnetic material has memory. It is this memory that is used in the medium

for storing information. $H_c$ is the coercivity, indicating the strength of magnetic field required to erase

the memory of a magnetic material. Magnetic materials with high $H_c$ are “hard” magnets, and are suitable

for medium applications if they have high $M_r$. On the other hand, magnetic materials with low $H_c$ are

“soft” magnets, and are candidates for head core materials if they have high $M_s$ and high m. $M_r$ and $M_s$

are the remanent and saturation magnetization, respectively, and their ratio is the remanent squareness.

The flux density corresponding to $M_s$ is $B_s$.

## Magnetic Media

Magnetic media are used to store information in a magnetic recording system. In order to increase the

areal density, we need to reduce flux change length B and track width w. Since B is limited by the term

$M_r d/H_c$, where d is the magnetic layer thickness, we can reduce B by either decreasing $M_r d$ or increasing

$H_c$. However, the amplitude of the magnetic signal available for reproducing head is proportional to the

term $M_r dw$. If we reduce track width w to increase areal density, we must increase $M_r d$ to avoid signal

FIGURE 36.14 Hysteresis loop of a magnetic material shows the nonlinear relationship between M and H which

results in magnetic memory. $B = \mu_0 (H + M)$

deterioration. In addition, if the magnetic layer is so thin that it causes thickness nonuniformity, more

noise will appear in the reproducing process. Therefore, the major requirements for magnetic layer are

high $H_c$, high $M_r$, and ease of making a uniform thin layer. Additional requirements include good

magnetic and mechanical stability.

There are two groups of magnetic media. The first group is called particulate media because the

magnetic materials are in the form of particles. This group includes iron oxide ($\gamma$-$Fe_2O_3$), cobalt-modified

iron oxide ($\gamma$-$Fe_2O_3$+Co), chromium dioxide ($CrO_2$), metal particles, and barium ferrite ($BaFe_{12}O_{19}$).

Some of these have been used in the magnetic recording for several decades. More recently, another

group of media has been developed largely due to the ever-increasing demand for higher storage capacity

in the computer industry. This group of media is the thin-film media, where the magnetic layer can be

made as a continuous thin film. Most materials in this group are cobalt-based metal alloys. Compared

with particulate media, the thin-film media usually have a higher coercivity $H_c$, a higher remanence $M_r$,

and can be deposited in a very thin continuous film. Table

36.8 lists $H_c$ and $M_r$ for some of the most popularly used particulate and thin-film media. Note that magnetic properties are affected by the fabrication process and film structure. Therefore, their values can be out of the ranges of Table 36.8 if different processes are used.

Magnetic media can be classified into three general forms of applications. Tape is the oldest form and remains an important medium today. It is central to most audio, video, and instrumentation recording, although it is also used in the computer industry for archival storage. Tape is economical and can hold a large capacity, but suffers slow access time. Hard disk is primarily used as the storage inside a computer, providing fast data access for the user, but having poor transportability. Flexible disk is designed for easy data transportation, but is limited in capacity. Besides these three general forms of applications, a hybrid of flexible and hard disk is being gradually accepted. It is a removable rigid disk capable of holding up to several gigabytes of digital data. In addition, magnetic stripes are getting wide use in different forms of cards.

The magnetic layer alone cannot be used as a medium. It needs additional components to improve its chemical and mechanical durability. Typical cross sections of a particulate magnetic tape and a thin film hard disk are shown in Figure 36.15. In the case of tape application, iron particles with typical size of 0.5 mm long and 0.1 mm wide are dispersed in a polymeric binder, together with solvents, lubricants, and other fillers to improve magnetic and mechanical stability. This dispersed material is then coated on an abiaxially oriented polyethylene terephthalate substrate. An optional back coat may also be applied to the other side of the substrate. The cross section of a hard disk is more complex. A high-purity aluminum–magnesium (5 wt%) substrate is diamond turned to a fine surface finish, and then electrolessly plated with a nonmagnetic nickel-phosphorus (10 wt%) undercoat. This layer is used to increase the

TABLE 36.8 Remanence ($M_r$) and Coercivity ($H_c$) Values of Some Commonly Used Magnetic Media (some values are from Reference 5)

| Group | Material | $M_r$ (kA/m) | $H_c$ (kA/m) | Application |
|---|---|---|---|---|
| Particulate | g-$Fe_2O_3$ | 56–140 | 23–32 | Floppy disk, audio, video, and instrumentation tapes |
| | g-$Fe_2O_3$ +Co | 60–140 | 44–74 | Floppy disk, audio, video, and instrumentation tapes |
| | $CrO_2$ | 110–140 | 38–58 | Floppy disk, audio, video, and instrumentation tapes |
| | $BaFe_{12}O_{19}$ | 56 | 58 | Floppy disk |
| Thin film | Co-Ni | 600–1100 | 30–85 | Hard disk |
| | Co-Fe | 1100–1500 | 60–150 | Hard disk |
| | Co-P | 600–1000 | 36–120 | Hard disk |
| | Co-Ni-Pt | 600–1100 | 60–175 | Hard disk |
| | Co-Cr-Ta | 350–900 | 55–190 | Hard disk |
| | Co-Cr-Pt | 300–750 | 56–200 | Hard disk |

hardness, reduce the defects, and improve the finish of the Al–Mg alloy, and is polished to a super surface finish. Next, an underlayer of chromium is sputtered to control the properties of the magnetic film, followed by sputtering the magnetic film. Finally, a layer of hydrogenated or nitrogenated carbon is overcoated on the magnetic film, and an ultrathin layer of perfluorinated hydrocarbon liquid lubricant is applied on top. The carbon and lubricant layers are used to improve the corrosion and mechanical resistance of the disk. For a 95 mm disk the finished product should have a surface flatness better than 10 mm and a tightly control surface roughness. In some applications, an arithmetic average roughness ($R_a$) of less than 0.5 nm is required.

## Magnetic Heads

Magnetic heads have three functions: recording, reproducing, and erasing. Usually for stationary head

applications such as tape drives, multiple heads are used to perform these functions. For moving head

applications such as disk drives, a single head is employed because of the requirements of simple

connections and small head mass for fast data access. Most of these heads are the inductive type, where

the fundamental design is an inductive coil and a magnetic core. The general requirements for the core

materials are high relative permeability m, high saturation flux density $B_s$, low coercivity $H_c$, high electric

resistivity r, and low magnetostriction coefficient l. Some of the properties for the commonly used core

materials are listed in Table 36.9.

FIGURE 36.15 Cross-sectional views of a particulate magnetic tape (top) and a thin film hard disk (bottom). TABLE 36.9 Relative Permeability (m), Saturation Flux Density ($B_s$), Coercivity ($H_c$) and Resistivity (r) Values of Some Commonly Used Magnetic Head Materials at Low Frequency (some values are from Reference 5) Material m $B_s$ (T) $H_c$ (A/m) r (mW ·cm) Application Ni-Fe-Mo 11000 0.8 2.0 100 Audio tape Ni-Zn 300–1500 0.4–0.46 11.8–27.6 10 11 Floppy and hard disk drives, video and instrumentation tapes Mn-Zn 3000–10000 0.4–0.6 11.8–15.8 10 6 Floppy and hard disk drives, video and instrumentation tapes Fe-Si-Al 8000 1.0 2.0 85 Floppy and hard disk drives, video and instrumentation tapes Ni-Fe 2000–4000 1.0 <10 20 Hard disk drives

The evolution of the magnetic head follows the selection of core materials, as shown in Figure 36.16.

Early heads used laminated molybdebum Permalloy (Ni-Fe-Mo, 79-17-4 wt%). These heads are inex

pensive to make, and have low $H_c$ and high m and $B_s$. The primary drawbacks are frequency limitation,

gap dimension inaccuracy, and mechanical softness. Frequency limitation is caused by the difficulty of making the lamination layer thinner than 25 mm. Eddy current loss, which is proportional to layer thickness and square root of frequency, reduces the effective permeability. As a result, laminated heads are seldom used for applications exceeding 10 MHz. Gap dimension inaccuracy is associated with the head fabrication process, and makes it unsuitable for high areal density applications. Lack of mechanical hardness reduces its usable life.

One way to reduce eddy current loss is to increase core material electric resistivity. Two types of ferrite material have high resistivity (four to nine orders higher than Permalloy) and reasonable magnetic properties: Ni–Zn and Mn–Zn. These materials are also very hard, elongating head life during head/medium contacts. The major deficiency of ferrite materials is their low $B_s$ values. In order to record in high $H_c$ media, high flux density B is needed in the head core. When the flux density in the core material reaches its saturation $B_s$, it will not increase despite the increase of recording current or coil turns. This saturation starts from the corners of the gap due to its geometry. To remedy this deficiency, a layer of

FIGURE 36.16 Schematic illustrations of (a) a laminated head, (b) cross-section of an MIG head, (c) cross-section of a thin film head, and (d) an MR sensor with leads.

metallic alloy material with much higher $B_s$ is deposited on the gap faces. This type of head is called the metal-in-gap (MIG) head. Sendust (Fe–Si–Al, 85–9.6–5.4 wt%) is one of the materials used for the deposition. MIG heads are capable of recording up to 100 MHz frequency and 180 kA/m medium coercivity.

Thin-film heads capitalize on semiconductor-like processing technology to reduce the customized

fabrication steps for individual heads. The core, coil, gap, and insulator layers are all fabricated by

electroplating, sputtering, or evaporation. Due to the nature of the semiconductor process, the fabrication

is accurate for small dimensions. Small gap dimensions are suitable for high linear and track density,

and small core dimensions allow the use of high $B_s$ Permalloy material (Ni–Fe, 80–20 wt%) as core with

low inductance for high data rate applications. Thin-film heads are used for high medium $H_c$, high areal

density applications. The high cost of the semiconductor-like process is offset by high throughput: a

150 ¥ 150 mm wafer can produce 16,000 nanoslider heads. One disadvantage is the limited-band record

ing capability because the small pole length limits low-frequency response and introduces undershoots.

A second disadvantage is the Barkhausen noise, which is caused by the relatively small number of magnetic

domains in the core. At present, thin-film heads are used up to frequencies of 80 MHz and medium

coercivity of 200 kA/m. MIG thin-film heads are also being used for high-coercivity applications.

An inductive head is often used for both recording and reproducing. The optimal performance cannot

be achieved because recording and reproducing have contradictory requirements for head design. To

solve this problem, the MR head has been developed. The MR head is for reproducing only, and an

inductive head is used for recording. As schematically shown in Figure 36.16, an MR head has a magne

toresistive element (MRE) and two electric leads. The MRE is a Permalloy stripe (Ni-Fe, 80–20 wt%),

with thickness t, width w, and height h. An electric current, with density J, passes through the MRE

through the leads. The electric resistivity of the MRE is a function of the angle q between J and MRE

magnetization M: (36.2)

where r q is the resistivity at q, r is the resistivity at q = 90˚, and Dr is the resistivity difference between

q = 0˚ and q = 90˚. Dr/r is the anisotropic MR ratio (AMR) of the MRE. Usually a transverse magnetic

field is applied to the MRE so that q = q 0 when the head is not reading a magnetic signal. Assume that

a magnetic signal from the medium rotates M from q 0 to q, then an electric voltage change v will be

detected across the MRE as the output signal: (36.3)

where q 0 is the bias angle and is set to 45˚ for good linearity. In practice, a longitudinal bias is also used

along the MRE width direction to stabilize the magnetic domain and reduce large Barkhausen noise. To

compare the output between an MR head and an inductive head, we write the inductive head output

using Faraday's equation: (36.4)

where n is the number of the coil turns, V is the medium velocity, f is the magnetic flux, and x is the

coordinate axis fixed on the medium surface. Equations 36.3 and 36.4 tell us that while inductive head

output is proportional to medium velocity and not suitable for low-velocity applications, the MR head

can be used for either high- or low-velocity applications. r r r r q q = + Ê Ë Á ˆ ¯ 1 2 D cos v Jw= ( ) r r r q q D sin sin 2 0 2 v nV x = d d f

Recording Process

We can imagine the recording process in two steps. First, the magnetic flux flowing in the head core generates a fringe magnetic field around the gap. Then the magnetic field magnetizes the magnetic medium and leaves a magnetization transition in it. Partly due to the nonlinear nature of the hysteresis loop, the recording process is so complex that there has been no rigorous explanation. However, we can still obtain significant insights into the recording process by using simple models if we keep in mind the limitations.

If we set the origin of a coordinate system at the center of the gap with x axis on the head surface and y axis pointing away from the head, then the longitudinal magnetic field $H_x$ and perpendicular magnetic field $H_y$ of this head can be expressed by the Karlqvist approximation [2]: (36.5) (36.6)

where n is the number of coil turns, i is the current in the coil, g is the gap length, l is the core length, $A_g$ is the core cross-sectional area, m is the relative permeability of core material, and $A_c$ is the gap crosssectional area. Both Equations 36.5 and 36.6 give accurate results for points 0.25g away from gap corners. Since longitudinal recording mode dominates the magnetic recording industry, we will focus on the field $H_x$ . Equation 36.5 shows that the contours of constant $H_x$ field are circles nesting on the two gap corners, as shown in Figure 36.17. The greater the diameter of the circle, the weaker the magnetic field. Assume a magnetic medium, moving from left to right with a distance d above the head, has a thickness d and a magnetization M pointing to the right. At some instant the recording current turns on and generates the magnetic field H x above the gap as depicted in Figure 36.17. On the circumference of H x = H c , half of the medium material has its magnetization reversed and half remains the same, resulting in a zero total magnetization. Since H x has a gradient, the medium closer to the gap (inside a smaller circle) gets its magnetization reversed more completely than the medium farther away from the gap (outside a bigger circle). Therefore, magnetic transition is gradual in the medium even if the change of recording current follows a step function. Assume the original magnetization is M r and the completely reversed magnetization is –M r , this gradual change of magnetization for an isolated transition can be modeled by [3]:

FIGURE 36.17 The constant horizontal fields of Karlqvist approximation are circles resting on the gap corners of a head, and the change of magnetization in the medium is gradual. H ni g lA A yg x y g x g c = + ( ) + ( ) È Î Í Í ù û ú ú p m tan 1 2 2 2 4 H ni g lA A x g y x g y y g c = + ( ) ( ) + + ( ) + È Î Í Í ù û ú ú 2 2 2 2 2 2 2 p m ln (36.7)

where x is the distance from the center of transition and a is a parameter characterizing the sharpness of the transition as shown in Figure 36.18. Assuming a thin-film medium and using the Karlqvist approximation for the head field, a is found to be [4–6]: (36.8)

where S* is the medium loop squareness and Q is the head-field gradient factor. For a reasonably well designed head, Q ª 0.8. It is obvious that we want to make parameter a as small as possible so that we can record more transitions for a unit medium length. If the head gap length g and medium thickness d are small compared with head/medium separation d, and the medium has a squareness of one, then

the minimum possible value of a is [7] (36.9)

In order to decrease the value of a and therefore increase areal density, we need to reduce medium

remanence M r , thickness d, head/medium separation d, and to increase coercivity H c .

Reproducing Process

In contrast to the recording process, the reproducing process is well understood. The flux density induced

in the head core is on the order of a few millitesia, yielding a linear process for easier mathematical

treatment. The head fringe field is the Karlqvist approximation (Equation 36.5) and the foundation is

the reciprocity theorem. For an isolated transition (Figure 36.18) with a thin magnetic layer d � d, the

induced electric voltage v for an inductive head is [7]: (36.10)

where m 0 is the permeability of vacuum, m is the relative permeability of the core, V is the medium

velocity, w is the track width, n is the number of coil turns, g is the head gap length, d is the head/medium

separation, and x is the distance between the center of the medium transition and the center of the head

FIGURE 36.18 An isolated arctangent magnetization transition from negative M r to positive M r . M M x a r = 2 1 p tan a S d Q S d Q M d QH r = ( ) + ( ) + ( ) + ( ) È Î Í Í ù û ú ú + + ( ) 1 2 1 2 2 2 * *d p d p d d p c a M H M H d M d H M H d m = ≥ < Ï Ì Ô Ô Ó Ô Ô r c r c r c r c d p d p d p d p 2 4 1 4 1 v x VwM n g lA A g x a d g x a d ( ) = + ( ) + + Ê Ë Á � �¯ + + Ê Ë Á � �¯ È Î Í ù û ú 2 2 2 0 1 1 m d p m r g c tan tan

gap. The term lA g /mA c is closely related to g for head efficiency. When a transition passes under the head,

its voltage starts with a very low value, reaches a peak, then falls off again, as shown in Figure 36.19, where the following typical values for a hard disk drive are used: V = 20 m/s, w = 3.5 mm, M r = 450 kA/m, d =

50 nm, n = 50, lA g /mA c = 0.1g. The effects of g and a + d are shown in Figure 36.19. Since a greater peak

voltage and a narrower spatial response are desired for the reproducing process, smaller g and a + d

values are helpful.

When an MR head is used for reproducing, the MRE is usually sandwiched between two magnetic

shields to increase its spatial resolution to medium signals, as shown in Figure 36.20. Since the MR head

is flux sensitive, the incident flux f i on the bottom surface of the MRE should be derived as a function

of the distance (x) between the center of MRE and the center of the transition [7, 8]: (36.11)

FIGURE 36.19 The reproducing voltage of an inductive head over an isolated arctangent transition shows the effects

of gap length g, parameter a, and head/medium separation d.

FIGURE 36.20 Schematic diagram of a shielded MR head with a shield to MRE distance g. f m d p i r x wM a d g f x g t a d f x t a d f x g t a d f x t a d ( ) = + ( ) + + ( ) + È Î Í ù û ú + ( ) + È Î Í ù û ú + Ï Ì Ô Ó Ô ( ) + È Î Í ù û ú ( ) + È Î Í ù û ú ¸ ý Ô þ Ô 2 2 2 2 2 0

where g is the distance between the MRE and the shield, t is the MRE thickness, and (36.12)

The angle between magnetization and current varies along the MRE height h. To find out the variation,

we need to calculate the signal flux decay as a function of y by (36.13)

where (36.14)

Then the bias angle q 0 and signal angle q can be calculated by (36.15)

and (36.16)

where f b is the biasing flux in the MRE and M s is the saturation magnetization of the MRE. Application

of Equations 36.15 and 36.16 to Equation 36.3 and integration over height h lead to (36.17)

For an MR head with a 45° bias at the center and small height h □ l c , the peak voltage is [6] (36.18)

The general shape of the reproducing voltage from an MR head is similar to that in Figure 36.19.

The study of an isolated transition reveals many intrinsic features of the reproducing process. However,

transitions are usually recorded closely in a magnetic medium to achieve high linear density. In this case,

the magnetization variation in the medium approaches a sinusoidal wave. That is, (36.19)

where l is the wavelength. The reproducing voltage in an inductive head becomes [9, 10] (36.20) f b b b b ( ) = + tan ln 1 2 1 f f s i c c y h y l h l ( ) = ( ) [ ] ( ) sinh sinh l gt c = m 2 sin q f 0 y y M t ( ) = ( ) b s sin q f f y y y M t ( ) = ( ) + ( ) s b s v Jw h y y y y h h = ( ) ( ) È Î Í ù û ú Ú Ú r r r q q D 1 2 0 0 2 0 sin sind d v JwM g t tgM g a d p r s ª + ( ) + ( ) È Î Í Í ù û ú ú 9 8 2 2 1 Dr d tan M x M x ( ) = r sin 2p l v x VwM ng g lA A e e g g x d ( ) = + ( ) ( ) Ê Ë Á Á Á □ ¯ □ □ □ m m p l p l p l p l pd l 0 2 2 1 2 r g c sin cos

This equation presents all the important features of the high-linear-density reproducing process. The

term exp(-2pd/l) is the spacing loss. It shows that the reproducing voltage falls exponentially with the

ratio of head/medium spacing to wavelength. The second term 1 – exp(-2pd/l) is the thickness loss. The

name of this term is misleading because its value increases with a greater medium thickness. However,

the rate of increase diminishes for thicker medium. In fact, 80% of the maximum possible value is

achieved by a medium thickness of 0.25l. The last term sin(pg/l)/(pg/l) is the gap loss. This term is based on the Karlqvist approximation. If a more accurate head fringe field is used, this term is modified

to sin(pg/l)/(pg/l)·(1.25g 2 – l 2 )/(g 2 – l 2 ) [11]. It shows a gap null at l = 1.12g, and limits the shortest

wavelength producible. These three terms are plotted in Figure 36.21. The most significant loss comes

from the spacing loss term, which is 54.6 dB for d = l. Therefore, one of the biggest efforts spent on

magnetic recording is to reduce the head/medium spacing as much as possible without causing mechan

ical reliability issues. For an MR head, the reproducing voltage is [11] (36.21)

Digital vs. Analog Recording

Due to the nonlinearity of the hysteresis loop, magnetic recording is intrinsically suitable for digital

recording, where only two states (1 and 0) are to be recognized. Many physical quantities, however, are

received in analog form before they can be recorded, such as in consumer audio and instrumentation

recording. In order to perform such recording, we need to either digitize the information or use the

analog recording technique. In the case of digitization, we use an analog-to-digital converter to change

a continuous signal into binary numbers. The process can be explained by using the example shown in

Figure 36.22. An electric signal V, normalized to the range between 0 and 1, is to be digitized into three

bits. The signal is sampled at time t = 1, 2, …, 6. At each sampling point, the first bit is assigned a 1 if

the value of the continuous signal is in the top half (>0.5), otherwise assigned a 0. The second bit is

assigned a 1 if the value of the continuous signal is in the top half of each half (0.25 £ V < 0.5, or >0.75), otherwise assigned a 0. The third bit is assigned similarly. The first bit is the most significant bit (MSB),

and the last bit is the least significant bit (LSB). The converted binary numbers are listed below each

sampling point in Figure 36.22. This process of digitization is termed quantization. In general, the final

quantization interval is

FIGURE 36.21 Spacing, thickness, and gap losses of an inductive head vs. frequency for the reproducing of a

sinusoidal medium magnetization. v M i w ht e e g g g t x r d µ ( ) ( ) Ê Ë Á Á Á ˆ ¯ ˜ ˜ ˜ + ( ) 4 1 2 2 2 Dp l p l p l p l p l p l pd l sin sin cos (36.22)

Where V is the total voltage range and N is the number of bits. Because we use a finite number of bits,

statistically there is a difference between the continuous signal and the quantized signal at the sampling

points. This is the quantization error. It leads to a signal-to-quantization-noise ratio (SNR) [12]: (36.23)

where P s is the mean square average signal power. For a signal with uniform distribution over its full

range, this yields (36.24)

For a sinusoidal signal, it changes to (36.25)

The SNR can be improved by using more bits. This improvement, however, is limited by the SNR of the

incoming continuous signal. The quantized signal is then pulse code modulated (PCM) for recording.

For analog recording, a linear relationship between the medium magnetization and the recording

signal is required. This is achieved through the anhysteretic magnetization process. If we apply an

alternating magnetic field and a unidirectional magnetic field to a previously demagnetized medium, and then reduce the amplitude of the alternating field to zero before we remove the unidirectional field,

the remanent magnetization shows a pseudolinear relationship with the unidirectional field strength H u

up to some level. Figure 36.23 shows such an anhysteretic curve. The linearity deteriorates as H u gets

greater. In applications, the recording signal current is biased with an ac current of greater amplitude

and higher frequency. Therefore, it is also termed ac-biased recording. Analog recording is easy to

implement, at the price of a lowered SNR because remanent magnetization is limited to about 30% of

the maximum possible M r to achieve good linearity.

Recording Codes

PCM is a scheme of modifying input binary data to make them more suitable for a recording and

reproducing channel. These schemes are intended to achieve some of the following goals: (1) reducing the

dc component, (2) increasing linear density, (3) providing self-clocking, (4) limiting error propagation,

and (5) achieving error-free detection. There are numerous code schemes; only three of the ones developed

early are shown in Figure 36.24. The earliest and most straightforward one is the return-to-zero (RZ)

FIGURE 36.22 Schematic illustration of the quantization of a continuous signal to three bits. DV V N = 2 SNR s = 12 2 P VD SNR = 2 2N SNR = ¥1 5 2 2 . N

code. In this scheme a positive and negative pulse is used to represent each 1 and 0, respectively, of the

data. The main drawback is that direct recording over old data is not possible due to the existence of

zero recording current between two data. It also generates two transitions for each bit, therefore reducing the linear density. In addition, it only uses half of the available recording current range for a transition.

The non-return-to-zero-invert (NRZI) method was developed to alleviate some of these problems. It

changes the recording current from one direction to the other for each 1 of the data, while making no

changes for all 0s. However, it has a strong dc component and may lose reproducing synchronization if

there is a long string of 0s in the input data. In addition, reproducing circuits are usually not designed

for dc signal processing. In frequency modulation (FM) code there is always a transition at the bit-cell

boundary which acts as a clock. There is also an additional transition at the bit-cell center for each 1

and no transition for 0s. It reduces the dc component significantly. The primary deficiency is the reduction

of linear density since there are two transitions for each 1 in the data.

The most popularly used codes for magnetic recording are the run-length-limited (RLL) codes. They

have the general form of m/n(d, k). In these schemes, data are encoded in groups. Each input group has

m bits. After encoding, each group contains n bits. In some schemes multiple groups are coded together.

d and k are the minimum and maximum 0s, respectively, between two consecutive 1s in the encoded

sequence. While d is used to limit the highest transition density and intersymbol interference, k is

employed to ensure adequate transition frequency for reproducing clock synchronization. The encoding

is carried out by using a lookup table, such as Table 36.10 for a 1/2(2,7) code [13].

Head/Medium Interface Tribology

As expressed in Equation 36.20, the most effective way to increase signal amplitude, therefore areal density,

is to reduce head/medium spacing d. However, wear occurs when a moving surface is in close proximity

to another surface. The amount of wear is described by Archard's equation:

FIGURE 36.23 An anhysteretic remanent magnetization shows a pseudolinear relationship with the applied unidi

rectional magnetic field to some H u level.

FIGURE 36.24 Comparison of some early developed codes.
(36.26)

where V is the volume worn away, W is the normal load, s is the sliding distance, H is the hardness of

the surface being worn away, and k is a wear coefficient. In order to increase medium hardness H, hard

$Al_2O_3$ particles are dispersed in particulate media and a thin layer of hard carbon (ª10 nm) with either

hydrogenation or natrogenation is coated on thin-film media of hard disks. A layer of liquid lubricant,

typically perfluoropolyethers with various end groups and additives, is applied on top of the carbon film

to reduce the wear coefficient k. Load is minimized to reduce wear while keeping adequate head/medium

dynamic stability. For applications where the sliding distance s is modest over the lifetime of the products

such as floppy disk drives and consumer tapes drives, the head physically contacts the medium during

operations. In the case of hard disk application, heads are separated nominally from the media by a layer

of air cushion. The head is carried on a slider, and the slider uses air-bearing surfaces (ABS) to create

the air film based on hydrodynamic lubrication theory. Figure 36.25 shows two commonly used ABS.

Tapers are used to help the slider take off and maintain flying stability. ABS generates higher-than-ambient

pressure to lift the slider above the medium surface during operations. The tripad slider is for pseudo

contact applications while the subambient-pressure (SAP) slider is for flying (such as MR head) applications.

Because the relative linear velocity between the slider and the medium changes when the head moves to

different radii, a cavity region is used in the SAP slider to generate suction force to reduce flying height

variation. The ABS is designed based on the modified Reynolds equation: (36.27) TABLE 36.10 1/2(2,7) Code Before Coding After Coding 10 0100 11 1000 000 000100 010 100100 011 001000 0010 00100100 0011 00001000

FIGURE 36.25 The ABS of (a) a tri-pad slider for pseudocontact recording and (b) a SAP slider for conventional

flying recording. V k Ws H = ∂ ∂ ∂ ∂ Ê Ë Á ˆ ¯ + ∂ ∂ ∂ ∂ Ê Ë Á ˆ ¯ = ∂ ( ) ∂ + ∂ ( ) ∂ + ∂ ( ) ∂ Ê Ë Á ˆ ¯ X PH Q P X Y PH Q P Y PH X PH Y PH T x y 3 3 L L s

where X and Y are coordinates in the slider longitudinal and transverse directions normalized by the

slider length and width, respectively, P is the hydrodynamic pressure normalized by the ambient pressure,

H is the distance between the ABS and medium surface normalized by the minimum flying height, Q is

the molecular slip factor, T is time normalized by the characteristic squeeze period, L x and L y are the

bearing numbers in the x and y directions, respectively, and s is the squeeze number. A full derivation

and explanation of the Reynolds equation can be found in Reference 14. At present, high-end hard disk

drives feature a flying height on the order of 20 to 50 nm.

When power is turned off, the slider in the popularly used Winchester-type drives rests on the medium surface. Although the ABS and medium surface look flat and smooth, they really consist of microscopic

peaks and valleys. If we model an ABS/medium contact by a flat surface and a sphere tip, the liquid

lubricant on the medium surface causes a meniscus force F m as depicted in Figure 36.26 [15]: (36.28)

where R is the radius of the sphere, g is the surface tension of the lubricant, q is the contact angle between

the lubricant and the surfaces, y is the sphere to flat surface distance, and h is the lubricant thickness.

Detailed analysis [16] shows that the static friction F at a head/medium interface is a function of several

parameters: (36.29)

where A is the ABS area, h is the peak density, E Φ is the effective modulus of elasticity, f is the peak height

distribution, s is the rms peak roughness, and s is the solid-to-solid shear strength. If friction F is too

large, either the drive cannot be started or the head/medium interface is damaged. While friction can be

reduced practically by reducing A, g, and increasing q, the most effective ways are to control h, s, h, and

f. Too thin a lubricant layer will cause wear, and too thick will induce high friction. This limits h to the

range of 1 to 3 nm. s is controlled by surface texture. Historically, texture is created by mechanical

techniques using either free or fixed abrasives. This leaves a surface with a random feature and is unsuitable

for controlling h and f. Recently, people started to use lasers [17]. This technique generates a surface

texture with well-defined h and f to improve wear and friction performance. Figure 36.27 shows AFM

images of a mechanical and a laser texture.

Optical Recording

The major obstacle to achieving higher areal density in magnetic recording is the spacing loss term. It is

a great engineering challenge to keep heads and media in close proximity while maintaining the head/

medium interface reliable and durable. Care must also be taken in handling magnetic media since even

FIGURE 36.26 Formation of meniscus between a sphere tip and a flat surface. F R y h y m = + ( ) 4 1 p g qcos F f h R A E s= Φ ( ) , , , , , , , , ,h g q f s

minute contamination or scratches can destroy the recorded information. In addition, the servo technique

of using magnetic patterns limits the track density to about one order lower than the linear density.

Optical recording, on the other hand, promises to address all these concerns.

Optical recording can be categorized into three groups. In the first group, information is stored in the

media during manufacturing. Users can reproduce the information, but cannot change or record new

information. CD-ROM (compact disk-read only memory) belongs to this group. The second group is

WORM (write once read many times). Instead of recording information during manufacturing, it leaves

this step to the user. This is usually achieved by creating physical holes or blisters in the media during

the recording process. Once it is recorded, however, the medium behaves like the first group: no further

recording is possible. The third group is similar to magnetic recording. Recording can be performed

infinitely on the media by changing phase or magnetization of the media. The most noticeable example

FIGURE 36.27 AFM images of (a) a mechanical texture and (b) a laser texture. (Courtesy of J. Xuan.)

in this group is the magneto-optic (MO) technique [18]. Only CD-ROM and the magneto-optic recording

are described in the following.

CD-ROM

Figure 36.28 shows the CD-ROM reproducing principle. Data are pressed as physical pits and lands on

one surface of a plastic substrate, usually polycarbonate. A layer of aluminum is deposited on this surface

to yield it reflective. Lacquer is then coated to protect the aluminum layer. During the reproducing

process, an optical lens is used to focus a laser beam on the reflective pit and land surface. The diameter

of the lens is D, the distance between the lens and the substrate is $h_3$, and the substrate thickness is $h_2$.

The diameter of the laser beam is $d_2$ when entering the substrate, and becomes $d_1$ when focused on the

reflective surface. The width of the pits are designed smaller than $d_1$. The reflected light consists of two

portions: $I_1$ from the land and $I_2$ from the pit. According to the theory of interference, the intensity of

the reflected light is (36.30)

where l is the wavelength of the laser and $h_1$ ª l/4 is the pit height. This leads to (36.31)

This change of light intensity is detected and decoded to yield the recorded data. The reflected light is

also used for focusing and track following.

The fundamental limit on optical recording density is the focused beam diameter $d_1$. For a Gaussian

($TEM_{00}$) laser, this is the diffraction-limited diameter at which the light intensity is reduced to $1/e^2$ of

the peak intensity: (36.32)

where q is the aperture angle. The following values are typical for a CD-ROM system: l (gallium arsenide

laser) = 780 nm, q = 27°, h 2 = 1.2 mm, D = 5 mm, h 3 = 4.2 mm. This yields d 1 ª 1.0 mm and d 2 ª 0.7 mm,

and sets the areal density limit of optical (including magneto-optic) recording to about 1 Mbit/mm 2 . For

most CD-ROM applications, the areal density is smaller than this limit, and a disk with 120 mm diameter

holds about 600 Mbyte information. In order to increase areal density, we can either reduce light

wavelength or increase numerical aperture. Much of the effort has been to adopt a new light source with

FIGURE 36.28 Schematic representation of the CDROM reproducing principle. I I I I I h = + + 1 2 1 2 1 2 4 cos p l I I I I I h I I I I h = + = ( ) + + = ( ) Ï Ì Ô Ó Ô 1 2 1 2 1 1 2 1 2 1 2 4 2 0 if there is a pit if there is no pit l d 1 2 ª l pq

short wavelength such as a blue laser. Increasing numerical aperture is more difficult because increasing

lens diameter is cost prohibitive and reducing h 2 or h 3 is reliability limited. Note that although the beam

size on the focus plane is on the order of 1 mm (d 1 ), it is two to three orders greater at the air/substrate

interface (d 2 ). This means that optical recording can tolerate disk surface contamination and scratches

much better than magnetic recording. However, the performance of optical recording does not match

magnetic recording in general. The data transfer rate of CD-ROM drives is expressed as multiple (¥) of

150 kB/s. Even for a 12¥ CD-ROM drive, the data access time and data transfer rate are still on the order

of 100 ms and 1.8 MB/s, respectively, while for a high-performance rigid disk drive these values are less

than 10 ms and greater than 30 MB/s, respectively.

Magnetooptic Recording

The primary drawback of a CD-ROM to an end user is its inability to record. This deficiency is remedied

by MO recording technology, as depicted in Figure 36.29. A linearly polarized laser beam is focused on

a layer of magnetic material, and a coil provides a dc magnetic field on the other side of the medium.

This dc magnetic field is too weak to affect the medium magnetization at normal temperature. The

recording process utilizes the thermomagnetic property of the medium, and the reproducing process is

achieved by using the Kerr effect. During recording, the medium is initially magnetized vertically in one

direction, and the dc magnetic field is in the opposite direction. The laser heats up the medium to its

Curie temperature, at which the coercivity becomes zero. During the cooling process, the dc magnetic

field aligns the medium magnetization of the heated region to the magnetic field direction. In the process

of reproducing, the same laser is used with a smaller intensity. The medium is heated up to its compen

sation temperature, at which the coercivity becomes extremely high. Depending on the direction of the

magnetization, the polarization of the reflected light is rotated either clockwise or counterclockwise (Kerr

rotation). This rotation of polarization is detected and decoded to get the data. The main disadvantage

of MO recording is that a separate erasing process is needed to magnetize the medium in one direction

before recording. Recently some technologies have been developed to eliminate this separate erasing

process at the cost of complexity.

The medium used in MO recording must have a reasonable low

Curie temperature (<300˚C). The materials having this property are rare earth transition metal alloys, such as $Tb_{23}Fe_{77}$ and $Tb_{21}Co_{79}$.

Unfortunately, the properties of these materials deteriorate in an oxygen and moisture environment. To protect them from air and humidity, they are sandwiched between an overlayer and a underlayer, such as SiO, AlN, SiN, and $TiO_2$. Another issue with the rare earth transition metal alloys is their small Kerr rotation, about 0.3˚. To increase this Kerr rotation, multiple layers are used. In the so-called quadrilayer structure (Figure 36.29b), the overlayer is about a half-wavelength thick and the underlayer is about a quarter-wavelength thick [18]. The MO layer is very thin (ª3 nm). Light reflected from the reflector is out-of-phase with the light reflected from the surface of the MO layer, and is in-phase with the light reflected from the inside of the MO layer. As a result, the effective Kerr rotation is increased several times.

FIGURE 36.29 Schematic illustrations of (a) MO recording/reproducing and (b) quadrilayer medium cross section.

Compared with magnetic recording, optical recording has the intrinsic advantages of superior reli ability and portability. However, its performance is inferior due to slower data access time and transfer rate. Another advantage of optical recording, higher areal density, has been disappearing or even reversing to magnetic recording. Both magnetic and optical recording will be continuously improved in the near future, probably toward different applications. Currently there are some emerging techniques that try to combine the magnetic and optical recording techniques. Table 36.11 is a short list of representative magnetic and optical devices for digital recording.

1. C.D. Mee and E.D. Daniel, Eds., Magnetic Storage Handbook, 2nd ed., New York: McGraw-Hill, 1996.

2. O. Karlqvist, Calculation of the magnetic field in the ferromagnetic layer of a magnetic drum, Trans. R. Inst. Technol. (Stockholm), 86, 3–28, 1954.

3. J.J. Miyata and R.R. Tartel, The recording and reproduction of signals on magnetic medium using saturation-type recording, IRE Trans. Elec. Comp., EC-8, 159–169, 1959.

4. M.L. Williams and R.L. Comstock, An analytical model of the write process in digital magnetic recording, A.I.P. Conf. Proc., 5, 738–742, 1972.

5. V.A.J. Maller and B.K. Middleton, A simplified model of the writing process in saturation magnetic recording, IERE Conf. Proc., 26, 137–147, 1973.

6. H.N. Bertram, Theory of Magnetic Recording, Cambridge, UK: Cambridge University Press, 1994.

7. C.D. Mee and E.D. Daniel, Eds., Magnetic Recording Technology, 2nd ed., New York: McGraw-Hill, 1996.

8. R. Potter, Digital magnetic recording theory, IEEE Trans. Magn., MAG-10, 502–508, 1974.

9. W.K. Westmijze, Studies on magnetic recording, Philips Res. Rep., 8, 161–183, 1953.

10. G.J. Fan, A study of the playback process of a magnetic ring head, IBM J. Res. Dev., 5, 321–325, 1961.

11. J.C. Mallinson, The Foundations of Magnetic Recording, 2nd ed., San Diego, CA: Academic Press, 1993.

12. M.S. Roden, Analog and Digital Communication Systems, 4th ed., Upper Saddle River, NJ: PrenticeHall, 1996.

13. P.A. Franaszek, Run-length-limited variable length coding with error propagation limitation, U.S. Patent No. 3, 689, 899, 1972.

14. W.A. Gross, L. Matsch, V. Castelli, A. Eshel, T. Vohr, and M. Wilamann, Fluid Film Lubrication, New York: Wiley,

1980.

15. J.N. Israelachvili, Intermolecular and Surface Forces, London: Academic Press, 1985, 224. TABLE 36.11 Digital Magnetic and Optical Storage Devices Description Manufacturers Approximate Price, $ Thin-film head for hard disk drive AMC, Read-Rite, SAE 6.00–9.00 MR head for hard disk drive AMC, Read-Rite, SAE, Seagate 8.00–12.00 Thin-film hard disk Akashic, HMT, Komag, MCC, Stormedia 7.00–10.00 Hard disk drive IBM, Maxtor, Quantum, Samsung, Seagate, WD 0.02–0.20/Mbytes

Floppy drive Panasonic, Sony 20.00–40.00

Floppy disk 3M, Fuji, Memorex, Sony 0.15–0.50

Removable rigid disk drive Iomega, Syquest 100.00–400.00 Removable rigid disk Iomega, Maxell, Sony 5.00–20.00/100 Mbytes Tape drive Exabyte, HP, Seagate 100.00–400.00 Backup tape 3M, Sony, Verbatim 4.00–25.00/Gbytes 8 ¥ CD-ROM drive Goldstar, Panasonic 100.00–200.00 Recordable optical drive JVC, Philips 300.00–500.00 Recordable optical disk 3M, Maxell, Memorex 3.00–15.00/650 Mbytes

16. Y. Li and F.E. Talke, A model for the effect of humidity on stiction of the head/disk interface, Tribology Mech. Magn. Storage Syst., SP-27, 79–84, 1990.

17. R. Ranjan, D.N. Lambeth, M. Tromel, P. Goglia, and Y. Li, Laser texturing for low-flying-height media, J. Appl. Phys., 68, 5745–5747, 1991.

18. J. Watkinson, The Art of Data Recording, Oxford U.K.: Focal Press, 1994.

19. G.A.N. Connell and R. Allen, Magnetization reversal in amorphous rare-earth transition-metal alloys: TbFe, Proc. 4th Int. Conf. on Rapidly Quenched Metals, Sendai, Japan, 1981.