NISTIR 4583

MEASUREMENTS FOR COMPETITIVENESS IN ELECTRONICS

First Edition

Prepared by the Electronics
and Electrical Engineering
Laboratory

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards and
Technology
Electronics and Electrical Engineering
Laboratory

April 1993

NIST SEEKS YOUR COMMENTS

NIST seeks your comments on measurement needs important to the competitiveness of the U.S. electronics industry. This document will be updated periodically to reflect the latest information on those measurement needs. Please address your comments to:

Managing Editor
Measurements for Competitiveness in Electronics
Electronics and Electrical Engineering Laboratory
National Institute of Standards and Technology
Building 220 - Room B358
Gaithersburg, MD 20899
Telephone: (301) 975-2220
FAX: (301) 975-4091

NISTIR 4583

MEASUREMENTS FOR COMPETITIVENESS IN ELECTRONICS

First Edition

Prepared by the Electronics
and Electrical Engineering
Laboratory

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards and
Technology
Electronics and Electrical Engineering
Laboratory

April 1993

U.S. DEPARTMENT OF COMMERCE
Ronald H. Brown, Secretary

NATIONAL INSTITUTE OF STANDARDS AND
TECHNOLOGY
Raymond G. Kammer, Acting Director

Bibliographic Information

Abstract

Measurements are used to determine the values of hundreds of important quantities in the electronics industry. Representative quantities are the widths of the interconnections within semiconductor integrated circuits, the attenuation of lightwaves in optical fibers, and the signal power from microwave satellite antennas. Measurement capability is a fundamental tool used to build the nation's high-technology products. As such, it is part of the national infrastructure for the realization of these products.

Measurement capability is critical to research and development, manufacturing, marketplace entry, and after-sales support of products. Thus measurement capability affects the performance, quality, reliability, and cost of products. The result of this pervasive impact is that the level of U.S. measurement capability places an upper limit on the competitiveness of U.S. products.

At present, U.S. industry is experiencing a major shortfall in the measurement capability needed for competitiveness in electronic products. This document identifies the measurement needs that are most critical to U.S. competitiveness, that would have the highest economic impact if met, and that are the most difficult for the broad range of individual companies to address. The measurement needs are reviewed for nine important fields of electronics, including semiconductors, magnetics, superconductors, microwaves, lasers, optical-fiber communications, optical-fiber sensors, video, and electromagnetic compatibility. These fields of electronics underlie more than $300 billion of electronic and electrical products manufactured in the U.S. each year.

This assessment provides the framework for an action plan to correct the shortfall in U.S. measurement capability in electronics and to advance U.S. competitiveness.

Keywords

electromagnetic compatibility; lasers; magnetics; measurements; microwaves; optical-fiber communications; semiconductors; U.S. competitiveness; video

Ordering

Copies of this document are available as Order No. PB93-160588 from the National Technical Information Service, 5285 Port Royal Road, Springfield, Virginia 22161, at (800) 553-6847 or at (703) 487-4650. The document contains 478 pages, cover to cover.

April 1993

Preface

BUILDING THE NATIONAL MEASUREMENT INFRASTRUCTURE FOR COMPETITIVENESS IN ELECTRONICS

The U.S. can strengthen its competitive position in world markets for electronic products by improving the measurement capability that supports the U.S. electronics and electrical-equipment industries. Directly affected are $300 billion of new products and 2 million high-paying jobs.

Measurements are a critical element of the U.S. infrastructure. They are part of the *tools* that U.S. industry uses to build the products of high technology. Industry cannot reliably build products whose performance and quality are better than its measurement capability. Therefore, the level of the nation's measurement capability limits the competitiveness of its high-technology products. At present, the U.S. is experiencing a major shortfall in the measurement capability needed for competitiveness in electronic and electrical products. The industries that make these products need improved measurement capability for hundreds of measured quantities. Each of these quantities is critical to the success of one or more important electronic or electrical technologies. Here are some examples:

> New generations of *semiconductor* integrated circuits that are smarter, faster, and cheaper cannot be made without better dimensional measurements of the microscopic circuit elements that they contain. Semiconductor manufacturing is so measurement intensive that the semiconductor industry estimates that 30 percent of the cost of manufacturing complex integrated circuits is in testing.
>
> The performance of *optical-fiber communications* systems for a national information highway, including fiber connections for individual homes and businesses, cannot be optimized without better measurements of the critical dimensions of precision connectors required for hundreds of millions of interconnections.
>
> *Microwave* integrated circuits for new commercial markets in personal communications services and intelligent vehicle-highway systems cannot be realized without precise measurements of microwave power losses in the circuits.
>
> Path-breaking applications of *lasers* in efficient manufacturing and in less-expensive and less-invasive medical procedures cannot be realized without better measurements for the quality of laser beams.

In fact, new measurement capability is needed across electronic and electrical technologies broadly to support every one of the four majors steps required for realizing competitive products.

> *Research and Development*: Scientists and engineers use measurements to probe and understand fundamental phenomena and to apply them to practical ends. They cannot fully exploit what they cannot "see" through the measurement process.

> *Manufacturing*: Engineers use measurements to design and control new manufacturing processes that reduce manufacturing costs and improve product quality. New management strategies for improving productivity and quality are sensitively dependent on measurement capability.
>
> *Marketplace Exchange*: Industry uses measurements to establish voluntary industry standards for the marketplace. These standards promote performance, quality, and compatibility in products. They create markets by improving buyer confidence. Industry also uses measurements to gain access to markets by proving that products conform with national and international standards.
>
> *After-Sales Support*: Measurements support the installation, maintenance, and operation of high-technology products to assure maximum economic benefit and continued customer satisfaction.

The shortfall in needed measurement capability is a long-standing one and has been noted repeatedly over the years by industry spokespersons. Here are some examples from the chapters in this document, with appropriate page references for additional information.

> The Semiconductor Industry Association has cited the shortfall in measurement capability as one of the most important problems impeding the progress of this industry (page 79).
>
> The National Conference of Standards Laboratories (NCSL), with broad participation from U.S. high technology companies and Government agencies, has surveyed, documented, and called for resolution of the measurement shortfall that is impeding U.S. progress in many technologies, such as microwaves, optical fibers, and lasers (pages 162, 265, and 202).
>
> The American National Standards Institute (ANSI), the American Society for Testing and Materials (ASTM), the Magnetic Materials Producers Association (MMPA), and the Institute of Electrical and Electronics Engineers (IEEE) have all voiced the need for improved measurement capability to support the U.S. magnetics industry (page 114).

In response, this document identifies the broad range of measurement needs that must be addressed to strengthen U.S. competitiveness in electronics. The focus here is on measurement capability that is *needed widely* in the U.S. electronics industry, that will have especially *high impact* if provided, but that is *beyond the resources* of the broad range of individual companies to develop.

This assessment has been developed by NIST staff in consultation with U.S. industry. That consultation has taken many forms including individual interactions, workshops, surveys, and reviews of draft chapters of this document.

This document has two primary purposes:

> Show the close relationship between U.S. measurement infrastructure and U.S. competitiveness, and show why improved measurement capability offers such high economic leverage.
>
> Provide a consensus on the principal measurement needs affecting U.S. competitiveness, as the basis for an *action plan* to meet those needs and to improve U.S. competitiveness.

This document has been prepared as part of NIST's responsibility to take due care of the nation's measurement health.

April 1993

Guide

ORGANIZATION AND CONTENT OF THIS DOCUMENT

THE CHAPTERS

While this document contains a lengthy analysis, you can find your way in it easily with the information provided here. This document contains 12 chapters. They are divided into two groups. The first three chapters are introductory in nature and are relevant to all of the following chapters. The remaining nine chapters address individual fields of electronic technology.

Each chapter begins with a two-page summary that provides ready access to the major points made in the chapter. These short summaries are found on the pages shown below. By selecting from these summaries, you can quickly access information on the subjects of most interest to you. For this reason, a separate summary has not been provided at the beginning of this document.

Introductory Information Page

Fields of Technology

For a quick overview, consider reading the summaries for Chapters 1, 2, and 3. Then read the summaries of any of the remaining chapters that address fields of technology of interest to you. A brief statement of the coverage of each chapter and appendix follows.

Introductory Information

The first three chapters introduce the subject of measurements and provide an overview of the products of the U.S. electronics and electrical-equipment industries.

Chapter 1, **Role of Measurements in Competitiveness**, shows why measurements are a fundamental part of the infrastructure of the nation. Chapter 1 also sets measurements in the context of the many other important factors that affect competitiveness.

Chapter 2, **NIST's Role in Measurements**, indicates the circumstances under which Government assistance to industry in the development of measurement capability is appropriate in pursuit of a strengthened national economy.

Chapter 3, **Overview of U.S. Electronics and Electrical-Equipment Industries**, introduces these industries through an overview of their major product lines. This chapter shows the various ways in which the products of these industries are commonly classified and how those classifications relate to the structure of this document.

Fields of Technology

The nine following chapters address individual fields of technology and have been arranged in a special order. The technologies on which most other technologies depend are introduced first. Thus the chapter on semiconductors appears first because most electronic technologies depend on semiconductor materials. In contrast, the chapter on video is located near the end because it depends on nearly every other technology discussed earlier.

Chapters 4, 5, and 6 of this document describe the measurement needs arising from three important materials technologies that underlie current and emerging electronic and electrical products. These chapters also describe the measurement needs of components and equipment based on these materials and not discussed separately in other chapters.

Chapter 4, **Semiconductors**, addresses both silicon and compound semiconductors and their use in components, including individual (discrete) electronic and optoelectronic devices and integrated circuits. Semiconductor components are central to all modern electronic products from consumer products to supercomputers.

Chapter 5, **Magnetics**, focuses on both magnetic materials and the components made from them. Magnetic materials are second in importance only to semiconductor materials for electronic products and play a central role in electrical products. This chapter also addresses the measurement needs of selected equipment critically dependent on magnetic materials, including magnetic information storage equipment, electrical power transformers, and others.

Chapter 6, **Superconductors**, examines superconductor materials and addresses both present and emerging applications of these materials in electronic and electrical products.

Chapters 7 through 11 describe the measurement needs associated with selected technologies of importance to U.S. competitiveness for current and emerging products.

Chapter 7, **Microwaves**, describes the highest-information-capacity radio technology. Microwave electronics provide the basis for modern and emerging wireless communications systems and radar systems. Included are new personal communications services with both local and worldwide access, intelligent vehicle-highway systems, and advanced audio and video broadcasting systems, among others.

Chapter 8, **Lasers**, addressed the single most important component for emerging lightwave systems used for manufacturing, medicine, communications, printing, environmental sensing, and many other applications.

Chapter 9, **Optical-Fiber Communications**, describes the highest-information-capacity cable technology. It provides the basis for national and international information highways of unprecedented performance and broad economic impact. Optical-fiber systems will be linked with microwave systems to interconnect mobile and portable users and to backup cable systems.

Chapter 10, **Optical-Fiber Sensors**, focuses on an emerging class of sensors that offers outstanding performance for a broad spectrum of applications in manufacturing, aerospace, medicine, electrical power, and other areas.

Chapter 11, **Video**, emphasizes advanced, high-performance systems, such as high-definition television, which offer, for the first time, simultaneous access to high-resolution, smooth motion, and great color depth. The chapter notes the potential of full-power implementations of video technology in interactive networked environments. The chapter contains a special focus on flat-panel displays.

Finally, Chapter 12, **Electromagnetic Compatibility**, describes the special challenges that the U.S. faces in maintaining electromagnetic compatibility among the many new products of electronic and electrical technologies. Such compatibility is essential if the full potential of all of the above technologies is to be realized without debilitating mutual interference.

Content of Chapters

Each of Chapters 4 through 12 contains four basic types of information:

Technology Review: The field of technology is reviewed to highlight and explain the special capabilities that make the technology important. This review introduces the technical concepts that are necessary for understanding the sections that follow.

World Markets and U.S. Competitiveness: The economic significance of the field of technology is highlighted through use of national and international market data for major products that employ the technology. Available information on the U.S. competitiveness is described.

Goals of U.S. Industry for Competitiveness: The goals that U.S. industry is pursuing to improve its competitiveness are discussed so that they can be related to requirements for new measurement capability supportive of the goals.

Measurement Needs: The new measurement capability that U.S. industry will need to enable it to achieve its goals is described. This discussion emphasizes measurement capability that is needed widely in U.S. industry, that will have high economic impact if provided, and that is beyond the resources of the broad range of individual U.S. companies to provide.

While the assessment of measurement needs in this document is wide ranging, not every field of technology important to the electronic and electrical-equipment industries has been covered. NIST plans to expand this assessment in future editions to include additional fields.

THE APPENDICES

The appendices provide definitions of the U.S. electronics and electrical-equipment industries. These definitions were used in preparing much of the economic information in the report.

> Appendix 1 describes the Standard Industrial Classification System that the U.S. Government uses for collecting data about U.S. industry. This appendix also lists publications in which the U.S. Government reports data on U.S. shipments.
>
> Appendix 2 provides a definition of the U.S. electronics industry in terms of the Standard Industrial Classification System.
>
> Appendix 3 provides a definition of the U.S. electrical-equipment industry in terms of the Standard Industrial Classification System.

TREATMENT OF ECONOMIC DATA

Certain standard practices have been followed in the treatment of economic data in this document to make that data easier to interpret.

Current Dollars

All data on markets and production are expressed in current dollars, unless constant dollars are specified. As a reminder, *current dollars* are dollars valued in the year for which the economic data are provided; they include the effects of inflation. *Constant dollars* can be valued in any year that precedes or follows the year for which the economic data are provided; they are adjusted to exclude the effects of inflation. Whenever constant dollars are used, the reference year is provided.

Nominal Growth Rates

Similarly, all growth rates are expressed in nominal terms, unless real terms are specified. *Nominal growth rates* describe the growth rate in current dollars (including inflation). *Real growth rates* describe the growth rate in constant dollars (adjusted to exclude inflation). Since the world generally experiences inflation rather than deflation, nominal growth rates are generally greater than real growth rates. The relationship among these quantities is this: nominal growth rate = real growth rate + inflation rate.

DATES OF PREPARATION

This document required two years for development. Each chapter or appendix was completed on the date shown on its first page. That date will provide you with a reference point in time for the economic and technological information in the chapter or appendix. These dates may be compared with the dates appearing on updated chapters or appendices in future editions.

TABLE OF CONTENTS

Chapter 9 OPTICAL-FIBER COMMUNICATIONS 217

Chapter 10 OPTICAL-FIBER SENSORS 303

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

ROLE OF MEASUREMENTS IN COMPETITIVENESS

June 1991

Chapter 1

ROLE OF MEASUREMENTS IN COMPETITIVENESS

SUMMARY

The level of measurement capability of the nation sets an upper limit on the competitiveness of its high-technology products. The effects of measurement capability on competitiveness are manifested in many ways. For example, manufacturers of high-technology products need a high level of measurement capability to:

- develop products with optimized performance, quality, and compatibility
- design and control manufacturing processes
- gain access to markets by proving compliance with marketplace standards
- reach agreement with buyers on product specifications and performance
- support products after sale

A high level of measurement capability is thus *one necessary factor* in achieving international competitiveness in high technology. There are many *other necessary factors*. Representative of these are:

- access to resources: capital, labor, equipment, technology
- protection of resources: patents, copyrights
- access to international markets: export and import restrictions, both tariff and non-tariff

All of these *other necessary factors* also present formidable barriers that must be overcome. Once they are overcome, however, the bottom line is this: manufacturers must deliver a product that is competitive on its own merits, and their ability to do this is highly sensitive to the level of measurement capability.

Measurement capability plays a special role in a free-market economy. A free-market economy relies on the ability of the marketplace to reach decisions that maximize economic efficiency. In areas of high technology, measurement capability is used by manufacturers and buyers to determine the performance, quality, and compatibility of products as a basis for marketplace decisions. Those decisions can be no better than the level of measurement capability on which they are based. This role is so critical that *the level of measurement capability sets an upper limit on the ability of a free-market economy to reach economically efficient decisions on high-technology products.*

Measurement capability is a *generic tool* that manufacturers use to make the products that they choose. The economic impact of measurement capability is high because measurements support every step that a manufacturer must complete to realize competitive products: research and development, manufacturing, marketplace exchange, and after-sales support. Because of this pervasive influence, measurement capability is often referred to as a *supporting technology* or as an *infratechnology*, that is, a technology that is essential to the development and utilization of other technologies. These terms reflect the fact that measurement capability has high leverage on the ability of other technologies to deliver.

Further, a high level of measurement capability is essential to the implementation of new management strategies for improving competitiveness, whenever the strategies contain technical elements. Examples of these strategies include total quality management, flexible manufacturing, joint technology development, and concurrent engineering. Measurement capability supports these strategies by improving: (1) the *control* that manufacturers exercise over all of the technical steps required to realize competitive products; and (2) the effectiveness of the technical *language* that manufacturers use in executing these strategies, particularly when collaboration or technology transfer are required.

These several roles give measurement capability *high leverage* on economic growth and international competitiveness. They also underscore the risks to competitiveness from underinvestment in the development of advanced measurement capability.

INTRODUCTION

Purpose of This Chapter

The purpose of this chapter is to show explicitly how measurement capability affects the competitiveness of high-technology products. The discussion is applicable to all high-technology industries. However, the examples provided are drawn from the electronics industry. The examples are set in small type.

Approach in This Chapter

Ultimately, the competitiveness of a product is determined by the decisions that a buyer makes in the marketplace. Therefore, this discussion begins by examining what the buyer demands in order to judge a product competitive. The discussion then moves backwards to the manufacturer in order to identify the challenges that the manufacturer must overcome to meet the buyer's demands for competitiveness. The manufacturer's challenges are examined in three parts relating to: (1) developing, manufacturing, marketing, and supporting the product; (2) raising and protecting required resources; and (3) implementing new management strategies for improved competitiveness. Throughout this discussion, the focus is on the role of measurements. The role of other factors important to competitiveness will be noted but will not be elaborated, since those factors are discussed well elsewhere. This omission is not intended to minimize the significance of other factors; in fact, their significance may vary from important to dominant.

Definitions

A few definitions at the outset will be clarifying. "Measurement capability", as used here, refers principally to *measurement methods* that are used for determining correct values of important quantities. Thousands of measurement methods are required for high-technology industries like electronics.

Here are examples of important quantities that require special measurement methods:

- radiation patterns of satellite antennas for telecommunications
- widths of interconnecting electrical lines in semiconductor integrated circuits for computers
- loss of strength of lightwaves as they pass through optical fibers in cross-country telephone lines
- sizes of microscopic magnetic domains that store information in rigid disks for computers
- current-carrying capacity of a superconducting wire for electromagnets in medical magnetic-resonance-imaging equipment

Measurement methods are supported by *measurement reference standards*, which are also an important part of "measurement capability". Measurement reference standards are physical objects, such as specially prepared materials or electronic devices. Measurement reference standards support the accuracy of measurement methods by providing stable, defined reference values against which unknown values are compared during the measurement process.

Measurement reference standards should be distinguished from *written standards*. Written standards are also important to competitiveness. Written standards are statements that specify a concept, procedure, or goal that is to be widely accepted. For example, a written standard for a *procedure* might specify how a measurement must be made; or a written standard for a *goal* might specify the minimum level of performance that an electronic product must attain. In most cases, when people speak of *standards*, they are referring to *written standards*, rather than to measurement reference standards.

Written standards are usually classified by three criteria, as shown in Table 1: (1) the subject specified in the standard, (2) the entity promulgating the standard, and (3) the elective nature of the standard. Several examples of the subjects specified in standards are shown in the table; there are many others. The meanings of the examples are probably clear, except perhaps for a compatibility standard. It specifies the properties that a product must have to be compatible with other products with which it will be used.

For example, one compatibility standard specifies the exact diameter that an optical fiber must have to fit precisely into a connector that joins two pieces of optical fiber end to end.

Sometimes a given written standard is incorporated into other written standards by reference; such incorporation is quite common for measurement standards. For example, a performance standard may specify that the measurement method described in a certain measurement standard must be used to determine if the performance goals have been met.

Written standards may be promulgated by different entities. *Industry* issues many standards. *Government organizations*, both national and local, issue standards. *International organizations*, including international regulatory bodies, also issue standards.

Written standards may be *voluntary* or *mandatory*. In the United States, most written standards used in high-technology fields are voluntary standards. Most of these are created by organizations established by industry itself; they are therefore called *voluntary industry standards*. National governments also create voluntary standards, as do international standards bodies with the participation of member countries. Governments at many different levels create *mandatory* standards at their levels of jurisdiction. Also, international regulatory organizations create mandatory standards at the international level. As an example, safety standards are often issued in mandatory form.

Table 1
Criteria for Classifying Written Standards

Subject Specified

definition	how an important term is defined	(concept)
measurement	how a measurement must be made	(procedure)
performance	level of performance a product must attain	(goal)
quality	level of quality a product must attain	(goal)
compatibility	level of compatibility a product must attain	(goal)
safety	level of safety a product must attain	(goal)

Promulgating Entity

industry
government organizations
international organizations

Elective Nature

voluntary
mandatory

BUYER'S VIEW OF COMPETITIVENESS

Ultimately, it is the buyer who determines whether a product is competitive in the marketplace. The buyer evaluates the competitiveness of a product on several bases, as shown in Table 2. The buyer requires that a product achieve certain desirable *product characteristics,* reflected in its performance, its quality (including its reliability), and its compatibility with other products in the buyer's system. The buyer also requires attractive *marketplace-exchange* factors, including price, timely availability, and an acceptable basis for reaching agreement with the manufacturer on product characteristics. Reaching agreement requires: (1) establishing meaningful technical specifications for the product, and (2) providing a means of proving compliance with those specifications and with prevailing written standards. Finally, the buyer may require after-sales *support* and will want to know how much of that support will be provided by the manufacturer. In particular, the buyer is concerned with the installation, maintenance, and daily operation of the product.

The *marketplace exchange* and *support* factors have associated *costs* which, when evaluated along with the price of the product, contribute to the buyer's assessment of the life-cycle cost of owning the product. The same factors are also conducted with varying degrees of speed, which is also of concern to the buyer.

Every factor in Table 2 is strongly affected by the level of measurement capability, as indicated by the annotations on the left side of the table. Product characteristics, price, and timely availability are highly dependent on the *manufacturer's* measurement capability. The other factors are highly dependent on both the *manufacturer's* and the *buyer's* measurement capability. The following sections explain the nature of these measurement dependencies by examining what the manufacturer must do to meet the buyer's demands.

Table 2
Buyer's Demands for Competitiveness

Product Characteristics

measurements →	performance
measurements →	quality/reliability
measurements →	compatibility

Marketplace Exchange

measurements →	price
measurements →	timely availability
	agreement with manufacturer
measurements →	specifications
measurements →	proof of compliance
measurements →	*cost of marketplace exchange*
measurements →	*speed of marketplace exchange*

Support

measurements →	installation
measurements →	maintenance
measurements →	daily operation
measurements →	*cost of support*
measurements →	*speed of support*

MANUFACTURER'S CHALLENGES FOR COMPETITIVENESS

The manufacturer must cope with a broad spectrum of challenges to meet the buyer's demands. The manufacturer's challenges can be divided into three groups: (1) developing, manufacturing, marketing, and supporting a product; (2) raising and protecting resources; and (3) implementing new management strategies. All of three of these groups of challenges are discussed in the sections that follow.

Developing, Manufacturing, Marketing, and Supporting a Product

To put a product into service to the buyer, the manufacturer must develop the product, manufacture it, market it, and often support it after sale. These basic steps are shown in Table 3 along with some of the concerns that the manufacturer must address for each step. The level of measurement capability affects many factors, as suggested by the annotations on the left side of the table.

Table 3
Manufacturer's Challenges for Competitiveness, Part 1 of 3:
Develop, Manufacture, Market, and Support a Product

Research and Development

measurements → discovery
measurements → product design

measurements → *cost of R&D*
measurements → *speed of R&D*

Manufacturing

measurements → process design
measurements → process control

measurements → *cost of manufacturing*
measurements → *speed of manufacturing*

Marketplace Exchange

access
- export restrictions
- import restrictions
  - tariff barriers
  - non-tariff barriers
    - nationalism
    - local content requirements
    - *measurements →* written standards compliance
    - *many others*

agreement with buyer
measurements → specifications
measurements → proof of compliance

measurements → *cost of marketplace exchange*
measurements → *speed of marketplace exchange*

Support

measurements → installation
measurements → maintenance

measurements → *cost of support*
measurements → *speed of support*

Research and Development

In the area of research and development, advanced measurement capability enables the *discovery* of fundamental phenomena of nature that later become the basis for new products. Without advanced measurement capability, the new phenomena might be missed entirely or might not be understood.

Years of expensive research time may be wasted, and often are, in the absence of adequate measurement capability. Not surprisingly, new discoveries are frequently linked directly to the development of new measurement capability. Further, advanced measurement capability accelerates the process of discovery by enabling the scientific community to compare findings on a common basis and, therefore, to reach definitive conclusions quickly. Measurement capability is therefore an essential part of the technical *language* needed for collaboration.

> For example, research toward new high-temperature superconductors has been impeded by the lack of a widely accepted measurement method for determining the current-carrying capacity of the superconductors. In the absence of such a method, researchers cannot readily compare and verify each others findings.

After discovery, advanced measurement capability is needed to exploit fundamental phenomena, whether discovered by the manufacturer or others, in a *product design* that can be manufactured efficiently and that has the levels of performance, quality, and compatibility demanded by the buyers. Individual product characteristics cannot be optimized unless they can be measured.

> For example, the information-handling capacity of microwave communications system, like those in satellites, increases as unwanted electronic "noise" is reduced. Reduction of that noise requires accurate ways of measuring it. Present measurements methods are not adequate to support the design process required to reduce noise to desired levels for many emerging microwave systems.

By facilitating discovery and product design, advanced measurement capability also reduces the *cost*, and increases the *speed*, of carrying out these processes.

Through these effects, the level of measurement capability employed in research and development directly impacts several of the buyer’s demands in Table 2: desirable product characteristics, attractive price, and timely availability.

Manufacturing

Once a product has been developed, advanced measurement capability is needed to *design manufacturing processes* that can make the product. These processes must provide the manufactured product with the desired levels of performance, quality, and compatibility.

> For example, an accurate knowledge of the microwave properties of materials is needed to design automated manufacturing processes for making microwave integrated circuits. Critical tolerances that must be achieved by the processes are sensitively dependent on those properties. Advanced measurement capability is needed to determine the properties.

Once the manufacturing process is designed, tight *process control* must be established to assure the quality of the product. That control is highly dependent on the performance of the measurements that monitor the key quantities that determine quality.

> For example, accurate measurements of the thickness and electrical properties of the thin layers within integrated circuits are essential to controlling the quality of those circuits during manufacture. Accuracy good enough to distinguish individual atomic layers will be required for the next generation of integrated circuits.

Since advanced measurement capability facilitates both process design and process control, it also reduces the *cost* and increases the *speed* with which the needed processes can be developed.

For many high-technology products, just the *direct cost* of making measurements during manufacturing is very high, entirely aside from the broader impact of measurements on product design and manufacturing processes.

> For example, about twenty percent of the cost of producing optical fibers is in measurements. An estimated thirty percent of the cost of complex semiconductor integrated circuits is in testing.[1] Twenty-five percent of the cost of producing microwave integrated circuits based on compound semiconductors is attributed to measurements.

Through its effects on manufacturing, the level of measurement capability directly impacts several of the buyer's demands in Table 2: desirable product characteristics, attractive price, and timely availability.

Marketplace Exchange

The fate of virtually every high-technology product during marketplace exchange is affected by measurements. Both *gaining access to the marketplace* and *reaching agreement with the buyer* are affected.

Gaining access to the marketplace often requires proving compliance with written standards that specify levels of performance, quality, and compatibility for high-technology products. Both national and international written standards may apply. The level of measurement capability determines whether the manufacturer can *prove compliance* with the standards and gain market access. The level of measurement capability also greatly affects the *cost* and *speed* of proving compliance.

> For example, every digital electronic product sold in the U.S. must first comply with a mandatory compatibility standard of the Federal Communications Commission. That standard assures that the product's emissions of unwanted electromagnetic radiation are sufficiently low that they will not interfere with other sensitive electronic products. Measurements of emissions are used to prove compliance.

Reaching agreement with the buyer requires first *agreeing on the specifications* for a product. The level of measurement capability determines whether key characteristics of the product, such as performance, quality, and compatibility, can be specified in a form useful to the buyer. Reaching agreement with the buyer also requires *proving compliance* with the product's specifications. The level of measurement capability determines whether the manufacturer and the buyer can reach agreement that the product really conforms to the specifications. Both the manufacturer and the buyer may make measurements on the product. For high-technology products, millions of dollars in penalties and bonuses in purchasing contracts can be at stake in this determination. The level of measurement capability affects both the *cost* and *speed* of the process of reaching agreement.

> For example, power transformers used by the electrical utilities must have extraordinarily low electrical losses, since those losses can cost more over the lifetime of a transformer than its purchase price. Advanced measurement methods are necessary to detect these losses, and severe penalties may be exacted by the buyer for the manufacturer's failure to meet agreed loss specifications.

By facilitating *access to the marketplace* and *ability to reach agreement*, the level of measurement capability directly impacts key buyer's demands from Table 2: timely availability and agreement with manufacturer.

There are dozens of *other barriers* that can limit access to the marketplace. They are primarily non-technical in nature. They include both export restrictions and import restrictions, as illustrated in Table 3. The import restrictions include both tariff and non-tariff barriers. The nature and importance of these other barriers has been addressed well elsewhere, most recently in a major study by the International Trade Administration.[2] Only representative barriers are shown in the table. Hurdling these other barriers is *just as necessary* as hurdling the barriers affected by measurement capability.

Support

After the sale of a product, measurements are often required to support the product. In particular, measurements may be required for the *installation*, *maintenance*, and *operation* of the product to assure proper performance. Depending on the support agreement between the manufacturer and the buyer, the manufacturer may assume some of the burden of installation and maintenance. The buyer must bear the rest. Both usually need a high level of measurement capability to meet the requirements for support. The installation of high-technology products often requires the integration of a number of purchased products into properly functioning systems of considerable complexity. Such integration can be highly measurement-intensive. For these several reasons, the *cost* and *speed* with which support can be provided is highly dependent on the level of measurement capability used in providing that support. Thus the level of measurement capability directly impacts the buyer's demands for the several types of support shown in Table 2.

> For example, electronic equipment in airplanes requires periodic testing to assure its proper functioning. The level of measurement capability used for that testing determines the validity of the tests, the speed of making them, and the cost of making them. Shortcomings can increase downtime and reduce safety.

Raising and Protecting Resources

The manufacturer faces a second set of challenges that affect competitiveness. They are outlined in Table 4 and relate to raising and protecting resources. The manufacturer must *raise resources* to enable him to conduct business. In particular, he must obtain capital, labor, matériel (equipment and supplies), and technology. The manufacturer must also *protect the resources* that he creates during product development to prevent appropriation by his competitors.

In coping with these challenges, the manufacturer's success is influenced by a number of concerns, both general and specific, as illustrated in Table 4. The general concerns have much in common from item to item in the table. The specific concerns reflect barriers or opportunities relevant to meeting the challenges. Only representative specific concerns are shown in Table 4. There are many others, and they are also important to competitiveness. The broader spectrum of concerns are discussed well in other publications, most recently in a major study by the International Trade Administration.[3]

Measurement capability enters here as a *technology* resource that the manufacturer must acquire in order to develop, market, and support his product. Measurement capability also enters here through

Table 4
Manufacturer's Challenges for Competitiveness, Part 2 of 3: Raise and Protect Resources

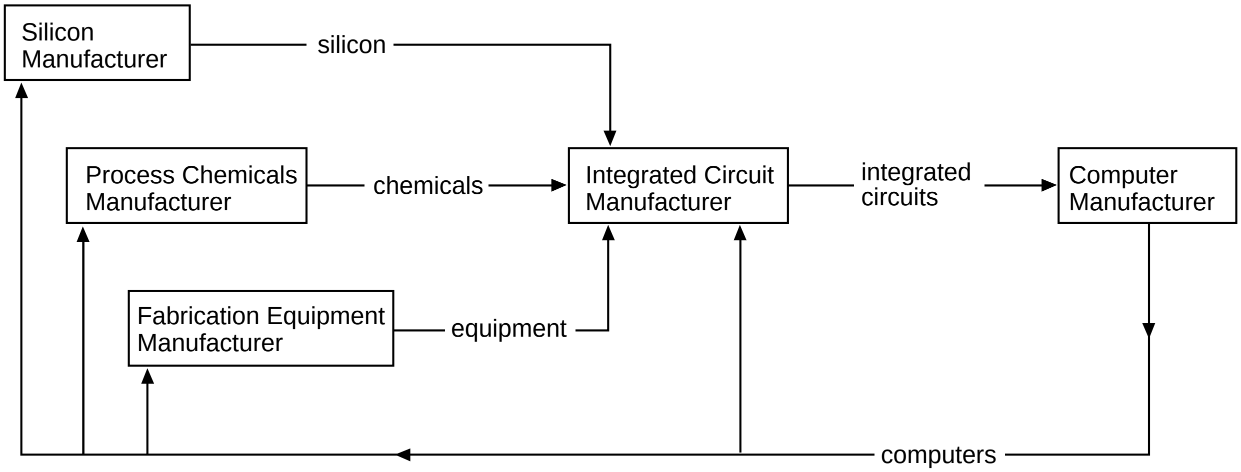
	Challenges	General Concerns	Specific Concerns (representative)
	Raise Resources		
	Capital	cost	stock market
	equity	availability	interest rates
	debt	"patience" or term	Federal debt
			Federal Reserve policy
			tax incentives
			R&D tax credit
			capital gains taxes
			direct Federal funds
	Labor	cost	long-term view
	managers	availability	teamwork view
	workers	effectiveness	education
			experience
	Matériel	cost	
measurements →	equipment	availability	
measurements →	supplies	quality	
	Technology	cost	patent laws
	supporting	availability	copyright laws
measurements →	**measurements**	quality	
	generic		
	other		
	Protect Resources		
	Intellectual Property Protection	cost	patent laws
		availability	copyright laws
		effectiveness	trade secrets

the matériel that the manufacturer buys to support his processes. The cost, availability, and quality of that matériel are highly dependent on the measurement capability of the supplier.

> For example, the starting materials purchased from a supplier by the manufacturers of semiconductor integrated circuits must have very low levels of impurities. Impurity requirements below 10 parts per billion are increasingly common. Impurity requirements below 10 parts per trillion are foreseeable.[4] The supplier must have extraordinary measurement capability to produce such low impurity levels. Both the supplier and the integrated-circuit manufacturer must have extraordinary measurement capability to confirm compliance with those levels at the time of sale.

The impact of a high level of measurement capability can be intensified when manufacturers find themselves in a buyer-seller loop. Such loops are increasingly common in high-technology fields because of the growing mutual dependency of the fields. Figure 1 provides an example of a buyer-seller loop involving the semiconductor and computer fields. The integrated-circuit manufacturer, in the center of the figure, buys silicon, process chemicals, and fabrication equipment from other manufacturers, makes integrated circuits, and sells the integrated circuits to many buyers, including computer manufacturers. The computer manufacturers sell computers to many buyers, including their own supplying manufacturers. The supplying manufacturers use the computers for many functions, including improving their process control and thus the quality of their products. Through this chain of events, the benefits of improved measurement capability -- introduced by any one of these manufacturers -- are passed on to the others in the buyer-seller loop, advancing the quality of the products of them all. This propagation of benefits is a major reason for the high leverage of measurement capability on competitiveness.

Figure 1
Propagation of Measurement Benefits in a Buyer-Seller Loop



Implementing New Management Strategies

Manufacturers face a third category of challenges, related to managing their resources. They struggle to implement new management strategies that will improve their competitiveness. These strategies must successfully respond to special dimensions of the challenges already discussed. Those special dimensions include: shorter marketplace lifetimes for products, demands for greater product diversity, and increased costs for the development of high-technology products. Representative responsive management strategies are shown in Table 5, along with the principal motivations for each. Each strategy has other motivations, too; for example, total quality management can reduce costs (by reducing waste, for example) as well as improve quality.

Common among the strategies is their strong dependency on the level of measurement capability, as suggested by the annotations on the left side of Table 5. In the simplest of terms, this dependency arises from the critical role of measurement capability in improving: (1) the *control* that manufacturers exercise over all of the technical steps required to realize competitive products; and (2) the effectiveness of the technical *language* that manufacturers use in executing the strategies, particularly those requiring collaboration or technology transfer. The examples in the following sections will be clarifying.

Table 5
Manufacturer’s Challenges for Competitiveness, Part 3 of 3:
Implement New Management Strategies

	Strategy	**Principal Motivation**
measurements →	total quality management	product quality
measurements →	flexible manufacturing	timely availability through increased product diversity
	collaboration among organizations	cost reduction through resource sharing
measurements →	cooperative development	
measurements →	technology transfer	
	collaboration within organizations	timely availability through reduced product development time
measurements →	teamwork	
measurements →	concurrent engineering	

Total Quality Management

Total quality management programs take a comprehensive approach to the improvement of product quality. These programs require control of every step necessary to develop, manufacture, market, and support a product. During manufacturing, both process variables and product characteristics must be controlled. Since control cannot be established over quantities that cannot be measured, the success of these programs is highly sensitive to the level of measurement capability. Further, total quality management programs generally cannot succeed without incorporating the suppliers of a manufacturer into the effort. The level of measurement capability determines how effectively the manufacturer and his suppliers can *communicate* the characteristics of the supplies and the manufacturing equipment that will affect the final quality of the manufacturer’s product.

Flexible Manufacturing

Manufacturers have traditionally competed on the basis of *economies of scale*, that is, on the ability to produce large quantities of a relatively small number of products with high efficiency. Increasingly, to maintain their competitiveness, manufacturers must compete on the basis of *economics of scope*, that is, on the ability to provide diverse products, often in relatively small quantities and often with relatively short marketplace lifetimes, while retaining very high manufacturing efficiency. Flexible manufacturing is an effort to respond by providing manufacturing approaches that are readily adaptable across a relatively broad range of product types and variations. The success of flexible manufacturing is particularly dependent on the *control* provided by a high level of measurement capability.

> For example, flexible manufacturing requires regaining control of manufacturing processes quickly and economically after they have been altered to accommodate a new product. The ability to regain that control is highly dependent on the level of measurement capability used in the processes.

Collaboration Among Organizations

Collaboration among organizations is a major way of reducing costs and gaining access to improved technology by sharing resources. Manufacturers may collaborate with each other, with government, or with educational institutions. They may conduct joint development of technology or engage in transferring technology among collaborators. Their success in these efforts is highly dependent on the sophistication of the technical *language* used for the collaboration or the transfer; and a key part of that language is the measurement capability employed.

> For example, transfer of sophisticated integrated-circuit manufacturing processes from one manufacturer to another, or even from one plant to another within a given company, cannot be accomplished without superb measurement capability, exercised mutually by the collaborating parties. Similarly, in areas of research, many findings cannot be shared between participants -- whether they are universities, government laboratories, or industry -- without a high level of measurement capability, exercised mutually by the participants, to enable findings to be consistently interpreted and reproduced.

In addition to the technical challenges of collaboration, manufacturers face major non-technical challenges, including anti-trust laws, conflict-of-interest laws, and general distrust.

Collaboration Within Organizations

Collaboration within organizations is a major way of improving the timely availability of new products by reducing product development time. The need to do this is driven by the fast pace of technological change and opportunity, and by the intensifying nature of competition in the marketplace. The lifetimes of new products *in the marketplace* are becoming shorter, whatever is happening to the lifetimes of products *in service*. Thus products late to market face not only the challenge of established competition but also the prospect of missing entirely the window of marketplace interest.

One approach to improving timely availability is to improve teamwork among those working on research and development, manufacturing, marketing, and support. Buyers may be incorporated in the team. The goal is to design a product that can be readily manufactured, marketed, and supported, and that is known at the outset to meet the specified needs of the buyers. Such collaboration requires an effective technical *language* to promote good communication; the measurement capability shared by the participants is a major part of that language, since it is used to specify, in exact terms, the characteristics of the product.

Collaborative activity is carried to a particularly intense level in *concurrent engineering* strategies, which attempt to collapse the product development cycle even further. In concurrent engineering strategies, teamwork is again employed; but, in addition, key steps of the product-development cycle, such as product design and manufacturing-process design, are conducted concurrently rather than sequentially. This approach enables, for example, scrutinizing alternative product designs for manufacturability very early in the design process. Concurrent engineering strategies carry high risk because they are very unforgiving of mistakes that might require redoing costly downstream steps.

Successful outcomes are highly dependent on the quality of the technical *language* used by the collaborators; and, once again, measurement capability is a major part of that language.

> Here is a hypothetical example of the role of measurement capability in new concurrent engineering strategies. An engineer designing a connector to join optical fibers will determine early what accuracy is required in the fit of the connector to the fibers in order to minimize the loss of light at the connection. He may make a number of critical measurements for confirmation. The design engineer will take this accuracy requirement to the process engineer who must develop the manufacturing process to make the connector. If the design engineer and the process engineer share the same high level of measurement capability, then the process engineer will be able determine if the accuracy required is attainable in a practical manufacturing process, even before the design is completed. Thus, a high level of measurement capability shared by the design engineer and the process engineer forms a key part of the technical language that they use when working concurrently to assure that time and money are not wasted in completing a design that would later be judged "not manufacturable".

CONCLUSION

A high level of measurement capability is essential to every step that manufacturers of high-technology products must complete to realize competitive products, including research and development, manufacturing, marketplace exchange, and after-sales support. Similarly, a high level of measurement capability is essential to the implementation of new management strategies designed to improve competitiveness, such as total quality management, flexible manufacturing, and concurrent engineering. For these reasons, progress in measurement capability is an essential goal of any effective strategy for improving the competitiveness of high-technology industries.

ENDNOTES

1. *Semiconductor Technology: Workshop Conclusions*, Semiconductor Industry Association, p. 19 (November 17-19, 1992). Contains views generated by 179 of the country's key semiconductor technologists, drawn principally from U.S. industry and its suppliers and customers, but including some individuals from academia, government agencies, and the National Laboratories.

2. *The Competitive Status of the U.S. Electronics Sector from Materials to Systems*, International Trade Administration, U.S. Department of Commerce, especially pp. 20-41 (April 1990).

3. *The Competitive Status of the U.S. Electronics Sector from Materials to Systems*, International Trade Administration, U.S. Department of Commerce, especially pp. 20-41 (April 1990).

4. A trillion is a million million or 10^{12}.

CHAPTER 2

NIST'S ROLE IN MEASUREMENTS

August 1991

Chapter 2

NIST'S ROLE IN MEASUREMENTS

SUMMARY

The National Institute of Standards and Technology (NIST) provides measurement capability to U.S. industry, government agencies, and educational institutions in support of key national goals: strengthening the U.S. economy, improving Government operations, and improving health and safety. A key focus of the effort to strengthen the U.S. economy is *improving U.S. international competitiveness*.

The measurement capability that NIST provides supports research and development, manufacturing, marketplace exchange, and after-sales support. The measurement capability is used by industry, governmental organizations, and educational institutions. This capability takes four principal forms: *measurement methods* that are used to determine the values of critical quantities, *measurement reference standards* that assure the accuracy of the methods, *measured reference materials data* of verified accuracy, and *techniques for analyzing measured data*.

NIST is the only Government agency with this measurement mission. Further, NIST is the only U.S. organization recognized by domestic and international organizations as the ultimate U.S. reference point for measurements. Through its support of research, development, and manufacturing, NIST's measurement capability underpins industry's pursuit of improved performance, quality, and compatibility for more competitive products. Through its support of marketplace exchange, NIST's measurement capability helps U.S. industry gain entry to international markets by providing the methods used to demonstrate compliance with international standards.

NIST does not provide measurement capability that other U.S. organizations can provide for themselves. Rather, NIST focuses on measurement capability that is beyond the reach of individual organizations for economic, technical, or other reasons, that is widely needed in industry, and that will have *high economic impact* if provided. Because measurement capability is a pacing factor in industry's pursuit of competitiveness, NIST endeavors to provide new measurement capability at the earliest time that it is needed by industry.

NIST employs a broad range of means for delivering new measurement capability to industry and government, including communications, joint activities, and paid services. The beneficiaries of NIST measurement capability function as manufacturers, buyers, users of services, sponsors or performers of research, and regulators.

NIST'S MEASUREMENT CAPABILITY

NIST provides four principal forms of measurement-related capability:

measurement methods

measurement reference standards
measured reference materials data
techniques for analyzing measured data

Measurement Methods

NIST develops new measurement methods for quantities not currently measurable. NIST develops new measurement methods to provide improved accuracy for quantities already measurable. NIST validates the accuracy of measurement methods already in service in industry so that they can be used with confidence throughout the industry. NIST resolves differences among competing measurements methods in use in industry, so that differing methods can be used interchangeably or can be selected on a sound basis.

Measurement Reference Standards

Virtually all measurements are based on comparing an unknown quantity to a known quantity or *reference standard*. For example, a very simple reference standard is a ruler that is held up to an object to determine its size, or the weights on the other side of a balance scale.

In electronics, measurement instruments contain internal reference standards. Electronic circuits in the instruments make comparisons between those standards and the unknown values of the quantities to be measured. The accuracy of the internal reference standards can be determined by comparison with even better reference standards.

NIST develops and maintains *national* measurement reference standards. They serve as the ultimate references for these comparisons and thus underpin the national measurement system. Through an expanding chain of comparisons, the accuracy of the national measurement reference standard for each quantity is related to the accuracy of large numbers of critical instruments in everyday service in industry, government, and educational institutions. Once so related, the accuracy of these instruments are said to be *traceable* (or related in a known manner) to the national measurement reference standards.

> For example, the national measurement reference standard for voltage is a sophisticated superconducting integrated circuit that produces highly accurate voltage levels. These accurate voltage levels are passed on to other organizations by means of *intermediate* measurement reference standards that are usually less accurate but more transportable and less expensive than the national standard. The voltage levels of the intermediate standards are determined to a known accuracy level by comparison (*calibration*) with the national standard. The intermediate standards are taken to user organizations where they are used to calibrate an even larger number of instruments that make voltage measurements on a daily basis. This process of transferring accuracy through successive calibrations makes the accuracy of instruments traceable to the national standard.

A similar process is used to establish the traceability of many other quantities to NIST's national measurement reference standards. In this manner, tens of millions of measurement devices are made traceable to NIST's national measurement reference standards every year.

Often, more than one intermediate standard is required. There is some unavoidable loss of accuracy with each successive calibration in a chain of intermediate standards. For this reason, the accuracy of NIST's national reference standards must be very high to assure that the accuracy at the end of the chain of calibrations is sufficient to meet users' needs. In some cases industry requires accuracy equal to, or very close to, the best that NIST can provide. In these cases, the calibration chain must be bypassed by a more direct special service.

Measured Reference Materials Data

NIST develops some measured data using its advanced measurement capability. NIST also validates data developed by industry or other organizations when such validation is necessary to achieve wider acceptance and use of the data in industry. Nearly all of these data focus on the properties of materials. Industry needs these data for reference purposes, particularly when designing products and developing manufacturing processes that are sensitive to the properties of materials. Most electronic products are highly sensitive to the properties of the materials from which they are made.

Techniques for Analyzing Measured Data

In selected instances, NIST will develop techniques for analyzing industry's measured data, especially when the techniques appear to have wide applicability within industry.

> For example, NIST is developing techniques called test strategies to identify the specific minimum set of measurements required to determine the performance of important classes of electronic components. These test strategies will be used by manufacturers to reduce the high cost of testing during manufacturing and to improve quality control.

WHY INDUSTRY CALLS ON NIST

NIST does not develop measurement capability that industry can provide for itself. Rather, NIST develops selected measurement capability that is beyond the reach of individual companies for economic, technical, or other reasons, that is needed widely in industry, and that will have high economic impact if provided. It is the goal of NIST to improve the measurement capabilities of U.S. industry as a whole. Understandably, this is not the goal of individual companies.

There are four key objectives that must be met if new measurement capability is to have the maximum positive impact on the economy:

(1) Identify the measurement problem early.
(2) Solve the identified measurement problem quickly.
(3) Achieve wide adoption of the measurement solution.
(4) Accomplish the above at minimum cost to the economy.

NIST has several characteristics shown below which, when taken together, enable NIST to accomplish the four objectives:

> NIST has a *long-range view* of the measurement needs of the nation.
>
> NIST has a *broad view* of the measurement needs of the nation across all sectors of the economy.

> NIST's *impartiality* is widely acknowledged. NIST is not a buyer or a seller, and NIST has no biases that favor buyers or sellers; rather NIST serves both. NIST is not a regulatory agency and thus poses no threat to industry. NIST protects proprietary information.
>
> NIST's *competence* in measurement capability is widely acknowledged by U.S. industry, by domestic and international buyers, and by international regulatory organizations whose rulings affect the access of U.S. products to international markets.
>
> NIST bears the *official imprimatur* of the U.S. Government as the lead agency for measurements.

Here is how these characteristics enable NIST to work with industry to meet the four objectives, and why industry generally cannot meet those objectives without NIST.

Identify the Measurement Problem Early

NIST's *long-range view* enables it to foresee the emergence of many critical measurement needs well in advance and to prepare to meet them. NIST's *broad view* helps it discover those measurement problems with the greatest economic impact across an entire industry. NIST's reputation for *impartiality* and *measurement competence* enables NIST to gain the trust of companies so that NIST can gain *access* to their information on measurement problems. This information is essential in defining correctly a new measurement problem that is afflicting companies widely. This document, *Measurements for Competitiveness in Electronics*, is part of NIST's effort to identify measurement problems early.

In comparison, individual companies, are not generally focused on measurement problems that cut across an industry. Also in the absence of NIST, they would be unlikely to share internal data with each other so that the pattern associated with a given measurement problem afflicting the industry as a whole could be recognized. NIST is trusted to protect proprietary data that may be revealed during the process of working closely with individual companies.

Solve the Identified Measurement Problem Quickly

NIST's *measurement competence* is instrumental in solving the identified measurement problem quickly. NIST's *impartiality* enables NIST to gain industry's cooperation in conducting tests in industrial settings to assure that the measurement problem really has been solved.

Once again, individual companies cannot gain such access to other companies. Also, to really *solve* the measurement problem, in the broad sense, NIST's solution must be applicable over industry-wide product types and performance levels. The development of such widely applicable measurement capability is beyond the motivations of individual companies which must focus more exclusively on the needs of their own product lines.

Achieve Wide Adoption of the Measurement Solution

Once NIST has developed and tested new measurement capability, the full economic benefits can be realized only if that capability is adopted widely by industry and is recognized by domestic and international buyers and, sometimes, by international regulators. Three of the characteristics of NIST mentioned above -- *impartiality*, *measurement competence*, and *official U.S. imprimatur* -- are the essential elements that enable NIST to achieve such acceptance.

Achieving wide adoption of new measurement capability is generally beyond the abilities or the motivations of individual companies, even when such wide adoption is in their interest for promoting buyer acceptance of their products. Understandably, an individual company's principal focus is on it own internal use of measurement capability.

Wide adoption is vitally important to buyers as well as to manufacturers. Without wide adoption by both manufacturers and buyers, marketplace transactions for high-technology products, which are generally based on proving compliance with product specifications, would not proceed efficiently. Also, without adoption by buyers, the buyers would lack the capability required to support their new purchases. Buyers often have less access to measurement capability than manufacturers.

Accomplish the Above at Minimum Cost to the Economy

NIST achieves high cost efficiency in the development of measurement capability. NIST expends resources only once to develop a measurement method that is then used by everyone. Far fewer resources are required than the sum total of the resources that individual organizations would expend to duplicate such work, even if they could. Thus, very high economic leverage is realized.

TIMING VERSUS IMPACT OF NIST'S MEASUREMENT CAPABILITY

New measurement capability has the highest economic impact when introduced as early in the product development cycle as possible. This is true because, as noted in Chapter 1, measurement capability affects virtually every step required to put a high-technology product into service to the buyer: research and development, manufacturing, marketplace exchange, and support. Measurement capability also affects the success of new management strategies for improving competitiveness, such as concurrent engineering and total quality management.

Measurement capability is a pacing factor in competitiveness in high-technology fields. Manufacturers cannot reliably produce products whose performance, quality, and compatibility exceeds the level of their measurement capability. Measurement capability can be appropriated by other nations in time. Consequently, maintaining international competitiveness in high-technology fields requires a relentless pursuit of progress in high-impact areas of measurement capability.

In response to this need, NIST strives to provide critical new measurement capability at the earliest possible time consistent with its resources.

HOW NIST DELIVERS NEW MEASUREMENT CAPABILITY

A key step in accomplishing adoption of new measurement capability is *delivery*. NIST focuses a very large portion of its energy on the delivery of the measurement capability that it develops. Delivery is accomplished in several ways: communications, joint activities, and paid services.

Communications

NIST communicates its findings to wide audiences through published descriptions of virtually all of its measurement capabilities, through talks at conferences, through telephone consultations with industry, government agencies, and universities, and through regional technology centers set up by NIST. Prior to publication of new findings, NIST has generally worked so closely with U.S. companies during the development and testing phase that they benefit from NIST's efforts very early in the process.

Joint Activities

NIST engages in a variety of joint activities with industry. NIST staff participate in industrial associations and professional societies. These organizations incorporate NIST measurement capability into voluntary industry standards. This is a very important method of promoting the adoption of improved measurement capability. Federal and state agencies promote adoption by requiring that the products that they purchase, or that they regulate, be evaluated or calibrated with the measurement capability that NIST has developed or that is traceable to NIST. For example, both the Department of Defense and the various state public utility commissions require traceability to NIST's measurements.

NIST conducts round-robin measurement comparisons with industrial and government participants. In a round robin, NIST circulates a physical object, such as a special optical fiber, to the organizations that wish to participate. They measure carefully certain important properties of that object, and they report their findings in detail to NIST. By studying the results of the measurements, NIST can determine if a new measurement method is needed. If so, NIST develops that method and trains industry and government representatives in its use. Then NIST conducts a second round robin to assure that all of the participants can measure the important properties accurately.

NIST also conducts cooperative research projects with industry and with government agencies. Employees from participating organizations come to NIST to work with NIST staff for six months to two years. This is one of the most effective methods for transferring measurement technology.

Paid Services

NIST offers selected measurement services to individual companies and to government organizations on a reimbursable basis. This type of support takes several forms:

> Other agencies transfer funding to NIST to support development of measurement capability needed for their specific systems.
>
> NIST provides copies of NIST-developed measurement reference standards needed by industry and government agencies for on-site verification of their measurement capability.
>
> NIST calibrates measurement reference standards that industry and government agencies send to NIST for that purpose.

NIST provides courses and on-site consultations that train industry and government staff in the use of advanced measurement capability.

BENEFICIARIES

These means of delivery touch a very wide range of beneficiaries including companies, government agencies, educational institutions, and individuals. The beneficiaries function in many different capacities -- as manufacturers, as buyers, as users of services, as sponsors or performers of research, or as regulators.

WHAT NIST DOES NOT DO

NIST does not generally develop products. However, NIST does develop measurement instruments needed for its internal work. These measurement instruments are sometimes adapted by industry for commercial applications and sold as products. Also, industry regularly implements NIST measurement capability in products and manufacturing processes of its own development. This is an important means of delivering NIST measurement capability to users widely.

With rare exceptions, NIST does not evaluate products. However, the measurement capability that NIST develops is very important for product evaluation. As noted in Chapter 1, such evaluations are critical to manufacturers and buyers in establishing that the performance of products meets the agreed specifications so that market transactions can be completed. Such evaluations are also critical to manufacturers in proving compliance with domestic and international marketplace standards, so that their products can enter those markets.

NIST is not a regulatory agency. With one exception, NIST does not develop or issue written standards or regulations that are binding on industry or other government agencies. [Written standards are defined and discussed in Chapter 1, Role of Measurements in Competitiveness, beginning on page 3.] However, NIST's measurement capability is regularly incorporated in voluntary standards developed by industry. In fact, this is a major means by which NIST's capability is delivered to industry *by industry itself*. The one exception to the development of written standards occurs in the computer area. NIST does develop written standards that affect computer network interfaces and data security practices within the U.S. Government only. The development is conducted in close cooperation with industry through voluntary standards organizations to assure that the standards are implementable by industry and are compatible with industry standards widely. The written standards become binding on specific Government agencies when approved by the Secretary of Commerce. However, neither NIST nor the Department of Commerce is responsible for the enforcement of the standards. Enforcement is handled by the participating agencies and by Government oversight organizations.

OTHER ROLES OF NIST

NIST conducts *fundamental research* of broad importance to the nation. This research extends beyond the fundamental research that NIST conducts as an integral part of its pursuit of major advances in measurement capability. The broader research enhances NIST's technical competence and often provides unforeseen benefits to the measurement mission.

NIST develops *generic technology* as well as measurement technology. Generic technology is not product specific but rather is broadly supportive of industry's efforts to realize new products. The development of generic technology is undertaken to demonstrate that a technical concept has potential for practical applications and to reduce the technical risk that might otherwise inhibit industry from pursuing product development. NIST also provides funding to industry groups for generic technology development. The funding is provided through a highly competitive process that requires cost sharing by industry.

NIST fosters the transfer of technology from U.S. Government laboratories and other sources to industry. This effort includes measurement technology, but also extends beyond it. The key current emphasis is the transfer of manufacturing technology.

CHAPTER 3

OVERVIEW OF U.S. ELECTRONICS AND ELECTRICAL-EQUIPMENT INDUSTRIES

March 1992

Chapter 3

OVERVIEW OF U.S. ELECTRONICS AND ELECTRICAL-EQUIPMENT INDUSTRIES

SUMMARY

The U.S. electronics industry and the U.S. electrical-equipment industry are both highly important to the U.S. economy. The electronics industry ships $267 billion of products annually (1990). Among U.S. manufacturing industries, it is second in shipments only to the chemical industry and employs more than twice as many people, 1.9 million (1990). The electrical-equipment industry ships $48 billion of products annually (1990). Included among the shipments of that industry are the products that support the electric utilities. These utilities provide $176 billion of electricity annually (1990) through the national power network.

Electronic products provide information and energy services at the most fundamental level. The largest segments of U.S. shipments of electronic products support information services and take the form of communications equipment (including search and detection equipment) and computers. Together, these two types of equipment account for more than half of all U.S. shipments of electronic products.

Electrical equipment provides energy services at the most fundamental level. The largest segments of U.S. shipments of electrical equipment are relatively close in size and include wire ($7.4 billion), motors ($6.9 billion), switchgear ($5.1 billion), batteries ($4.6 billion), and transformers ($4.1 billion), all for 1990. More than 80 percent of all electricity generated is converted to non-electrical forms before it is used. Thus electricity is used primarily as a means of transferring energy. The dominant non-electrical form to which electricity is converted is motion; the motors that produce that motion consume about 60 percent of all electricity generated in the U.S. (1985).

The international markets for electronic and electrical products are highly competitive. Improving U.S. competitiveness will require outstanding performance in the development of products based on a broad spectrum of electronic materials. For the electronics industry, outstanding performance in the development of products based on semiconductor materials is particularly important. Semiconductors materials are critical to the performance of the largest dollar fraction of U.S. shipments of electronic products. Similarly, for the electrical-equipment industry, outstanding performance in the development of products based on magnetic materials is particularly critical. Finally, improved competitiveness will require developing products that can operate at frequencies from 0 hertz (direct current) to above 10^{18} hertz to meet the demand for an increasing diversity of services.

IMPORTANCE OF THE INDUSTRIES

The electronics industry is critically important to the U.S. economy. Among U.S. manufacturing industries, the electronics industry employs the largest number of people, 1.94 million (1990), as shown in Table 6.[1] In shipments, the electronics industry is a close second at $267 billion (1990) to the chemical industry. U.S. shipments of electronic products are dominated by non-defense applications, including commercial, non-defense government, and consumer applications which together account for $208 billion, or 78 percent (1990).[2] Electronic products have extraordinary leverage; they provide intelligence and control functions in virtually all modern products.

Table 6
Five Largest U.S. Manufacturing Industries (1990)

Manufacturing Industry	Shipments ($billions)	Employment (thousands)
Chemical[1(a)]	289	858
Electronics[1(b)]	267	1937
Automotive[1(c)]	234	836
Petroleum Refining[1(d)]	159	72
Aerospace[1(e)]	124	814

The U.S. electrical-equipment industry is smaller than the industries in Table 6 but has enormous economic impact. That industry is defined here to include the products that supply electricity and the products that convert electricity to non-electrical forms. U.S. shipments for this industry are $48 billion (1990). The equipment used to supply electricity supports the electric utilities, among other industries. The electric utilities provide $176 billion of electricity annually (1990).[3] The electrical equipment used to convert electricity to non-electrical forms supports many products, such as home appliances, automobiles, aircraft, and a large part of all modern manufacturing equipment.

STRATEGIES FOR DESCRIBING ELECTRONIC AND ELECTRICAL PRODUCTS

The products of the electronic and electrical industries can be classified according to several different strategies. Alternative strategies are listed in Table 7. These strategies provide a useful framework for understanding the industries and for aggregating economic data for the industries. Each of these strategies will be discussed briefly below.

Table 7
Strategies for Describing Electronic and Electrical Products

1. basic functions
2. societal services
3. stages of assembly
4. consuming economic sectors
5. technologies

Basic Functions and Societal Services

The first two strategies, basic functions and societal services, are usefully addressed together. In the most general sense, electronic and electrical products perform basic functions that fall into the following two classes: *information* and *energy*. The basic functions are outlined in Table 8, along with some examples of each. These basic functions are the elemental technical functions (such as storing information) that are the building blocks of the societal services. The societal services are described at the more general level that people normally think of as services (such as computing). Examples of societal services are shown in Table 9. They are representative of societal services relevant to electronic and electrical products. Societal services may be described in many different ways. There is no unique, non-overlapping set of such services.

Table 8
Basic Functions of Electronic and Electrical Products

Classes of Basic Functions	Basic Function	Examples
Information		
	generate	microwave radar systems, optical-fiber sensors
	manipulate	semiconductor microprocessors in computers and other products
	transfer	microwave and optical-fiber communications networks, cable television, wired conventional telephone networks
	store	magnetic rigid-disk drives, semiconducting integrated-circuit memory
	display	cathode-ray picture tubes, liquid-crystal displays, printers
Energy		
	generate	electric generators, solar cells, lasers for cutting, microwaves for heating
	manipulate	semiconductor power transistors, switches, and relays
	transfer	wires, transformers, insulators
	store	batteries, superconducting magnetic coils
	convert	motors, light bulbs, heating elements, electrodes for electrolytic processes

The relationship between basic functions and societal services is readily illustrated. For the *information* class, electronic products provide the following basic functions: *generate, manipulate, transfer, store, and display information*. These basic functions can be used to provide a variety of services at the societal level. *Computing* is an example; it requires performing all five of the basic functions. *Communications* is another example. Depending on the nature of the communications under discussion, it, too, may require performing all five of the basic functions. Electrical equipment, as defined here, does not provide information functions.

For the *energy* class, both electronic and electrical products provide the following basic functions: *generate, manipulate, transfer, and store electricity*. They also *convert* electricity to non-electrical forms (such as motion or light). Most electricity is converted to non-electrical forms to serve its intended application. Only a small part of all electricity generated is used in electrical form (as electronic equipment uses it). Thus electricity is put into an electrical form primarily to facilitate its transfer. An example of an energy service at the societal level is the supplying of *electric power* by the national electrical power system. The electric utilities that provide this service carry out the first three basic functions under energy: *generate*, *manipulate*, and *transfer*. They do not yet have a satisfactory means of storing large amounts of electrical energy. Conversion to non-electrical forms is done by the products that use the electricity.

Societal services often draw on basic functions from both classes: *information* and *energy*. For example, manufacturing control is really a combination of information and energy functions in the following sense. *Information* is generated by a host of sensors in a manufacturing process to determine temperature, position, size, and many other types of information. The information is then manipulated in computers to make decisions about process steps, and *energy* is transmitted to devices that implement those decisions.

Table 9
Societal Services
(examples)

computing
communications
entertainment
transportation
defense
medicine
manufacturing
electrical power

Stages of Assembly

The third strategy -- stages of assembly -- divides products according to the stage of the manufacturing process in which they fall, as shown in Table 10. That is, materials are used to make components; components are assembled into equipment; and pieces of equipment are assembled into systems. This is a useful structure when products are bought and sold at every stage, enabling marketplace data to be captured. Generally, such sales *do* occur at every stage. However, this strategy becomes less useful as companies become more vertically integrated and consume the products of an earlier stage without presenting them at a marketplace interface. When this occurs, it is difficult to capture consistently the value of intermediate products that companies produce for internal consumption.

Table 10
Stages of Assembly

Stage	Example
materials	semiconductor silicon
components	integrated circuit
equipment	computer
systems	computer network

Consuming Economic Sector

The fourth strategy -- consuming economic sectors -- divides products by the sector of the economy in which they are consumed. Commonly used subdivisions for this strategy are shown in Table 11. [In some taxonomies, *industrial* is a subset of *commercial*.]

Table 11
Consuming Economic Sectors

residential (or consumer)
commercial
industrial
government (and other)

Technologies

Finally, electronic and electrical products can be described by the technologies that they employ. There a variety of useful subcategories that can be created. The choice depends on the purpose intended. For the purposes of this document, the two most useful subdivisions are these: (1) the electronic *materials* from which the products are made, and (2) the *frequency ranges* in which the products operate. These are useful subdivisions because the measurement needs of electronic and electrical products differ greatly when different materials or different frequencies of operation are involved. For example, measurements used to characterize solid semiconducting materials bear little resemblance to those used to characterize liquid insulating materials. Similarly, measurements used to determine power at the frequencies used in electrical power systems, 50-60 hertz, bear little resemblance to the measurements used to determine power at the frequencies used in optical-fiber communications systems, 10^{14}-10^{15} hertz. Put another way, electronic and electrical products which have both materials and frequency ranges in common tend to have many measurement needs in common, too.

Multiple Strategies

While the above strategies for describing the products of the electronic and electrical-equipment industries have been presented individually, they are often used in combination. In some cases, one strategy provides the primary means of description, while another strategy provides subdivisions under the first. In other cases, the mixing of the strategies is more complex.

U.S. ELECTRICAL-EQUIPMENT INDUSTRY

The U.S. electrical-equipment industry can be conveniently described using the first strategy -- basic functions -- as depicted in Table 8. Further, since all of the products of this industry address energy, rather than information, they can be classified according to subdivisions in the energy section of Table 8.

Definition

The industry must be defined carefully because there are many products that are powered by electricity, but not all of these products can reasonably be defined as electrical equipment. The major categories of products included in the definition, and excluded from it, are shown in Table 12. The first major category of included products are those used to supply electricity. These products generate, transfer, manipulate, and store electricity. Manipulation, as used here, means controlling electricity and converting it into *other electrical forms* (alternating current to direct current, for example). The second major category of included products bridges those that supply electricity and those that apply it to practical purposes. The products in this category convert electrical energy into the four principal *non-electrical forms of energy* used in practical electrical applications: motion (motors), light (lamps), heat (resistive devices), and electrolytic action (principally carbon and graphite electrodes).

Table 12
Definition of U.S. Electrical-Equipment Industry:
Major Categories

within definition	electrical supply equipment	
		generation
		transfer
		manipulation
		storage
	electrical conversion equipment	
		motion
		light
		electrolytic action
		heat
outside of definition	electrical applications equipment	
	electronic equipment	

The definition excludes products that employ electrical components for practical applications. Such products are highly diverse. They include, for example, household appliances, transportation equipment, and manufacturing equipment. Products of this type are excluded from the definition for two principal reasons: (1) Most of them are as much the products of other industries as they are of the electrical-equipment industry; and (2) they are difficult to bound because electricity is used so widely. Representative products of these types are listed on page 445 in Table 161 in Appendix 3. Also excluded from the definition are electronic products. For the most part, they are the products that apply electricity in electrical form rather than as motion, light, heat, or electrolytic action.

Viewing all of these factors together, an electric motor is considered electrical equipment since it converts electricity into motion; but products that employ electric motors, such as household appliances and automobiles, are not. Similarly, light bulbs are considered electrical equipment but not lighting fixtures. Heating elements are considered electrical equipment but not electrical heating systems. Electrodes are considered electrical equipment but not full electrolytic processing facilities. Electrolytic action, for processes such as making aluminum, is important in the industrial sector of the economy but not in the commercial and residential sectors. The other forms of energy are important in all three sectors.

Essentially, the category *electrical conversion equipment* contains the electrical components that other products use to apply electricity to practical ends. This definition of the electrical-equipment industry ends with these components. That is, this definition ends when the energy takes a non-electrical form. The definition is described on page 437 in Table 160 in Appendix 3 in terms of the Standard Industrial Classification System used by the U.S. Government for classifying products.

The basic-function strategy works well for describing the products of the electrical-equipment industry because most of the products perform a single function. In this sense they are analogous to the components in the electronics industry. This is true even though most products of the electrical-equipment industry are referred to as *equipment*. That term is used because so many of the products are physically large. For example, a power transformer in the electrical-equipment industry may be meters tall and handle megawatts of power (millions of watts). It is treated like equipment. It is incorporated directly into systems, not into other equipment like a component would be, according to the stages of assembly in Table 10. In contrast, a power transformer in the electronics industry is a low-power, small device and is considered a component. It is incorporated into electronic equipment. Yet functionally speaking, the power transformer in the electrical-equipment industry and the power transformer in the electronics industry are identical: they both are used to reduce or increase the voltage of an alternating current.

U.S. Shipments

U.S. shipments of the products of the electrical-equipment industry are \$48 billion (1990); the types of products and levels of shipments are shown in Table 13.[4] The fine print in Table 13 shows the breakdown of the broader product categories where that breakdown is illuminating. The breakdown is finer for headings indented from the left and for the matching shipment values indented from the right. That is the material indented the most from either side is the most finely divided. Each indented column of shipment numbers totals to the value that is just above it and to the right. Apparent totaling errors of \$0.1 billion in the table, and in the other tables in this chapter, are caused by the effects of rounding.

The largest segment of electrical supply equipment serves the basic function of the transfer of electricity, about \$21 billion. This figure has been swollen slightly, however, by inclusion of all electrical wire and insulation, no matter what other products may later be built with them. For example, the wire used in generators, relays, and motors is included. Annual shipments of such wire (called magnet wire) account for about \$1 billion per year of the \$7.4 billion shown. Similarly, some of the \$0.7 billion of insulation products shown is used in the products located under other basic functions.

If the data in Table 13 are viewed at the level of the fine print, the largest segment of U.S. shipments is wire (\$7.4 billion), which is a *passive product* in the sense that it does not operate on electricity to control it or to change its form. The next largest segments are all *active products* and include motors (\$6.9 billion), switchgear (\$5.1 billion), batteries (\$4.6 billion), and transformers (\$4.1 billion), in that order. Automotive electrical equipment is so significant for this industry that it merits a separate breakout under several categories. Across all categories, automotive electrical equipment totals $7.1 billion, exclusive of batteries. The portion of the batteries that should be included in automotive electrical equipment is not known. Thus automotive electrical equipment represents at least 15 percent of all electrical equipment.

Table 13
U.S. Shipments of Electrical Equipment (1990)

	U.S. Shipments ($billions)		
Electrical Supply Equipment			37.4
generation		2.5	
generators and parts	1.3		
automotive generators and parts	1.1		
transfer		21.3	
transformers	4.1		
insulation	0.7		
wire	7.4		
wiring devices	7.0		
automotive wiring and distribution devices	2.1		
manipulation		9.0	
switchgear	5.1		
relays, industrial controls, and apparatus	1.4		
automotive electrical equipment	2.5		
storage		4.6	
storage batteries	3.2		
primary batteries, wet and dry	1.4		
Electrical Conversion Equipment			10.8
motion		6.9	
motors and parts	6.0		
automotive motors and parts	0.9		
light		2.7	
heat		0.9	
elements and electrodes	0.4		
automotive spark plugs	0.5		
electrolytic action		0.3	
Total			48.2

Included within *electrical supply equipment* is all of the equipment that provides $176 billion per year of electricity for the U.S. (1990).[5] The generation of electricity represents 11 percent of the total energy use in the U.S. (1985).[6] The breakdown of electricity use by economic sector served is summarized in Table 14 for 1990.[7] Industrial use is greatest, but only just barely; residential use is nearly as great.

Table 14
Use of Electricity by Economic Sector (1990)

Economic Sector	Percent
residential	34
commercial	28
industrial	35
government (and other)	4
Total	*101

* rounding error

Included within *electrical conversion equipment* is equipment that converts electricity to the four non-electrical forms in which it is principally used: motion, light, heat, and electrolytic action. The breakdown of the use of electricity by end-use energy form is shown in Table 15 for 1985, the most recent year for which this breakdown could be found.[8] The percentage shown for *light (and other)* includes the used of electricity in electrical form, as well. All of the end-use energy types are important in the industrial sector of the economy, and all but electrolytic action are important in the residential and commercial sectors. It is noteworthy that 60 percent of all electricity used powers motors across the entire economy. Since 11 percent (1985) of the national energy budget goes to generate electricity,[9] this means that about 6 to 7 percent of the national energy budget goes to power electric motors. Clearly, electric motors have significant importance to the national economy.

Table 15 provides evidence for the fact that electricity is used primarily as a convenient means of transferring energy and secondarily as an end-use energy form. The devices that use electricity as an end-use energy form constitute a minority. They are principally the electronic products.

Table 15
Use of Electricity by End-Use Energy Type (1985)

End-Use Energy Type	Percent
motion	60
light (and other)	21
heat	16
electrolytic action	4
Total	*101

* rounding error

U.S. ELECTRONICS INDUSTRY

The products of the U.S. electronics industry are highly diverse. Classifying them presents unusual challenges. Most classification schemes employ a hybrid structure, drawing from several of the strategies described earlier. A hybrid structure enables groupings that are easily recognized by broad audiences.

Definition

The Electronic Industries Association (EIA)[10] has defined the electronics industry using a hybrid structure that draws principally from the strategies for the stages of assembly, the consuming sectors, and the societal services. The key categories resulting are outlined in Table 16. Generally speaking, the materials and components categories of the stages of assembly have been merged under the heading *electronic components*. The equipment and systems categories of the stages of assembly have also been merged, and then they have been subdivided according to a combination of the societal services and the consuming sectors. The categories based on the consuming sector are *industrial electronics* and *consumer electronics*. However, virtually all of consumer electronics are devoted to the societal service of entertainment; computers and other high-end electronic products used in the home are included in other categories.[11] A full description of EIA's definition, expressed in terms of the Standard Industrial Classification (SIC) System used by the U.S. Government for classifying products, is contained in Appendix 2, beginning on page 421.

Table 16
EIA's Definition of U.S. Electronics Industry:
Major Categories

- electronic components
- communications equipment
- computers and peripheral equipment
- industrial electronics
- electromedical equipment
- consumer electronics
- electronic-related products and services
  - aerospace equipment
  - automatic controls, industrial apparatus, and other instruments
  - systems integration and computer services
  - motor vehicles
  - electronic-related office equipment

Three of the categories under *electronic-related products and services* reflect electronic content for the named product lines, not captured elsewhere in the table. The three categories are these: (1) aerospace equipment; (2) automatic controls, industrial apparatus, and other instruments; and (3) motor vehicles.

One category in the table, *systems integration and computer services*, contains software products that are so closely related to electronic equipment and systems as to be considered part of those products. Other software products, less closely related, are *not* counted as part of electronic products, even though such software products may be necessary to realize the full range of services from electronic equipment and systems.

U.S. Shipments

Using this structure, U.S. shipments for the electronics industry are shown in Table 17 and Table 18.[12] Major categories of products are shown and partial breakdowns are provided. The approach to showing finer detail through indentation from both the left and the right is the same as that used in Table 13. Some of the finer breakdowns end with an *other* category to pick up product groups too small to merit separate lines in the table. Adjacent to each such *other* entry, the largest subelements are listed, in parentheses, in order of the size of U.S. shipments, beginning with the largest.

Table 17
U.S. Shipments of Electronic Products (1990), Part 1 of 2

			U.S. Shipments ($billions)
Electronic Components[12(a)]			55.2
electron tubes		2.6	
solid-state (semiconductor)		22.0	
integrated circuits	15.5		
other (wafers, packages, transistors, diodes, rectifiers, etc.)	6.5		
parts		10.2	
other components[12(b)]		20.4	
printed circuit boards	6.5		
printed circuit assemblies	7.1		
other (microwave components, power supply converters, cables, etc.)	6.8		
Communications Equipment			66.6
commercial, industrial, and military communications equipment[12(c)]		15.4	
mobile and portable communications equipment	5.6		
space satellite communications systems	2.6		
other receivers and transmitters, except amateur and citizens	2.1		
other (fiber optics, antennas, amateur, citizens, etc.)	5.1		
telephone and telegraph equipment[12(d)]		15.2	
central office switching equipment	4.8		
carrier line equipment	3.6		
other (other switching, switchboards, modems, telephone sets, etc.)	7.2		
intercommunications systems, alarm systems, and traffic control equipment[12(e)]		2.2	
broadcast, studio, and related electronic equipment[12(f)]		1.9	
search and detection, navigation and guidance systems and equipment[12(g)]		31.9	
radar systems and equipment	7.7		
missile and space vehicle equipment	4.4		
navigation systems	4.8		
counter measures equipment	2.8		
light reconnaissance and surveillance systems	2.6		
special electronic and communication intelligence equipment	2.3		
other (sonar, specialized command and control, etc.)	7.5		
Subtotal [Part 1 of 2]			121.8

The middle columns in Table 17 and Table 18 provide an intermediate level of detail that is useful for determining the largest sectors of the electronics industry. Among the equipment categories, the largest sectors are, first, *search and detection, navigation and guidance systems and equipment* ($31.9 billion) and, second, *computers* ($30.6 billion). The next largest sectors are electronic products for

Table 18
U.S. Shipments of Electronic Products (1990), Part 2 of 2

	U.S. Shipments ($billions)		
Computers and Peripherals			56.2
computers[12(h)]		30.6	
general purpose	23.6		
special purpose	3.2		
parts for computers	3.8		
peripherals[12(i)]		25.6	
magnetic and optical data storage devices and media	9.3		
printers and plotters	4.9		
terminals	3.4		
parts for peripheral equipment	5.8		
other (other input/output devices, etc.)	2.2		
Industrial Electronics[12(j)]			19.7
control and processing equipment		9.6	
control equipment	4.1		
processing equipment	1.7		
display and control instruments	3.8		
testing and measuring equipment		5.4	
other (nuclear electronic, robots, etc.)		4.8	
Electromedical Equipment[12(k)]			7.5
diagnostic		2.7	
other (patient monitoring, irradiation equipment, therapeutic equipment)		4.8	
Consumer Electronic Products[12(l)]			7.2
television receivers		5.0	
other (recorders, phonographs, radio/TV chassis, audio/video tape, etc.)		2.2	
Electronic-Related Products and Services[12(m)]			54.5
aerospace equipment (aircraft, missiles)		27.6	
automatic controls, industrial apparatus, and other instruments		1.1	
systems integration and computer services		21.3	
motor vehicles		2.8	
electronic-related office equipment		1.7	
Subtotal [Part 2 of 2]			145.1
Total [Part 1, from Table 17, plus Part 2]			*266.8

* rounding error

aerospace equipment ($27.6 billion), *peripheral equipment* for computers ($25.6 billion), and *systems integration and computer services* ($21.3 billion). Thereafter come *commercial, industrial, and military communications equipment* ($15.4 billion) and *telephone and telegraph equipment* ($15.2 billion). Note the strong presence of computer and communications applications in virtually every one of these largest equipment sectors. Among components, the largest sector is *solid-state (semiconductor) components* ($22.0 billion). These components provide the basis for nearly every product of the electronics industry.

VIEW OF THE INDUSTRIES BY TECHNOLOGY

The measurement needs of the electronics and electrical-equipment industries are best understood by looking at them by technology. The reason for this, as noted above, is that measurement needs tend to follow technologies. Both a breakdown by materials and a breakdown by frequency are useful. A breakdown by materials is particularly convenient for grouping materials and components with common measurement needs. A breakdown by frequency is particularly convenient for grouping equipment and systems that have common measurement needs.

The discussion that follows describes the important subcategories under both materials and frequency and indicates where within this document these important subcategories are discussed.

Materials

Electronic and electrical materials can be broken down into a number of categories, as shown in Table 19. Six classes are shown along with key examples. The classes are defined by the predominant property of interest for the material.

Table 19
Electronic and Electrical Materials

Class of Properties	Electronic Material
conducting	semiconducting normal conducting/resistive superconducting
insulating	dielectric
magnetic	magnetic
optical	optoelectronic semiconducting electro-optic solids liquid crystals electroluminescent magneto-optic dielectric
thermal	thermionic
chemical	electro-chemical

Under the conductive class, conducting and resistive materials are listed on a single line because they differ principally in the degree of their conductivity. Some materials serve more than one class:

semiconducting materials serve both the conducting and the optical classes; dielectric materials serve both the insulating and the optical classes.

Importance

It is useful to consider the importance of these materials to electronic and electrical products. Importance can be measured in many ways. One way is to determine *which single material is most critical to the performance of the primary function of a product*, and then to compare the values of U.S. shipments of products dependent on the different materials. Viewed in this way, semiconductor materials are by far the most important materials for electronic products. In contrast, for electrical products, magnetic materials are the most important, through their impact on motors, generators, transformers, relays, and other products. Semiconductor materials and magnetic materials are addressed separately in Chapters 4 and 5 of this document.

A second way to measure importance is to ask *which materials are present in the greatest value of U.S. shipments* of products, however critical or incidental the role of those materials is to the performing of the primary function of those products. Against this criterion, normal conducting and dielectric materials are the most important for both electronic and electrical products since they are present in virtually all electronic and electrical products.

It is useful to recognize that the sensitivity of particular products to the performance of their constituent materials can vary considerably from product to product. Some products demand unusually high levels of performance from materials that may seem at first glance to play only a secondary role. Dielectric materials provide a good example. They become critical materials whenever they are subjected to high electrical stresses or to high frequencies. High electrical stresses may arise when dielectric materials are asked to provide electrical insulation between points of very high voltage or very close proximity. The former case arises in electrical power applications. The latter case arises in integrated circuits, where circuit elements are measured in micrometers (10^{-6} meter) or even nanometers (10^{-9} meter). High-frequency properties become critical in microwave and optical circuits. Dielectric materials are not addressed in a separate chapter in this document. Rather, they are discussed in several chapters along with the products that place the special demands upon them.

Similarly, the performance of conducting materials becomes critical whenever they are asked to carry high current densities, that is, whenever the total current level carried is high for the size of the conductor carrying the current. This situation may arise when the current is very high or when the size of the conductor is very small. Again, these two extremes arise, respectively, in power system components and in interconnecting "wiring" within integrated circuits. Like dielectrics, conducting materials are not addressed in a separate chapter in this document. Rather, special concerns related to conductors are discussed in chapters addressing affected products.

Sometimes the current levels required in a product are simply too high for a normal conductor to handle. In these cases, such as in powerful electromagnets, only superconducting materials may have the necessary current-carrying capacity. Because superconductors are so fundamentally different from conducting materials, they are addressed separately in Chapter 6.

Optical materials, many of which are based on semiconductors, underlie an increasingly broad range of electronic products. Many needed functions in electronics can be performed only by light or best by light. Optical materials are highly diverse and are discussed as part of the broader treatment of products operating at light frequencies in Chapters 8, 9, and 10 principally.

Table 20
Electronic Materials and Supported Components

- semiconducting
  - diodes
  - transistors
  - integrated circuits
  - lasers
- conducting/resistive
  - wires
  - radio-frequency waveguides
  - resistors
  - switches
  - connectors
- superconducting
  - Josephson junctions
  - integrated circuits
  - electromagnets
- dielectric
  - optical-frequency waveguides
  - capacitors
  - insulators
- magnetic
  - inductors (coils)
  - transformers
  - motors
  - generators
  - relays
  - recording heads
- optoelectronic
  - modulators
  - optical switches
  - optical-fiber sensors
  - liquid-crystal displays
- electroluminescent
  - television picture tubes
  - computer monitors
- magneto-optic
  - optical storage diskettes
  - optical fiber sensors
- thermionic
  - television picture tubes
  - computer monitors
  - microwave power tubes
- electro-chemical
  - batteries

Thermionic materials and electro-chemical materials are not discussed in this document. Thermionic materials are the critical elements in electron tubes and cathode-ray tubes, such as those used for microwave transmitters, television receivers, and computer monitors. Electro-chemical materials are the critical elements in batteries. Batteries are a vital part of the future of electronics. Increasingly, electronic products are being made small and portable; they rely increasingly on batteries for power.

Components Based on the Materials

Components made from electronic and electrical materials are highly diverse. Table 20 repeats the electronic materials introduced in Table 19 and provides examples of important components made from each. Virtually all components are made from more than one material; they are listed in Table 20 under the material most responsible for their primary function.

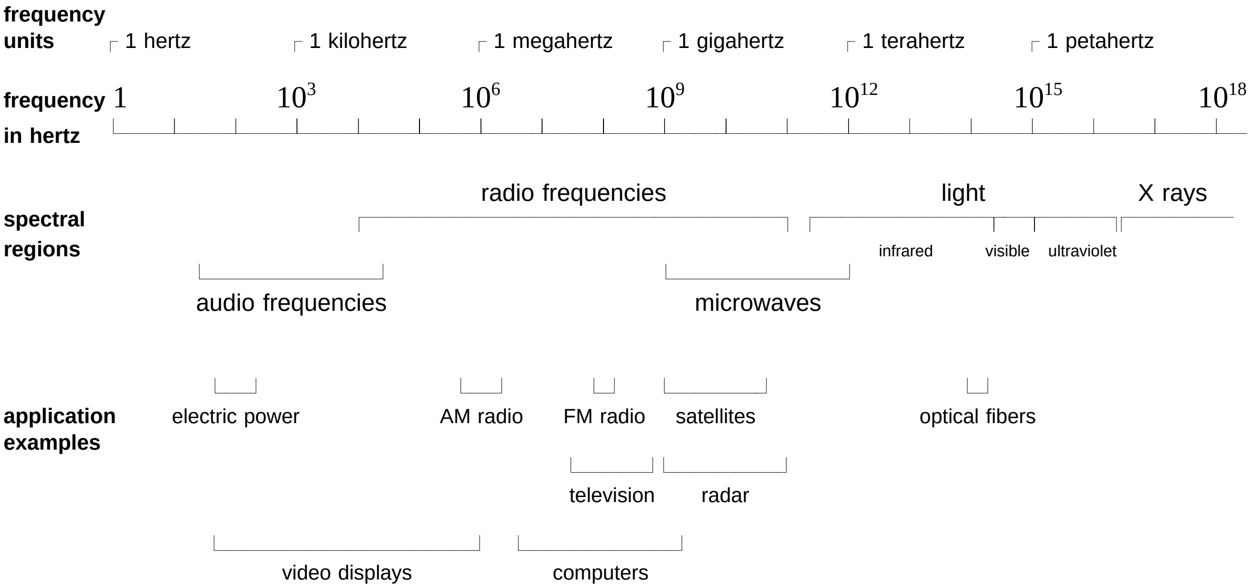
Frequency

Electronic products operate over a very wide range of frequencies from 0 hertz to above 10^{18} hertz. Electrical products operate only at the lowest frequencies, mostly at 50 to 60 hertz and seldom above 400 hertz.

Figure 2 below shows this broad spectrum of frequencies along the horizontal "ruler" near the top of the figure. Above the "ruler" are the names of the frequency units used to express frequencies. Thus 10^9 hertz is called 1 gigahertz. These frequency units are used throughout this document. Under the "ruler" are the names commonly given to important regions within the frequency spectrum: audio frequencies, radio frequencies (or radio waves), microwaves, light (or lightwaves), and X rays.

Thereunder are examples that show the range of frequencies over which familiar products operate. Frequencies right down to 0 hertz (direct current) are important, even though the figure itself begins with 1 hertz.

Figure 2
Overview of the Frequency Spectrum



Electronic products accomplish their tasks by employing electrons and electromagnetic fields. The two are closely related. Electromagnetic fields are produced by the motion of electrons. When electrons accelerate or decelerate, electromagnetic fields are radiated; that is, the fields move through space, carrying energy. These radiated fields provide the means for wireless communications systems, radar, lasers, light bulbs, and many other electronic and electrical products.

The motion that leads to radiation is produced by two principal mechanisms: (1) the back and forth action of electric currents in materials, or (2) the movement of electrons during a change from one discrete energy state to another within atoms, molecules, or crystals. An example of (1) is the alternating current set into motion in the antenna of a television station. This alternating current causes a television signal to be radiated (or transmitted) to listeners. An example of (2) occurs when an orbiting electron in an atom jumps from one energy state to another and releases electromagnetic energy equivalent to the energy difference. These energy states are roughly equivalent to different orbits around the nucleus of the atom; the orbits closer to the nucleus have lower energy than the orbits that are more distant. Generally speaking, process (1) is used to generate electromagnetic radiation with frequencies below and up into the microwave region, and process (2) is used to generate electromagnetic radiation at higher frequencies, particularly those of light and X rays. The frequencies produced by process (1) match the rate of back and forth motion of the electrons. The frequencies produced by process (2) are proportional to the change in energy level of the electrons as they change energy states within atoms, molecules, or crystals.

Table 21 shows the primary focus of the chapters in this document *in terms of* the materials and the frequency ranges discussed above. Chapters 4 through 6 are most simply characterized by the *materials* that they address. These materials are used for products that operate at a broad range of frequencies from dc to light. Chapters 7 through 10 are most simply characterized by the *frequency ranges* that they address. The remaining chapters cut across so many materials and frequency ranges that they are not so easily characterized. For all of the chapters, the discussion in each chapter describes in detail the broad range of products affected by the technologies in that chapter.

Table 21
Materials and Frequency Focii of Chapters

		Principal Focii	
Chapter	Title	Materials	Frequency Ranges
Chapter 4	Semiconductors	**semiconducting**	multiple
Chapter 5	Magnetics	**magnetic**	multiple
Chapter 6	Superconductors	**superconducting**	multiple
Chapter 7	Microwaves	multiple	**microwave**
Chapter 8	Lasers	multiple	**light**
Chapter 9	Optical-Fiber Communications	multiple	**light**
Chapter 10	Optical-Fiber Sensors	multiple	**light**
Chapter 11	Video	multiple	multiple
Chapter 12	Electromagnetic Compatibility	multiple	multiple

FUTURE EDITIONS

Not all products of the electronics industry and the electrical-equipment industry with critical measurement needs are addressed in this First Edition of *Measurements for Competitiveness in Electronics*. Future editions are contemplated and will contain additional chapters to cover other important areas.

ENDNOTES

1. All shipments figures in the table represent industry data, rather than product data, except those for the electronics industry. Industry data reflect the value of all products and services sold by establishments in the named industry, whether or not the products are classified in that industry. Product data reflect the value of all products classified in the named industry sold by all industries. Industry data are usually several percent higher than product data for a given product classification. For this reason the value of all chemical products shipped (product data), if it could be known, would likely be closer to the value of all electronic products shipped. Note, too, that there is some overlap in the products listed in the table. Some electronic products are included in the automotive and aerospace industries. This overlap arises because there is no set of codes in the Standard Industrial Classification (SIC) System, on which all of the figures in the table are based, that is devoted exclusively to the electronics industry. The superscripts in the table refer to the notes that follow: (a) *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 11-1 (January 1991). (b) *1991 Electronic Market Data Book*, Electronic Industries Association, pp. 4, 113 (1991), and reflecting a small correction from EIA. (c) The figures shown include both the motor-vehicle and supporting parts industries. *1992 U.S. Industrial Outlook*, pp. 36-1 and 36-15. (d) The employment figure is from 1989, the year of most recent data, and is used as an estimator for 1990. *1992 U.S. Industrial Outlook*, p. 4-1. (e) *1992 U.S. Industrial Outlook*, p. 21-1.

2. *1991 Electronic Market Data Book*, Electronic Industries Association, p. 4 (1991).

3. Preliminary estimate for 1990 from the Edison Electric Institute, as of December 9, 1991.

4. The references used for determining U.S. shipments were all developed by the Bureau of the Census, U.S. Department of Commerce and include the *Current Industrial Reports*, the *1987 Census of Manufactures*, and the *1987 Annual Survey of Manufactures*. Data on shipments for 1990 were not available for all products; in such cases the two years of most recent data (selected from 1990, 1989, 1988, 1987, and 1982 in different cases) were used in a straight-line projection to obtain an estimate of the 1990 level. The data are collected in the categories of the Standard Industrial Classification (SIC) System. Because only part of the products in some of the SIC categories are applicable to the electrical equipment industry, and because the fraction applicable is not known with precision, estimates had to be made to develop a total figure for the industry. This has led to higher levels of uncertainty in some product categories than in others. For example, the actual value associated with *electrical insulation*, in particular, may vary considerably from that reported here.

5. Preliminary estimate for 1990 from the Edison Electric Institute, as of December 9, 1991.

6. *Electrotechnology Reference Guide: Revision 1*, Resource Dynamics Corporation, prepared for the Electric Power Research Institute, Report No. EPRI EM-4257, p. S-2 (August 1988).

7. *Statistical Yearbook of the Electric Utility Industry/1990*, Edison Electric Institute, p. 44 (October 1991).

8. *Electric Motors and Drives: Markets, Trends and Opportunities: Phase 1 - Motors*, Resource Dynamics Corporation (prepared for the Electric Power Research Institute), p. 4 (January 1991).

9. *Electrotechnology Reference Guide: Revision 1*, Resource Dynamics Corporation, prepared for the Electric Power Research Institute, Report No. EPRI EM-4257, p. S-2 (August 1988).

10. *1991 Electronic Market Data Book*, Electronic Industries Association (1991).

11. In its *1991 Electronic Market Data Book*, EIA groups the categories of *computers and peripherals*, *industrial electronics*, and *electromedical equipment* together under a parent category called *computers and industrial electronics*.

12. All data in Table 17 and Table 18 come from the *1991 Electronic Market Data Book*, Electronic Industries Association (1991) from the pages shown. These data in turn are based almost exclusively on data collected by the Bureau of the Census of the U.S. Department of Commerce. The superscripts in the table refer to the notes that follow: (a) p. 68; (b), p. 68, however, the breakdown shown was estimated by NIST directly from data in "Semiconductors, Printed Circuit Boards, and Other Electronic Components, *Current Industrial Reports*, No. MA36Q(89)-1, Bureau of the Census, U.S. Department of Commerce, p. 1 (September 1991); (c) p. 36; (d) p. 38; (e) p. 36; (f) p. 36; (g) p. 37; (h) p. 48; (i) p. 52; (j) p. 48; (k) p. 51; (l) based on revised data computed with EIA assistance from the *Current Industrial Reports*, of the Bureau of the Census, U.S. Department of Commerce, including "Radio and Television Receivers, Phonographs, and Related Equipment", No. MA36M(89)-1, pp. 1-3 (October 1991), "Computers and Office and Accounting Machines", No. MA35R(89)-1, p. 4 (October 1991), and "Communication Equipment, and Other Electronic Systems and Equipment", No. MA36Q(89)-1, p. 1 (September 1991); (m) pp. 99-101.

CHAPTER 4

SEMICONDUCTORS

May 1991

Chapter 4

SEMICONDUCTORS

SUMMARY

Electronics affects every aspect of life today. Electronic systems are integral parts of the nation's defense, communication systems, transportation, medical care, banking, manufacturing, entertainment, etc. Electronics is one of the largest manufacturing segments of the U.S. economy, shipping some \$266 billion of products in 1990 and growing at 4.7 percent annually. It employs 1.94 million people (1990), more than twice as many as the chemical industry, which is slightly larger in shipments. Semiconductor devices are absolutely essential components of all electronic products today. Without semiconductors, electronic products and systems cannot be made or operated. The U.S. semiconductor industry produced \$29 billion of semiconductor devices in 1989 (42 percent of the world market) and is growing at a rate of 16 percent annually. Japan produced \$33 billion, 46 percent of the world market. From 1985 to 1989 the U.S. lost 11 percentage points of the world market and Europe lost nearly 3; Japan and the rest of the world were the gainers. [*Editor's note:* After the writing of this chapter in May 1991, the U.S. achieved small gains in its world market share of production in both 1991 and 1992. The U.S. reached approximate parity with Japan in 1992, with strong competition continuing.] U.S. suppliers sold \$4.5 billion of semiconductor-manufacturing equipment in 1989, a 48-percent share of the world market. [*Editor's note:* After the writing of this chapter in May 1991, the U.S. achieved an increase to above 50 percent in 1992.]

The competitive problems faced by the U.S. semiconductor-related industries are well known. Many significant studies have emphasized the seriousness of the U.S. deficiency in semiconductor-manufacturing technologies. Among the studies are those by National Research Council (NRC) on *Advanced Processing of Electronic Materials in the United States and Japan* (1986), by the Defense Science Board (1987), and most recently by the National Advisory Committee on Semiconductors on *A Strategic Industry at Risk* (1989) and *Preserving the Vital Base* (1990). Worldwide, the \$71-billion semiconductor-device market, the \$9.7-billion semiconductor-manufacturing-equipment market, and the \$9.5-billion semiconductor-materials market are at stake (all 1989). Japanese companies, while highly competitive with one another, are the beneficiaries of a variety of well-funded long-term programs of generic technological development initiated by their government. The U.S. has responded by establishing SEMATECH, a consortium to rebuild U.S. excellence in semiconductor manufacturing. About half of its funds come from the 14 member companies and half from the U.S. government.

Semiconductor integrated circuits dominate today's semiconductor products. Integrated circuits may incorporate just a few transistors or millions of transistors, plus other components, all formed on a single piece of semiconductor material and interconnected to perform a useful electronic function. Integrated circuits are sometimes called *chips* or *ICs*. Modern integrated circuits are made with manufacturing processes that must form circuit elements with submicrometer dimensions (below 10^{-6} meter) and that press science and skill to very expensive limits. The breadth of science and technology used in semiconductor manufacturing is rapidly expanding. Problems of imperfect understanding and inability to measure what is being done abound. Integrated circuits today are built with great ingenuity by using very complex processes to deposit and remove surface films. The chemistry and physics of these processes are not well understood. The surfaces on which films are deposited affect both the deposition processes and the properties of the semiconductor devices in ways that are evident but not understood; these effects are becoming ever more significant. A clear view of what is being made requires electron microscopes, but accurate dimensional measurements with these are not now possible. Soon it will be necessary to create ultra-small patterns using x-rays or other new techniques that will require totally new measurements. Control of contamination from particles and stray impurities is pressing the limits of measurement and control. As device dimensions shrink, the details of their behavior change significantly. Previously negligible effects become important and must be accounted for in engineering design. New device theory and computer models of device behavior are needed.

Semiconductor manufacturing is a measurement-intensive business. The industry estimates that 30 percent of the cost of complex integrated circuits is in testing. As semiconductor devices become smaller and chips become complex, the demands on measurement technology increase. Better sensitivity, improved accuracy, totally new measurements, and the pressure to automate the manufacturing process are all needs that must be met. New materials and new processes bring with them new measurement problems that must be solved.

Many kinds of manufacturing equipment contain measurement systems to control their own operation. The strong present-day move to greater automation means that these measurements have a very great effect on the overall control of the entire manufacturing process. It is difficult enough to measure many of the process variables with human aid. It is much more difficult to obtain accurate data without it. All of these factors add greatly to the challenges of providing measurement capability that is needed to support U.S. industries that make manufacturing equipment, semiconductor materials, and devices.

These needs are enumerated in the reports of recent workshops and technology assessments, which almost invariably include requests for NIST's help in finding solutions to them. The reports include the following: *Semiconductor Research Corporation Technology Trends Assessment*, Wild Dunes, SC (September 6-8, 1989); *SEMICON/Southwest 90 Technical Conference Proceedings: Metrology for Advanced Materials/Process Characterization*, Dallas, Texas (January 31-February 1, 1990); *Semiconductor Research Corporation Topical Research Conference on Metrology for Semiconductor Manufacturing*, Santa Fe, New Mexico, (February 21-22, 1990); and *Workshop on Silicon Materials for MEGA-IC Applications*, New Orleans, Louisiana (January 30-31, 1991).

DESCRIPTION OF THE TECHNOLOGY[1]

Semiconductor technology continues to advance on a wide front. For example, the scale of dimensions used in semiconductor-manufacturing processes for making integrated circuits (ICs) is rapidly shrinking, for two reasons. Smaller devices are faster, and more of them can be put on one chip. Both of these results allow electronic systems to have greater levels of performance. Table 22 shows this evolution.

Semiconductor manufacturing is a complex sequential process of up to 500 steps. Most of these steps cause irreversible changes, so if a processing error occurs during any step the wafers are useless.

Table 22
Evolution of Electronic Devices

Dimension	Technology	Time
centimeter (cm)	Vacuum Tube	1925-1960
millimeter (mm)	Transistors	1950-
micrometer (µm)	Integrated Circuits	today
sub-micrometer	Advanced Integrated Circuits	1990s

Measurement is not always possible after every step. Even when measurements are possible, inaccuracies can cause defective material to appear to be within process control limits, thus leading to spending money on further processing of material that is in fact scrap. Measurement uncertainties also limit the degree of process control that is possible.

Each new generation of semiconductor devices presents manufacturing difficulties that must be overcome. Yet the first company to sell a new device has the great advantage of being able to get a high price for the product, until the entry of competitors in the market starts price competition. In deciding when to enter the market, the advantage of early entry must be weighed against the consequences of possible loss of control of the process and inability to deliver the product.

Improved measurement allows better control over the process and lets one get to the market earlier. Product yields at the early stages of manufacturing are typically from 5 to 15 percent. Experience with the new product usually improves process control so the yield rises rapidly (and costs drop in proportion). This is a further incentive for early product introduction. Early high prices help pay the costs of low yield. By the time the competition enters the market, the first producer has his costs well in hand and can make money even at reduced prices. The later entries must suffer low yields and at the same time low prices until yields can be raised. This pattern has been repeated many times in the semiconductor industry and is well known to both the winners and the losers in the market.

Materials

Semiconductor devices are made in a very wide spectrum of types to serve varying applications. All, however, depend on the properties of semiconductor[2] materials for their operation. By far the majority of semiconductor devices are made from silicon, which is one of the most abundant elements. Compound semiconductors are used for optoelectronic devices and for very high-speed devices, including microwave integrated circuits. Compound semiconductors are usually formed of arsenic, antimony, phosphorus, aluminum, gallium, or indium in various combinations, though compounds of many other materials are semiconductors having specialized uses. Gallium arsenide (GaAs) is the most commonly used compound-semiconductor material. Table 23 shows some attributes of silicon and the more common compound-semiconductor materials.

Regardless of the semiconductor material that is used, devices can be made only if the material is extraordinarily pure. Typically, semiconductor-grade silicon may contain oxygen at a few parts per million, carbon at less than one part per million, and metallic impurity elements at the part per trillion

Table 23
Silicon Compared with Compound Semiconductors

	Silicon	Compound Semiconductors
Number of starting elements	one	two to six
Availability of starting materials	very abundant	scarce (aluminum less so)
Available purity of starting elements	excellent	fair to poor
Speed of devices	fast	2-3 times faster
Ease of fabrication	difficult	very difficult
Ability to make light	no	yes
Ability to detect light	yes	yes
Typical materials	silicon	gallium arsenide gallium aluminum arsenide indium phosphide

level[3]. Certain electrically-active impurities (boron, phosphorus, antimony, and aluminum) are controlled to a few tens of parts per trillion. As the silicon is processed into devices, these latter four elements (called dopants, to distinguish them from all other impurities) are intentionally added under strict controls to adjust the electrical properties of the various regions of silicon that form the finished product. The dopants are themselves in a highly pure state, to avoid unwanted impurities being inadvertently added to the silicon.

Various other chemicals are used in semiconductor-device manufacturing. These include solvents (alcohols, acetone, etc.), acids (nitric, sulfuric, hydrofluoric, etc.), gases (nitrogen, oxygen, hydrogen, argon, and numerous specialized gases), and water. To consider the latter as a chemical may seem surprising, but, like the rest of these materials, it must be supplied in very pure condition to avoid contamination of the silicon. Pure water is an aggressive solvent.

Manufacturing

Semiconductor manufacturing is done in special facilities (clean rooms) in which the temperature, humidity, and particulate-contamination levels are tightly controlled. All equipment and facilities are specially designed to avoid contamination. Workers are dressed in special clothing, because people continually shed particles from the skin. This fanatical attention to cleanliness is necessary to avoid contamination of the silicon. Even a particle smaller than a hundredth as large as the cross-section of a hair can cause an integrated circuit to be defective. Pollen, germs, smoke, and other common particles are much larger in size.

The process[4] of making semiconductor devices starts with a very flat, round wafer of silicon. The material supplier has made the wafer from a larger ingot typically three or four feet long and composed of a single crystal[5] of silicon. The basic electrical properties of the wafer are controlled by the material supplier to meet the device manufacturer's specifications. The wafer is typically 100 to 200 millimeters (4 to 8 inches) in diameter and flat to within less than 3 micrometers. This departure from flatness would, if the wafer were the size of a baseball field, amount to only an eighth of an inch.

In addition, the structural perfection of the wafer is important. That is, every atom in the wafer should be in its proper position with no atomic sites unoccupied. This ideal can be closely approached but not entirely met. The presence of oxygen leads to an imperfection called oxygen-induced stacking faults, for example. These and other crystalline imperfections at too great a concentration cause problems in device manufacture. So do minute departures from flatness on an atomic scale, called micro-roughness. This fault has only been known to be important in recent years. Its causes and cures, and its measurement, are still imperfectly understood. And, like everything else used in making semiconductor devices, wafer surfaces must be free of particulate contamination.

The device process consists of forming various films in or on the wafer surface, creating patterns in each film as it is created. This builds a structure of perhaps 20 different layers, all interconnected laterally and vertically to form a complex network of individual transistors[6] and electrical connections among them. Films may be of insulating materials such as silicon dioxide (quartz) or of conducting materials such as heavily doped silicon or metals. They are applied by evaporation or sputtering in a vacuum (for metals), or by chemical processes involving reactions promoted in a mixture of gases by heat or electrical energy in a partial vacuum.

The patterns are formed by first coating the surface of the wafer with a photoresist. This is a paint-like material that, when exposed to ultraviolet light through a patterned glass plate, forms on the wafer surface an image of the pattern on the plate. The process is much like taking a photograph, except that the image in this case is either black (resist remains after development) or white (no resist). The difficulties in this process can be appreciated by considering that the details in the image are as small as 0.6 micrometer (less than 1/100 the diameter of a hair), that a pattern must align with those formed earlier to a tenth of that amount, and that the machine that does this work has about a quarter of a second to align and expose each image. (Lenses that can form such detailed images can do so only over an area a little more than half an inch square, so the process must be repeated many times to cover the entire wafer.) After the pattern is reproduced in the photoresist, leaving resist in some places and not in others, the film thus exposed can be removed by chemical etching, covered by material added in several ways, or modified by ion implantation. The rest of the resist is then stripped away, and the next film is applied. The procedure of this paragraph is repeated as many times as necessary, typically about 20 in more advanced products.

The last few layers form the electrical connections among the transistors that comprise the circuit, and also provide small connection pads around the edge of each completed integrated circuit. The circuits may now receive an initial electrical test by contacting the pads with an array of needles. The location of defective circuits is identified so they will not be processed further. Depending on the size of each circuit and the size of the wafer, there may be from 70 to several hundred circuits in a regular pattern on the wafer. The individual circuits are then separated by cutting the wafer with a fine diamond saw.

At this point, the integrated circuit (or other device) is electrically completed but is still exposed to contamination and physical damage. It must be enclosed in a protective “package”, which may be a metal and ceramic enclosure or molded from plastic. In addition to protecting the silicon chip, the package provides more robust electrical connections to the circuit and a pathway for removing the heat generated by the device. The latter is crucial for large integrated circuits and for any other devices that handle appreciable amounts of power.

Complete final testing is only possible after the chip has been packaged. This is the product that is sold to the customer. Any effect of the package on the performance of the device, which may be slight or great, must be taken into account in the device manufacturer’s specifications for the device. More important for power device types is the heat dissipation provided by the package. Extensive testing heats the chip, and without the package the device can become too hot and be damaged. The economics of chip and final testing is changing. Increased attention to process control raises yields sufficiently in some cases to make the cost of such testing unattractive. The payback is better for increased measurement early in the process, which leads to making the product right the first time.

Market Environment

Semiconductor devices have been made commercially for over 35 years. The course of development started with making individual transistors and diodes (which early replaced most vacuum tubes in electronic equipment), continued with the integration of these devices into simple integrated circuits containing only a few transistors, and advanced to the manufacture of today’s integrated circuits that contain as many as ten million transistors. Integrated circuits containing well over 200 million transistors are in development. Throughout this period the cost of semiconductors, per electronic function that they provide, has dropped dramatically.

The driving force behind this remarkable accomplishment has been the demand of electronic systems manufacturers and their customers for products providing more and more power, speed, convenience, etc., all at low costs. The growth in power of computers stems directly from this process. Semiconductor-device manufacturers have met this demand by making the individual devices smaller and smaller, and packing more of them on a single piece of silicon. Synergy between the computer industry and the integrated-circuit manufacturers has been essential. Computers were necessary to design integrated circuits, which quickly became far too complicated to design manually. The resulting integrated circuits allowed more powerful computers to be made, which were used for the next generations of circuit design, and so on. Today, almost without exception all semiconductor-manufacturing equipment contains computer-based controls. The equipment is tied together and directed by computers. Material flow and process control in semiconductor plants is managed by computers. For many years the final testing of the product has only been possible with computer-controlled equipment.

The march toward smaller features in increasingly complex circuits has brought with it very high costs. Clean-room space costs from $2000 to $7000 a square foot, and is expensive to operate and maintain. Equipment is both expensive and quickly becomes obsolete by advances in the technology. Integrated-circuit fabrication facilities now cost $500 million or more each for a commercially viable scale of operation, and the billion-dollar plant is on the horizon. Individual processes are more complex, and there are more of them. The allowable variances in materials properties and in process execution are becoming smaller.

The steady thrust toward greater numbers of increasingly smaller devices on chips that do ever more complex things, made cheaper with more expensive facilities and tools in fierce competition with every advanced nation in the world, must eventually approach some physical limits. In fact, the rate at which circuit complexity increases has slowed somewhat. For years circuits doubled in complexity every two years. Lately it has taken three years, and soon it will require four. But at least until the end of this decade there seems to be no insuperable obstacle to the continuation of these trends.

Nevertheless, one should bear in mind that the great majority of electronic products are manufactured using semiconductor devices that are not made with leading-edge technology. Logic circuits, such as those used in desktop computers, are only about one tenth as complex (in terms of the number of transistors in a single integrated circuit) as the most advanced memory devices. A very large part of the world's electronic work is done with systems that use devices of older generations. They are well proven, cheap, and perfectly adequate. In this context, advanced technology allows these older products to be made inexpensively and reliably in great volumes. Products that are very difficult to make today at any cost may become easy to manufacture with the passage of a few years. They become the cash cows that help pay for the newest and latest advances.

ECONOMIC SIGNIFICANCE AND U.S. COMPETITIVENESS

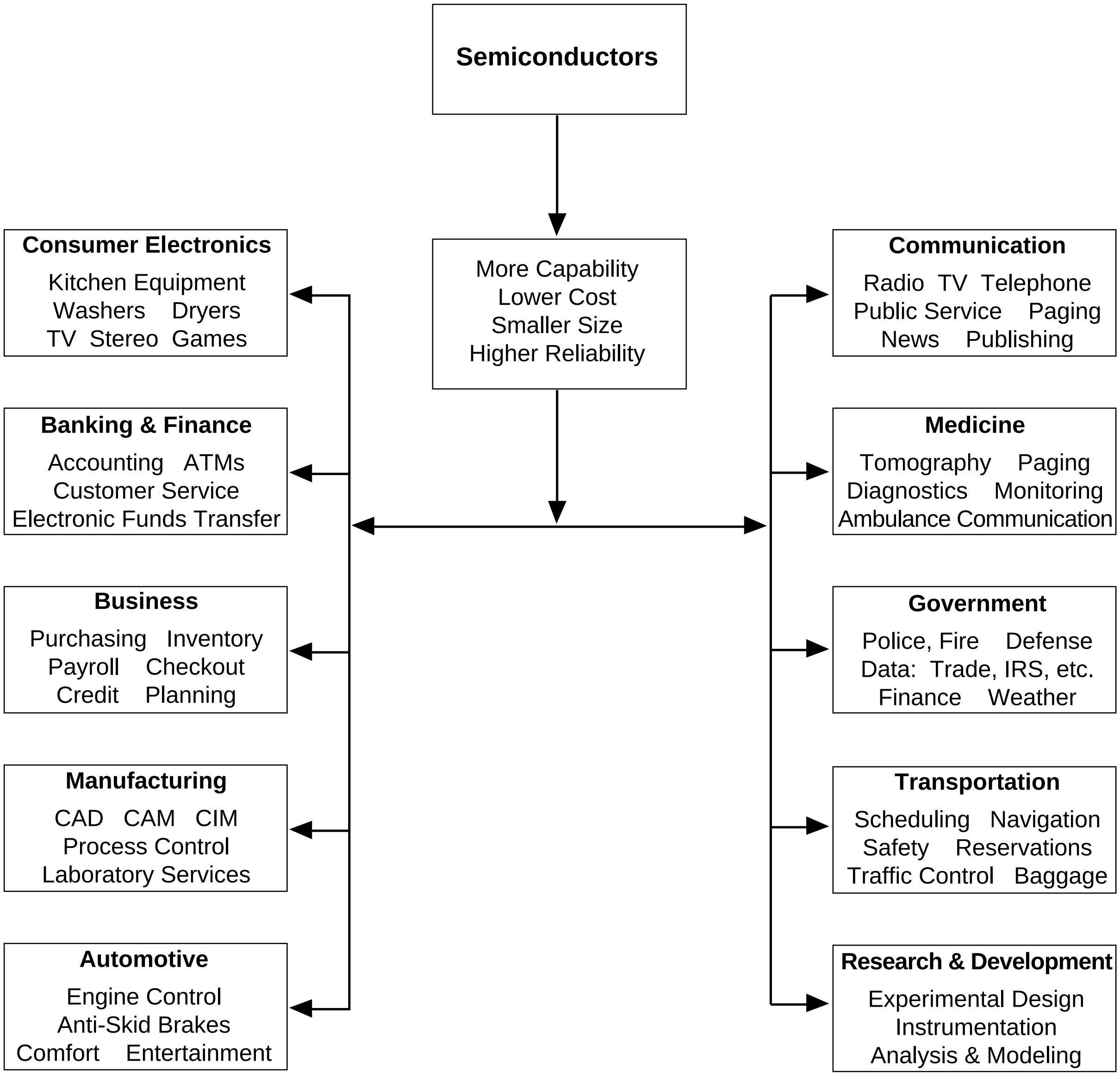
Semiconductor technology is unique in being not only the foundation of a large industry in the U.S. and the world but also the key technology required for making *all* electronic products. Figure 1 illustrates the point. Electronics is one of the largest manufacturing segments of the U.S. economy, shipping about $266 billion of products in 1990 and growing at 4.7 percent annually.[7] This industry employed 1.94 million people in 1990, more than twice as many as the chemical industry, which is slightly larger in shipments. Total world production of electronic products is believed to be somewhat greater than twice U.S. production.

Semiconductor devices are absolutely essential components of all electronic products today. These products in turn are used by every human activity in the industrial world. "Over the past three decades, no single field of science or engineering has had a greater impact on the national productivity and quality of life than has semiconductor electronics. Semiconductors have revolutionized the communications, entertainment, and transportation industries, and have created the computer industry."[8] Without semiconductors we would have little or no communication, transportation, medical care, entertainment, banking, defense, or other services, as we know them today. We have become dependent on electronics and thus dangerously vulnerable as a nation to the loss of the semiconductor industry. Just such a loss is occurring.

"Information, communications, computation, and control technologies are also crucial elements of our national defense, both in its management and its weapons systems. Because it is so pervasive, this area of engineering can be considered basic to all others as the world moves into the information age."[9]

The trend in market shares for production of semiconductor devices is illustrated in Table 24.[10] Data in the table for years before 1970 include merchant producers, exclude captive producers, and are listed by nationality of the supplier rather than by region of production. Data for 1970 and later include both merchant and captive producers and are listed by region of production. Data for 1990 and beyond are estimated. The table shows that the U.S. semiconductor industry produced $29 billion of semiconductor devices in 1989 (42 percent of the world market).[11] U.S. production levels from 1985 to 1989 corresponded to a compound annual growth rate of 16 percent. Japan produced $33 billion of products in 1989, or 46 percent of the world market, with a compound annual growth rate of 32 percent from 1985 to 1989. The net effect of these figures is that from 1985 to 1989 the U.S. lost 11 percentage points of the world market, and Europe lost nearly 3; Japan and the rest of the

Figure 3
Contribution of Semiconductors to the Economy



world were the gainers. [*Editor's note:* After the writing of this chapter in May 1991, the U.S. achieved small gains in its world market share of production in both 1991 and 1992. The U.S. reached approximate parity with Japan in 1992, with strong competition continuing.[12]] While the U.S. has been holding its competitive position in basic research, the principal competitive challenge continues to come in product implementation.

In the related area of semiconductor-manufacturing equipment, U.S. suppliers sold \$4.5 billion of products in 1989, a 48-percent share of the world market of production. [*Editor's note:* After the writing of this chapter in May 1991, the U.S. achieved an increase to above 50 percent in 1992.]

Table 24
Semiconductor Industry Production of Devices
($millions)

Year	US		Japan		Europe		Rest of World		Total
	$M	%	$M	%	$M	%	$M	%	$M
1952	19	100							19
1955	40	100							40
1960	594	97.2	8	1.3	9	1.5			611
1965	1,122	69.3	183	11.3	315	19.4			1,620
1970	2,151	65.8	587	18.0	532	16.3			3,270
1975	3,367	60.1	1,081	19.3	1,155	20.6			5,603
1980	9,839	54.2	5,359	29.5	2,892	15.9	80	0.4	18,170
1985	16,319	52.7	10,862	35.1	3,104	10.0	674	2.2	30,959
1986	17,598	45.5	16,546	42.8	3,574	9.2	990	2.6	38,708
1987	21,295	45.5	20,021	42.8	4,324	9.2	1198	2.6	46,838
1988	28,110	43.4	28,705	44.3	5,178	8.0	2788	4.3	64,781
1989	29,387	41.6	32,558	46.1	5,186	7.4	3459	4.9	70,590
1990e	30,271	43.1	31,337	44.6	4,799	6.8	3771	5.4	70,178

Industry Structure

An understanding of the semiconductor industry can only be gained by considering not only the semiconductor-device manufacturers and their products but the key parts of the industrial infrastructure that support semiconductor manufacturing. The evolution of semiconductor technology affects the device manufacturers and the materials and equipment companies alike. All must work together, and do. The discussion that follows thus deals first with the suppliers and then with various categories of semiconductor products and techniques needed to make them. All of these industrial groups serve a worldwide market.

The sizes of the companies in these businesses vary widely. Semiconductor material, gas, and chemical suppliers tend to be divisions of large chemical companies. Some specialty gas and chemical suppliers are small, providing specially purified and packaged products solely for semiconductor-manufacturing purposes. Equipment manufacturers are mostly small companies. There are a few medium-sized ones, most of which are mini-conglomerates formed by acquisition of small companies. The principal exceptions to the pattern are in Japan, where a small number of the large device manufacturers (Hitachi, Toshiba, and NEC, primarily) either make semiconductor-manufacturing equipment or have subsidiary companies that do.

Companies that make semiconductor devices vary widely in character. Some (e.g., IBM, Delco, Digital Equipment) make devices entirely for their own internal purposes. These are "captive" semiconductor-device manufacturers. Others (e.g., AT&T, Motorola, Texas Instruments, most Japanese firms, Siemens, Philips) make devices both for internal use and external sale. They are thus captive and merchant manufacturers at the same time. The remainder (e.g., Intel, National) are almost totally merchant device manufacturers. This group also includes dozens of small companies, mostly in the U.S. Most U.S. semiconductor-device manufacturers are the product of the entrepreneurial approach that has flourished here but is quite exceptional elsewhere.

The abilities of these varying sizes of firms to compete in international markets differ markedly. The large firms may do as they choose with some success. Small ones often have great difficulty, though not necessarily. A small company with a strong proprietary product position, good management, and a conservative balance sheet can do well worldwide.

KEY TECHNOLOGIES

Materials

Silicon

Two German firms (Wacker Chemitronic and Dynamit Nobel, now part of MEMC) produced over half of the world's silicon for most of the past 15 years but no longer do. Both companies have strong research activities that continue to provide steady advances. The Japanese have four large silicon producers each with good research operations. Collectively, they are now larger producers than the Germans. Other Japanese companies, notably Nippon Steel, have entered the silicon business in addition to the four referred to. Both the German suppliers and some of the Japanese ones have major facilities in the U.S. U.S.-owned producers account for less than 5 percent of the world's silicon production.

About 18,000 metric tons of raw, high-purity semiconductor silicon are produced each year. Losses during the process of making this material into polished wafers consumes about half. The remainder, mostly in the form of one-side-polished wafers about 0.7 millimeter thick and ranging in diameter from 100 to 200 millimeters, is used to make devices. Wafer use is reported in area terms, and was 1940 million square inches (110 million wafers) in 1989.

Compound Semiconductors

Only about 20 tons of GaAs, the principal compound-semiconductor material, are produced annually. Sumitomo Electric makes about 60 percent and is a leader in research and development for product improvement. The balance is made by numerous small firms that, in the U.S., tend to be owned by much larger companies. For GaAs substrate material, the U.S. is behind and holding its position in basic research and in development but not doing so well in production.

There are serious structural and purity problems in these materials. Source materials are often not available in the necessary purity. Crystals of compound materials are much more prone to breakage than is silicon; reduction of internal stresses is a necessity.

The U.S. is doing better in research and development on compound materials having complex layered structures that are used to make semiconductor lasers and some specialized transistors. The U.S. is roughly even with Japan in basic research, although here the work is done largely in universities and by military contractors (major exceptions: Bell Labs and IBM), while in Japan the research is in industrial and government laboratories. They are thus in a better position to bring the fruits of the research to market for civil applications.

Chemicals and Gases

Semiconductor production uses a wide range of chemicals and gases. Unfortunately, on the scale that most chemicals are produced, the quantities of any given chemical that are used are small. Thus the suppliers in the U.S. have tended not to put forth as much effort in producing chemicals of exceptional purity as the suppliers in Europe or Japan. Some chemicals and gases are difficult to ship internationally, so U.S. suppliers have been shielded from foreign competition to some extent. A recent change has been the purchase of several U.S. suppliers of gases and gas-handling equipment by European and Japanese companies, thus internationalizing the gas business by ownership.

Plastic Molding Materials

Plastic molding compounds are used in forming the body of finished semiconductor devices. The plastic enshrouds the silicon chip and can contaminate it if the molding compounds are not specially purified. As in the case of chemicals, U.S. suppliers have not pursued semiconductor applications with vigor. Most plastic molding materials now come from Japan.

Photomasks and Mask Substrates

Photomasks are made of glass or fused quartz plates (the substrate) on which the master patterns are created for making images on the silicon wafer that define the configuration of each layer in an integrated circuit. Photomask patterns must be perfect. The substrate must be quite flat. There can be no flaws on the surface or in the body of the substrate that would cause errors in the projected image on the wafer. By providing superior quality at lower prices, the Japanese have acquired almost the entire substrate business (Hoya) and a substantial part of the merchant photomask business (Dai-Nippon Printing). Roughly half of the photomasks used are made in-house by the device manufacturers, but on purchased substrates.

Manufacturing Issues

Manufacturing Equipment

The U.S. has led this $9.7 billion industry (1989) from the start. It is under a strong challenge from Japan, which has risen in world market share from 24 percent in 1985 to 42 percent in 1989, mostly at the expense of the U.S., whose market share fell from 61 to 48 percent in the same period. [*Editor's note:* As indicated above, after the writing of this chapter in May 1991, the U.S. achieved an increase to above 50 percent in 1992.] Japanese companies now make the best step-and-repeat patterning gear and the fastest final test equipment.[13] Because much of the technology for semiconductor manufacturing is embodied in equipment, loss of leadership in this field would be very serious indeed. The U.S. industry now has less than half of the world market, and this share is steadily declining.

Europe has several small firms with products that sell competitively in the U.S., but overall they are not a major competitor. The U.S. leads in ion implantation and in some kinds of film deposition (silicon epitaxy, for example). The U.S. also leads in many areas of measurement instrumentation. Most firms are small, having less than $100 million in annual sales. A protracted slump in semiconductor-device demand in 1985-86 hit equipment manufacturers doubly hard, because device factories were overbuilt in 1984. This excess capacity meant equipment sales were very low for years.

Exports to the PRC were the one bright spot for firms with equipment that could be licensed for shipment, but by no means all firms could take advantage of this opportunity.

This picture is changing. Equipment previously in use is becoming obsolete (a five year life span is a long one in the semiconductor industry). Automation, regardless of the cost, is required to manufacture sub-micrometer geometries at acceptable yields. People shed too much dirt and must be excluded from clean factories. This equipment is remarkably sophisticated. It is built in clean rooms and designed to be used and maintained in clean rooms without emitting contaminant fumes or particles. Most of it is at least microprocessor controlled. It is also expensive. Single items costing over a million dollars are common. It now costs about $500 million as a minimum to build and equip a semiconductor plant of economically viable scale. There are currently 1000 semiconductor plants in the free world, though only a few percent of these are of this modern expensive vintage.

Device Design

This is the only area in which the U.S. leads everyone and is holding position. The most successful computer-aided design systems are U.S.-made. This is a software-intensive field. The major device and process modeling codes, used everywhere, originated at U.S. universities. Japanese device manufacturers are becoming more proficient in modeling and design as they close the software gap.

Process Control

The Japanese manufacturers excel at process control as the key ingredient in quality control. The overall result is the production of high quality, reliable products at high yield, i.e., low cost. Japan’s effectiveness in the semiconductor business is directly traceable to this way of doing business. U.S. companies became lax about quality control during the 1970s, were abruptly awakened by the Japanese, and have been trying hard to re-establish parity. There has since been a great improvement in U.S. quality, but still not enough. European device manufacturers have had a similar history. Siemens has teamed up with Toshiba and is absorbing the Japanese process and quality control ethic with very positive effects on yield.

Process Automation

Automation of a semiconductor plant is expensive. Mitsubishi has two automated plants making memory chips and is building a third. Their yield has significantly improved (by 20 percent) because there are fewer operator errors and because there are fewer operators, the major source of contamination in semiconductor plants. The two existing plants, one built in 1983 and the other in 1985, can produce 18 million memory devices a month with 650 employees; the total capital cost was $455 million, a substantial price for that time. There are several other fully automated Japanese plants, but few elsewhere in the world.

Some industry standards exist for the necessary communication between machines and host computers. Much manufacturing equipment is equipped to meet these standards either as a standard feature or as an option. Other standards, such as those for the physical locations of facilities connections, work surface height, points for supply and removal of work in process, and means for interfacing with mobile robots, are only beginning to be developed. No standards exist for integrating the information gathered from all sources (equipment, measurement tools, material flow data, environmental conditions, safety alarms, and the like) in a semiconductor plant. A few proprietary data management systems are available on the market. Few semiconductor firms are willing to accept these as offered, but have developed their own; none of these proprietary systems can be considered standard.

Once the various data-communication systems are in place, one must deal with a flood of numbers that may be useful to machines and computers but which entirely overwhelms human beings. The conversion of data into *information* that can be understood by people and used to manage an automated factory is an additional very difficult step. Some type of machine-learning system will be needed to search the streams of data for patterns that indicate when exceptional situations exist and suggest appropriate actions to deal with them.

Devices

The larger application areas for semiconductor devices are given in Table 25.[14] The table shows North American consumption; for Japan and Europe the military segment is less and the consumer segment is larger, but the pattern is similar.

Advanced microcircuits

The world market for the principal classes of semiconductor devices is shown in Table 26.[15] The data in the table include intracompany consumption for merchant suppliers and exclude consumption by full captives. The competitive circumstances surrounding each of the several types of advanced integrated circuits are different. Four types, memories, microprocessors, custom logic circuits, and logic circuits made from compound-semiconductor materials will be mentioned here. In recent times, U.S. research and development for the most advanced circuits was done either as part of the Very-High-Speed-Integrated-Circuit (VHSIC) Program of the U.S. Department of Defense, or in the larger commercial houses, of which many, like IBM, are captive producers. The diffusion of technology from captive suppliers or VHSIC contractors into the merchant suppliers was slow, so the U.S. did not obtain maximum benefit from its expenditures. The establishment of SEMATECH is helping to improve the transfer of technology within the 14 member companies. Japanese research and development is done by the large suppliers for both captive and commercial purposes, by government laboratories, by Nippon Telegraph and Telephone (which has a strong technology transfer activity to its suppliers), and by government-sponsored cooperative research programs in the industrial laboratories. Coupling to the commercial market is thus quite effective.

Japan chose to enter the memory business aggressively in the early 1970s to make these highly standardized circuits in large quantities. The development chronology is shown in Table 27.[16] Its success was such that most U.S. suppliers left the field. Whether the U.S. can recapture this business is quite debatable. Some of the larger European firms are still making memories, but are not reported to be making money.

The U.S. has maintained a strong position in microprocessors, partly because the U.S. is still leading the world in software; the development of the software for a new microprocessor is at least as expensive as developing the chip. Once accepted in the market, microprocessors have a long life. Chip copyright laws have helped reduce the design copying that was a problem for a time. Until recently, Japanese firms have had to license the right to produce microprocessor designs from the

Table 25
1990 North American Semiconductor Consumption
($millions)

Segment	Value $M	CAGR* %
Data Processing		
Computer	4,749	13
Data storage	1,373	17
Data terminals	215	7
Input/output	455	8
Dedicated systems	175	3
Subtotal	6,968	13
Communications		
Customer premises	1,271	13
Public telecom	437	11
Radio	528	18
Broadcast & studio	156	14
Other	112	14
Subtotal	2,504	14
Industrial		
Security, energy management	176	12
Manufacturing systems	1,057	12
Instrumentation	502	14
Medical equipment	457	11
Commercial aviation	177	12
Other	340	19
Subtotal	2,708	13
Consumer		
Audio	42	11
Video	686	11
Personal electronics	51	10
Appliances	323	9
Other	35	7
Subtotal	1,138	10
Military	2,176	5
Transportation	1,319	11
TOTAL	16,812	12

*CAGR = compound annual growth rate 1987-91

original producer, but indigenous designs are beginning to have some success. The U.S. will continue to lead for a few years, but after that the future is uncertain.

Custom, or application-specific, integrated circuits take many forms. These different forms are shown in Table 28.[17] Custom circuits are used in growing numbers in applications for which very few of a given kind of device will ever be needed, such as short-production-run instruments, or in applications in which very large numbers of identical circuits are required, such as cars or television sets. This only appears to be a contradictory statement. In the first example, the cost and size of an instrument can be reduced by using an easily customized gate array. The simplest ones are derived from standard designs of individual logic elements (gates) with final metallization patterns to connect the gates into circuits for special purposes. This is an economical approach for small quantity requirements. Others examples are microprocessors that have been customized by providing them with fixed, pre-programmed memory that dedicates the microprocessor to a single application. This approach is used for applications requiring functions too costly when implemented by the standard cell approach. Still other types are designed from scratch for specific purposes. This is costly, so it is reserved for cases (the second example above) in which large numbers of identical parts will be needed. Until recently, the U.S. had a strong position in custom circuits. Now the Japanese are as competitive in custom circuits as the U.S. is.

Table 26
World Market for Various Types of Semiconductor Devices
($ millions)

Type	1980	1985	1990	1994	1999
Bipolar digital	2,374	3,672	4,089	4,577	4,185
Memory	572	589	497	421	239
Logic	1,802	3,083	3,592	4,156	3,946
MOS digital	4,715	10,122	34,474	75,630	174,079
Memory	2,230	3,821	17,078	38,300	91,985
Microprocessors	862	2,748	7,781	17,486	39,410
Logic	1,623	3,553	9,615	19,844	42,674
Analog	2,457	4,761	8,974	17,558	32,434
Subtotal ICs	9,546	18,555	47,537	97,765	210,688
Discrete	3,883	4,576	7,649	11,873	17,264
Optoelectronic	689	1,226	2,109	3,641	5,843
Total semiconductor	14,118	24,357	57,295	113,279	233,795

Logic devices using compound semiconductors are only beginning to appear in useful quantities. The growing availability of these devices is shown in Table 29.[18] The data in the table include intracompany consumption for merchant suppliers and exclude consumption by full captives. Devices based on compound semiconductors are faster than those based on silicon but are much more expensive. For this reason compound-semiconductor devices are attractive only for applications demanding their very high speed, such as supercomputers and microwave devices. The technology required to make them is useful in making advanced optoelectronic devices and will be used for optical computing chips in the future, if present expectations come true. The present level of

Table 27
Dynamic Random-Access Memory Chronology

Size	Critical Dimension (micrometers)	First Paper		Production Quantity 1M/year	Production Quantity 100M/year	
16K	3.6	1976	Intel	1977	1980	Intel
64K	2.5	1978	NTT	1980	1983	Fujitsu
256K	1.5	1980	NEC, Toshiba	1983	1986	Hitachi
1M	1.0	1984	NEC, NTT	1985	1988	Toshiba
4M	0.8	1986	TI	1988	1991	Hitachi
16M	0.6	1988	Toshiba, Hitachi & Matsushita	1991	1994	----
64M	0.4	1991	Toshiba, Matsushita Mitsubishi & Fujitsu	1994	1997	----
256M	0.3	1994	----	1997	2000	----
1024M	0.2	1997	----	2000	2003	----

Table 28
Estimated Worldwide Consumption of Application-Specific Integrated Circuits
($millions)

	1986	1988	1990	1992	CAGR(%)*
Gate Arrays	1,798	3,164	4,851	7,832	28
Programmable Logic	308	583	918	1,161	25
Total semicustom	2,106	3,747	5,769	8,993	27
Cell-Based ICs	674	1,263	2,091	3,611	32
Full-Custom ICs	2,160	2,395	2,148	2,088	(1)
Total ASIC	4,941	7,404	10,008	14,692	20

*CAGR = compound annual growth rate 1986-92.

integration is much lower than for advanced silicon chips, in part because of the problems with the materials mentioned earlier. Again, work done by contractors of the U.S. Department of Defense is not well coupled into commercial firms.

Table 29
Worldwide GaAs Semiconductor Merchant Consumption
($millions)

	1986	1988	1990	1992	CAGR(%)*
Analog/Linear ICs	107	177	294	532	31
Digital ICs	51	96	208	577	50
Subtotal	158	273	502	1,109	38
Optoelectronics					
Discretes**	1,558	2,021	2,216	2,915	11
ICs	28	67	162	253	44
Subtotal	1,586	2,088	2,378	3,168	12
Microwave discrete	203	244	288	345	9
Total	1,947	2,605	3,168	4,622	16

*CAGR = compound annual growth rate 1986-92.
**Discrete devices are mainly light-emitting diodes and lasers.

Optoelectronics

Optoelectronic devices are made from compound-semiconductor materials and include light-emitting diodes (LEDs), semiconductor lasers, and some photodetectors. The most demanding applications are in optical communication systems, but enormous numbers of LEDs are made for indicating lamp applications. All uses of compound materials, including the logic circuits mentioned above, account for about 5 percent of the dollar volume in semiconductor devices but use only 0.1 percent as much material in weight terms as is used for silicon devices. Compound-semiconductor materials are intrinsically expensive and difficult to process, and thus are only used when there is no alternative.

Much U.S. research is done in universities, and done well. In Japan, the research outside of company laboratories is done by government laboratories and consortia. On balance, the U.S. is ahead in research and the Japanese are ahead in implementation. This area of semiconductor technology is only just beginning to show its potential. Optical communication is growing rapidly. Faster computers are always sought. Optical computation is a tempting objective, though it will be realized only at some time in the future.

More information on the use of compound-semiconductor materials for optoelectronic applications may be found in Chapter 8, Lasers, beginning on page 183, in Chapter 9, Optical-Fiber Communications, beginning on page 217, and in Chapter 10, Optical-Fiber Sensors, beginning on page 303.

Microwave Integrated Circuits

These special-purpose circuits are used for satellite communication, military systems, and an increasing number of civil applications at frequencies from 1 to 100 gigahertz. Europe, Japan, and the U.S. are developing these for telecommunication and other applications. The U.S. and the Europeans also have military interests. In Europe and Japan the civil applications under development include direct TV broadcast from satellites to homes and vehicle detection and communication systems to control traffic, give routing information to vehicles, and warn drivers of impending collisions.

The distinguishing feature of these circuits is that their interconnections are specially designed for good high-frequency properties, which requires that the area taken by interconnections be large. The approach provides uniform, easily designed circuits having wide frequency response and relatively small size and weight. The high operating frequency leads to special problems. Unique design software is required. Chip testing is difficult; conventional methods cannot be used. At the same time data transfer rates are rising for conventional integrated circuits; 100-megahertz clock frequencies are used in the latest microprocessors and require microwave considerations in the design. Package and internal conductor design for “ordinary” integrated circuits will soon have to be done using microwave approaches. More information on microwave integrated circuits may be found in Chapter 7, Microwaves, beginning on page 147.

Power Devices

High-power semiconductors for power-frequency applications up to 400 hertz are used in power supplies, some very large (e.g., 1200 volts at 250,000 amperes). Due mainly to underinvestment in research and development and in modern facilities by U.S. companies and the existence of large electric railway systems (major customers) in Europe and Japan, the U.S. is behind and losing ground. This market is not likely to be recovered. Other devices, used in switching power supplies, high power audio, and radio frequency amplifiers, are still technically viable in the U.S.; here the U.S. is about even and will probably remain so for some time.

TRENDS IN THE TECHNOLOGY

The dominant trend in semiconductor technology is toward more complex integrated circuits, for the purpose of supplying more useful functions and reduced cost to the user. This has been the direction of the industry for 25 years. No change is in sight for at least the next decade. All other trends are consequences of this single one. Some of the consequences of the trends are illustrated in Table 30.[19] But these consequential trends involve major technological achievements on a wide front. For this reason, the 40-year-old semiconductor industry and its suppliers are still riding (creating?) a wave of "emerging technology."

Table 30
Integrated-Circuit Manufacturing Trends

DRAM size	4M	64M	256M	1024M	bits
Date of first functional device[1]	1987	1991	1994	1997	
1M units manufactured	1988	1994	1997	2001	
Minimum line width	0.8	0.4	0.3	0.2	µm
registration tolerance	±0.2	±0.1	±0.08	±0.05	µm
line width tolerance	±0.1	±0.05	±0.04	±0.02	µm
killer defect size	0.15	0.08	0.06	0.04	µm
Chip size	90	200	300	450	mm^2
Wafer diameter*	150 (1991)	200 (1997)	250 (2000)	300 (2005)	mm
Lithography field size	15×15	20×20	25×25	30×30	mm
Number of mask levels, average	18	24	28	32	
Single wafer processing	40	60	80	100	%
Liquid processes	35	25	15	0	%
Gaseous processes	65	75	85	100	%
Wafer factory capital cost	$250M	$500M	>$1B	>$1.5B	
Factory automation	30	60	90	100	%
Equipment cost (relative)	1×	1.5×	2×	2.5×	
Throughput, wafers/month	25K	20K	20K	15K	
Processed wafer direct cost	$250	$500	$800	$1250	each
Processed wafer cost[2]	$417	$917	$1633	$2917	each
Chips per production wafer[3]	184	146	152	146	units

[1] As reported at the International Solid State Circuits Conference.
[2] Includes depreciation of factory capital cost assuming 5-year straight-line depreciation.
[3] At a production rate of 100M units/year.

Attributes of materials have always defined the limits of semiconductor technology. The need for high purity has been mentioned before. Device manufacturers have continually pressed materials suppliers for improved purity levels, to an extent that at times seems to be an article of faith rather than supported by hard information. (Getting such information is expensive and time-consuming.) Part of the reason for this is the difficulty of measuring purity at the levels the device manufacturers desire. Another reason is that in the past there have been episodes of catastrophic device yield loss that have been traced to impure materials. The device manufacturers, once burned, are twice shy.

A significant change is beginning to occur in the way materials are specified. In the past, a given property of a material was required to have a value between an upper and a lower bound. For example, a metal rod might be specified to have a minimum diameter of 0.248 inch and a maximum diameter of 0.252 inch. Any value meeting these conditions was acceptable, and other values were not. The implicit but seldom stated assumptions involved in this kind of specification are first, that *any* value that was acceptable was as suitable for the application as any other, and second, that measurements could cleanly distinguish whether or not the material met the specification. For many semiconductor applications, neither one of these assumptions is true. There is often a preferred or "target" value for the property. Further, as the specified range of values of a property becomes squeezed ever smaller, the uncertainties in measurement often are about the same as the range itself. It is thus not possible to tell if the specification has been met or not.

Now, rather than guaranteeing that an attribute will always meet a given condition (requiring the materials supplier to be very conservative in specifying his product), a target value is set based on the user's requirements. Both supplier and user of the material understand that at times this value will not be met, so statistical limits are negotiated on the frequency and degree of deviations that may occur because of variations in the supplier's process and in both parties' measurement capabilities. There must be a closer relation between the two parties for this approach to work. In its optimal form, this relationship is collegial and mutually supportive in addressing problems, rather than the traditional adversarial one. This approach is new in U.S. practice, but derives directly from Japanese practice. The originator of the idea is W. Edwards Deming, but the implementation is due to Prof. Genichi Taguchi.

A similar approach to cooperative development and improvement of manufacturing equipment is also developing. There have been examples of this in the past, but now cooperation is becoming the norm. SEMATECH started to pursue this path vigorously in 1990. The funding of development programs is shared between SEMATECH and the equipment firm under the condition that the new tool will be first available exclusively to SEMATECH and its member companies. When these needs have been met, the supplier is free to sell to anyone. In addition, SEMATECH is trying valiantly to change the frequently adversarial relationship between suppliers and their customers into a cooperative one, to reflect more constructively the real interdependence between them. If this effort succeeds, it will be a major contribution to improving U.S. competitiveness in the semiconductor field.

The entire approach to manufacturing semiconductors is changing in directions that rely increasingly on more and better measurements and a stronger base of fundamental knowledge of each process.[20] This evolution from empirically based, manual operations to knowledge-based automated operation is illustrated in Table 31.[21]

Concerns about the cost of equipment, its rapid obsolescence rate, and the difficulties of contamination control have led to a concept of modular equipment ("cluster tools") that is being seriously pursued. The concept involves a number of tools that perform a sequence of different processes "clustered" about a core module. The core provides material handling in a clean vacuum environment and is the communications link between the tools surrounding it and the computers that manage materials flow and the processes executed by the entire cluster of equipment. The expectation is that the functions of the central core will be relatively invariant over time and will need replacement only rarely. If a process tool becomes outdated, it can be replaced *if there is a standard interface* between the core and the tool. The first industry standard for this interface has just been established.

Table 31
The Move Toward Intelligent Microelectronics Manufacturing

1970-1980's		Mid to Late 1990's
Empirically developed and "force-fit" into commercial equipment	*Processes*	Chemical models, developed in concert with the equipment
Batch processing generically developed for broad families of processes	*Equipment*	Fluid and heat models developed in concert with each process
1950's "Western Electric" rules. Control based on measurements on pilot wafers. Control to specification limits.	*Process Control*	Adaptive, robust, with artificial intelligence (AI). Based on thermodynamic-kinetic models. Control to target value.
All measurements on off-line equipment	*Sensors*	Smart microsensors, with computing at the sensor
Data gathering and lot manual planning at fixed times	*Factory Control*	Recipe downloading, AI tracking intensive, real time process control

The shrinking scale of features on the chip leads directly to tighter limits on processes. Film thicknesses and properties must be more closely controlled. The *absolute* uncertainty must decrease in proportion to thickness to maintain at least the same *percentage* of uncertainty. In addition, as the entire sequence of manufacturing processes acquires more steps, the variations in each step must be decreased to maintain control over the whole.

Today's processing equipment makes a few in-line measurements of its own performance, such as detecting the end point of an etching process[22]. Most such control is still rather rudimentary. Film deposition and removal processes involve chemical and physical phenomena that are generally not well understood in detail. Control tends to be of the "black box" type, in which it is assumed that if flow rates, temperatures, and similar input conditions are maintained then the product of the process will not vary. The process itself is the black box, about which little is known. This approach does not work well for many reasons. Not every condition that affects the process is measured and controlled. For example, keeping a uniform flow rate of an input gas stream does not help if the composition of the gas is not also well controlled. Many chemical reactions are strongly influenced by small changes in the level of trace impurities. Thus, the performance of equipment can vary with the time since it was last cleaned. Small changes in the impurity level of input gases can also have disproportionate effects.

Successful automation of the entire fabrication process requires detailed, accurate information concerning the execution of every process step. The amount of data involved is shown in Table 32.[23] For some process steps this can be approached, but others change the silicon wafer in ways that are not immediately measurable. Information about such steps comes later, if at all, and corrective information must be fed back up the line of processes. Most processes result in an irreversible change to the wafer, so if there is an error in processing only two alternatives exist. Either the wafer must be scrapped, or (in only some cases) a change in a later process can have at least a partially corrective effect. In the latter case, correction information must be sent ahead of the wafer. Obviously it is much preferable to have information on process execution available immediately, so measurements made on the process line are more useful than those that require laboratory services and thus do not deliver information right away.

Table 32
Information Quantity Trends

	LSI	VLSI	ULSI
Products (DRAM)	16K	256K-1M	4M
Throughput, wafers/month	10,000	30,000	50,000
Total process steps per lot	100	230-400	550
Number of types of equipment	40	100	120
Total equipment count	70	300	400
Number of process conditions	200	800	1500
Database records/lot for stable production	100	5,000	10,000
Database records/month	4×10^4	6×10^6	2×10^7

From time to time, a given process reaches some limit beyond which it cannot be pressed. Over the past decade, as device feature sizes have dropped, the end of optical imaging has repeatedly been foreseen. So far, ways have been found to meet the challenges and optical imaging is still the mainstay of the industry. It appears certain that optical techniques will serve at least to dimensions of 0.35 micrometer and perhaps to 0.25 micrometer[24]. Present state-of-the-art production devices are being made at 0.8 micrometer. Achieving good imagery at 0.25 micrometer requires the use of very short wavelength ultraviolet light.[25] If the wavelength is reduced beyond that point, the materials used for making lenses become opaque. This physical reason makes the argument for reaching the end of the optical road more convincing than it has been in the past. Before the end of the decade a new imaging method must be ready.

The choices are to use either electron beams or X-rays. Electron-beam equipment is now universally used to make the master patterns on photomasks. A fourth generation of equipment is under development, so the capabilities and limitations of the method are well known. Electron beams can write patterns at least as small as 0.1 micrometer with excellent quality. The process is too slow, however, because each point in every pattern must be written one at a time. Exposure of a 150-millimeter wafer takes 20 minutes or more, depending on the complexity of the pattern. Optical equipment now in use takes less than two minutes without regard to pattern complexity, because the image of the whole chip is exposed at one time.

X-ray processes are more like optical ones. Entire patterns are exposed at one time, through a mask. But the technology is far less well developed than electron beam methods. Suitable X-ray sources are largely still laboratory instruments, not production tools. Most suffer from low intensity, except for synchrotrons. These machines were originally developed for nuclear physics research. Such machines are large and expensive. Smaller models are being developed for patterning wafers, but they will still cost upwards of $25 million each without the wafer positioning equipment. X-rays have been used in the laboratory to make very attractive patterns, so in principle the method will work. But serious problems must be overcome in developing X-ray sources, masks, resists, and wafer positioning equipment. The latter must be much more precise than that now used for optical imaging purposes. For use with a synchrotron, the wafer must be held in a vertical plane rather than horizontally, which introduces mechanical complexities. How this matter will be resolved is not yet clear.

Until a few years ago, film removal was typically done by wet chemical etching. This method is simple and inexpensive, but has the problem that the etching process proceeds laterally in the film as fast as it does vertically, producing edges of holes in the film that slope at about 45 degrees. If the film is much thinner than the smallest feature in the pattern one is trying to make, this is not a serious problem. But for most films, the film is now thicker than the smallest feature and wet chemical etching cannot be used. The lateral etching, coming from both sides of a small feature, leaves only a vestige of the original film. (The films are somewhat thinner than before, but the features are much smaller.) A new method was developed using gaseous etching in a partial vacuum. The gases are activated by an electrical discharge, which makes them more chemically active and also gives the gas molecules an electric charge. By applying an electric field, the gas molecules are directed more or less vertically against the horizontal wafer surface, allowing etching to proceed faster in the vertical direction than in the lateral direction. This produces film edges that are nearly vertical, which is the desired result. This process is called plasma etching, because the electrically activated gas atmosphere is a “plasma”. Balancing these advantages is the fact that the bombardment of the silicon wafer by the high-velocity gas molecules can cause structural damage to the surface layers. This damage, if not avoided or later removed by a suitable heat treatment, adversely affects the electrical properties of the devices being made. Detection and measurement of this damage is a difficult and incompletely solved problem.

Plasma processes can be used to deposit films as well, but for a different reason. Films can be deposited by chemical reactions in a furnace, or they can be grown by exposing the silicon wafer to oxygen in a furnace, forming silicon dioxide. Other processes also require heating the wafer. Whenever the wafer is heated, atoms in the silicon or in the films move around by diffusion, attempting to redistribute themselves uniformly rather than staying where the device designer intended they should be. As devices become smaller, the total time at high temperature experienced by a wafer must be strictly limited. Alternative ways to accomplish the purpose must be found. In plasma processes the energy required to stimulate the desired chemical process is provided by the electrical activation of the gas, rather than by the heat of a furnace. Plasma deposition processes are widely used, but are well-known generators of particles that can fall on the wafer surface. Ways to avoid this problem are being sought. Other means of adding energy to cause chemical reactions to proceed include the use of ultraviolet light or radiation from lasers. These methods are not yet in use but are being studied.

Particulate contamination during the manufacturing process has always presented problems. Smaller feature sizes mean that smaller particles, if they should fall on a wafer, can make a chip defective. Particles of concern are those of a size greater than roughly a fifth of the smallest feature dimension. To make circuits having 0.25 micrometer feature sizes, particles larger than 0.06 micrometer must be excluded. Particles can fall out of the air, settle from liquids, or come from wearing moving parts of equipment. Some processes themselves generate particles by such means as nucleation and condensation from the gas phase in plasma and chemical vapor deposition equipment. Rough wafer handling can cause the edges to chip, generating silicon particles. People shed particles. Removal of all particles is not possible. Filters can remove them down to some critical size, but are less effective below that size.[26] The limit for filters in the room air supply is now about 0.1 micrometer.

Even detecting and counting very small particles is difficult. The usual method for counting particles in a gas or liquid stream uses an intense light, often a laser, shining across and through the stream. The light scattered from particles is detected. Larger particles scatter more light than smaller ones, so it is possible to determine the size distribution as well as to count them. But when the particles are of the same dimensions as the wavelength of the light (about 0.5 micrometer), individual particles do not scatter light in this way. Finding particles on a surface, such as that of a wafer, uses similar approaches and suffers from the same limitations.

Particles are not the only kind of contamination to be avoided. Dissolved impurities in liquids can leave residues on surfaces. Process materials themselves, if not properly removed, are contaminants in later steps. Photoresist is a good example.

New difficulties exist within the integrated circuits themselves as the dimensions become smaller. The electrical connections on the surface of the chip are very thin and narrow. While they carry only small electrical currents, the small cross-sectional area of these conductors is such that the density of the current is extremely high: approaching a million amperes per square centimeter, more than 1000 times that in ordinary house wiring. The stream of electrons, which is the electric current, can actually move the atoms of the conductor. Over time, this rearrangement of material (electromigration) may cause voids to develop in the conductor, further reducing the current-carrying capacity. Eventually the conductor melts and the circuit fails. The insulating layers of the chip, which separate electrically conducting regions and avoid short circuits, are highly stressed. These can fail with the passage of time. As the volume of individual parts of the circuit becomes smaller, previously insignificant effects of the boundaries of the part become more noticeable because the ratio of the surface area to the volume increases. These issues require research into causes of failures and ways to prevent them, and into the details of operation of extremely small transistors so that integrated-circuit designers can cope successfully with these difficulties.

Additional trends also exist:

Increasing wafer size, increasing the need for uniform process control over larger areas.

Automated wafer handling, because of greater wafer weight and the need to reduce particulate contamination and wafer damage.

Single-wafer processing, to reduce the size and control the operating cost of equipment and increase the uniformity of processing. (Most present-day equipment processes many wafers at once.)

Statistical process control is becoming more prevalent.

Greater application of computers at all levels, for data communication and processing, adaptive process control, faster information feedback (thus quicker process correction). The data come primarily from metrology tools.

More in-process measurement. For example, Japanese experience[27] shows that adding five percent to the cost of equipment for in-line sensors and data links results in:
42 percent reduction in cycle time
50 percent increase in output
32 percent increase in equipment uptime
25 percent reduction in direct labor hours.

Optimal use of these measurements also requires that process monitoring data, lot history data, and device functional test data be readily available for diagnostic work to guide corrective actions. Machine learning (or "artificial intelligence") tools can help greatly in identifying data patterns that characterize problems and also indicate paths to solve them.

GOALS OF U.S. INDUSTRY FOR COMPETITIVENESS

The primary goal of U.S. manufacturers of materials, equipment, and semiconductor devices is to improve their competitive position while still making money. Their competitive position has been steadily eroding, as reflected in the share of the world market now held by U.S. companies in these industry segments. Many causes for this loss have been cited. In materials and devices, there was a period of U.S. insensitivity to the achievements of the Japanese in improving quality (which directly improves manufacturing cost). This was brought forcefully to national attention in 1980 by Hewlett-Packard's description[28] of its experience with various suppliers of memory integrated circuits. The defect rates at incoming inspection were dramatically better for products from Japanese suppliers than from U.S. suppliers. The latter have now improved their quality significantly. Some are fully competitive in this respect now.

U.S. silicon suppliers, in common with their competitors, have suffered from persistently low market prices, in part caused by industry overcapacity. The U.S. suppliers cut their research-and-development efforts to reduce costs, but the German and Japanese firms did not. They thus learned from their development work how to reduce costs by improving efficiency. Ultimately the U.S. operations were sold to the foreign firms, and now industry rumors state that these U.S. operations with foreign owners are profitable. Companies must be willing to take a long-term approach to compete effectively with either European or Japanese firms, but that is not a popular approach with investors in this country.

The equipment manufacturers, as a group, have kept doing their research and development to the extent that small firms can, and have lost market share only with a fight. The competitive advantage in this field comes to equipment suppliers that are part of, or subsidiaries of, large companies. Those are mostly Japanese. In addition, there is a strong tendency for the Japanese device manufacturers to buy from indigenous suppliers. Customer support is better from local suppliers anywhere. U.S. firms with strong support operations in Japan have generally done well. A few U.S. firms manufacture in Japan for the Asian market and find this presence to be a considerable advantage. Given equal levels of technology in equipment, the firm supplying the better customer service has the advantage.

These three examples illustrate that what happens in the marketplace hinges on improved service to the customer. The technological factors that allow a company to perform better in the market need closer examination. Improved quality and reliability, better-functioning products, more repeatable operation, and lower costs come from purer and more uniform materials, closer control of processes, process and equipment improvements, better design techniques, and more effective ways of handling information.

MEASUREMENTS TO MEET THESE GOALS

New or better measurements are implicit in all of these approaches to improvement in service and competitiveness. The very significant role of measurements is evident in the industry's estimate that 30 percent of the cost of complex integrated circuits is in testing.[29]

For materials, there is an obvious need for providing purer and more uniform (better controlled) materials and for closer control of processes. One cannot control what one cannot measure.

For improving processes, equipment, designs, and information handling, the role of measurements is more subtle but equally important. Many processes are imperfectly understood. Their execution, and the design and operation of the equipment that performs the processes, are often based on empirical information, lacking anything better. Gaining an improved understanding of processes involves measuring what is occurring inside the process, learning about the physics and chemistry that govern the process, and obtaining quantitative, detailed knowledge of why the process works, what limits its performance, what to change to make improvements, and what to monitor to be sure the process is running right. These measurements and their interpretation are often difficult, but the research only needs to be done once. The results immediately translate into better control, more efficient processes, and the engineering data needed to design improved equipment. Numerous workshops and symposia of SEMATECH and the Semiconductor Research Corporation (SRC)[30] that were sponsored by Semiconductor Equipment and Materials International (SEMI),[31] American Society for Testing and Materials (ASTM),[32] NIST, and others have identified the need for this kind of information.

Quality improvement results from proper application of the information gained by improved measurements. First, all sources of variation (in materials, processes, tools, measurements) are eliminated or reduced as far as possible. Then design alternatives are chosen to make the processes and products as insensitive to residual variations as possible. Both of these steps involve carefully planned, statistically analyzed experiments. The end result is then the starting point for repeating the whole procedure. One never stops doing this, because new knowledge, tools, and experience become available to refine the manufacturing process still further.

The design of integrated circuits is based on detailed knowledge of the operation of the individual transistors that comprise the circuit. Whenever the scale of dimensions of the circuit is changed, the behavior of the transistors changes, often in ways that were not foreseen. Computer models of devices are widely used to predict the performance of these circuit elements, based on the physics of transistors and solid state materials. Certain physical assumptions are built into all computer models, and as devices are scaled down these assumptions may no longer be correct. This leads to errors in the predictions of the model. Corrections of these problems involves searching out the causes, improving those parts of the computer model, and verifying the changes by measurement of transistors specially made to provide the data to show that the problem has been solved. Again, such work only needs to be done correctly once to be of benefit to the entire industry.

The example given above requires long-term research to deal with a problem. Many other measurement developments need substantial effort over a long time to create a generic solution. Few companies are willing to address such problems except to the limited extent necessary to deal with them as near-term obstacles. If the problem arises again during the development of the next generation of devices, it must be dealt with a second time. Device companies are not usually in the business of creating complete solutions to measurement problems. Such work is a diversion from product development. Similarly, universities do not often solve measurement problems well. The time required is longer than the typical tenure of a graduate student. Only a few laboratories, such as NIST and some of its counterparts in other countries, have a mission to do this kind of work.

Information handling is central to management of semiconductor manufacturing at every level. Engineers need detailed technical data to detect or anticipate problems. Equipment gathers data on its operations and its condition. Instruments evaluate the results of manufacturing processes. All of these data come from measurements of some kind. One prefers that the measurements be done correctly and the data be significant. But the best measurements are of no value unless the results are communicated to those who need them. This does not occur as simply as it might, and sometimes not at all. Many kinds of computers are involved, speaking different languages. Some are on differing types of networks that cannot work together. These communication problems need solution.

Even so, a torrent of data assaults the process engineer, most of it reporting that everything is under control and thus needing no action. Computer tools to sift out the data patterns that indicate trouble are badly needed, and are just beginning to emerge. The more effective tools use adaptive computer approaches called artificial intelligence, or machine learning. They can recognize data patterns that typify a given problem and alert the engineer to the possibility of a specific difficulty. They can also “learn” from experienced people, and incorporate that experience into their analyses.

EVIDENCE OF NEED FOR IMPROVED MEASUREMENTS

ASTM Committee F-1 on Electronics sponsors workshops on measurement issues in semiconductor manufacturing that provide reports describing the need for new measurement standards and the metrology on which they would be based. ASTM F-1 has a planning subcommittee on standard reference materials (SRMs), established in 1989. At its first meeting, in addition to developing a list of proposals for SRMs from NIST, the subcommittee decided to start work on consensus-calibrated materials. It has also asked NIST to be the repository for materials used in interlaboratory experiments for establishing the accuracy and bias of new test methods, these materials having a lasting value for several measurement validation purposes. Both of these ASTM activities are valuable sources of industry advice on current needs for measurements, especially relating to materials and in-process topics. SEMI standards committees have discussed needs for measurement work on moisture in gases, contaminating particles, feature size and layer-to-layer registration of patterns, gas and liquid flow, trace impurity levels in chemicals, and thermal properties of semiconductor devices and their packages.

Several technical presentations at a symposium on measurements for the semiconductor industry, sponsored by ASTM, SEMI, and NIST early in 1990, contained lists of quantities for which better measurements are needed. SEMATECH held workshops during 1989 on lithography (pattern production) and in 1990 on film deposition processes that provided much useful guidance. The SRC Technical Advisory Board meets for three days annually to assess the long-range trends in the industry that will define its research agenda. The SRC also held two topical research conferences on temperature measurements and on metrology for semiconductor manufacturing in February 1990. The reports of these conferences[33] contain very good current consensus summaries of measurements needed through the end of the century, to the extent they can be identified in detail.

Technical discussions with SEMATECH staff support the need for

> **improved measurements** of feature size and layer-to-layer registration of patterns, thickness and interface characteristics of thin transparent films (including multilayer films), and intensity and spectral distribution of short wavelength ultraviolet light used for exposing photoresist;
> **calibration services** for ultraviolet photometers and for flow-measuring devices;
> **evaluation of techniques** for measuring the concentration and spatial distribution of chemical species in plasma discharges and determination of their chemical reactivity;
> **reference materials** to calibrate instruments that measure the depth distribution of dopants in silicon;
> **improved understanding** of the chemical and physical processes involved in depositing films on surfaces;
> **improvements in test devices** used to assess the execution of many fabrication processes;
> **development of computer aids** to detect data patterns indicating out-of-control processing conditions;
> **standard data systems architectures** that allow communication between disparate computer systems and networks effectively and transparently; and
> **comparisons** among several commonly used device modeling codes to determine whether or not the results from the various codes agree.

The Semiconductor Industry Association (SIA) has identified metrology at one of the "critical technology competencies" in which the U.S. has a shortfall that is impeding progress toward industry goals. The Association names specific measurement needs that are especially important, such as dimensional measurements to support lithography and contamination measurements. The SIA explicitly calls on NIST for assistance as the "only place in the U.S. where the broad range of measurements needed for semiconductor processing are routinely and systematically developed."[34]

The development cycle for a new generation of semiconductor devices takes more than a decade, starting with basic research and extending through industrial technology development, new manufacturing-tool development, and new device-manufacturing development to the time when devices are qualified by customers for their use. At any time, about four such cycles are in process, because new device generations appear in the marketplace about every three to four years. Metrological research-and-development needs to take place throughout these cycles.

MEASUREMENT NEEDS

Throughout its history, the semiconductor industry has required measurements that were beyond the current state of the art. These needs have been one of the major driving forces on company and university researchers, instrument manufacturers, NIST, and industry standards organizations like ASTM and SEMI to develop new scientific information, tools, and measurement methods. In spite of the impressive number of techniques that now exist, better ones are still needed. There is no apparent end to this continuing demand.

There are, however, three major differences in the needs of today and through the 1990s as compared with those of perhaps a decade ago. One is the very much wider range of technologies now in use by the industry and the consequent broader spectrum of scientific and technical skills needed to address the measurement needs. The second is the greatly increasing difficulty in pushing forward the frontiers of measurement. Detection sensitivities of one part per trillion (a million million) are needed in some analytical measurements. As the devices get smaller, the volume of material to be probed is also smaller. The spatial resolution of many measurements should be, but currently is not, smaller than the minimum feature size of patterns on wafers. These are formidable challenges. Without the needed measurements, the material and equipment suppliers and the device manufacturers often are left groping in the dark in their efforts to improve quality and yield of their products. Third is the greater need for real-time and in-situ measurements.

Measurement techniques are often different for those that are performed in *real time, in-situ, in-line,* or *off-line*. Real time metrology involves continuous measurement within the process tool and feedback for dynamic control of the process. In-situ measurements are also performed inside the process tool but after the process step is completed. In-line measurements are made on external equipment after the wafer is removed from the process tool; the wafer is returned to the process flow. Off-line measurements are usually more time-consuming and may be destructive; the wafer is generally not returned to the process flow. The first three types of measurement must be non-destructive. Reproducibility[35] is essential for all of these kinds of measurements. Accuracy[36] is very desirable for process control and indispensable if technology is to be transferred.

Materials

High-priority new or improved measurements needed for silicon are listed in Table 33.[37] Lenses used to project the patterns on wafers have a very shallow depth of field (the region of sharp focus), which is approaching one micrometer. This means that the area of the wafer to be printed, now about 0.4 inch square but growing over time, must be at least that flat when mounted on a vacuum chuck. Distortion of the chuck itself as a result of the internal reduced pressure and any external forces must be taken into account.

Table 33
Selected Measurement Needs for Silicon

Property	Range	Uncertainty
wafer diameter	up to 200 mm	<0.05 mm
wafer flatness	20 µm	<1 µm
wafer thickness	up to 725 µm	<1 µm
site flatness	<1 µm	<0.1 µm
site thickness variation	0.5 µm	<0.1 µm
site size	20×20 mm	
resistivity	0.001-200 ohm-cm	<0.5%
resistivity radial uniformity	0-5%	<0.5%
oxygen content	0-40 ppm	<1%
carbon content	<200 ppb	<5%
epitaxial layer thickness	1-50 µm	(0.2 µm +5%)
sheet resistivity	varies	<10%
surface impurity level	$<1\cdot10^{11}$ cm^{-2}	$<1\cdot10^{10}$ cm^{-2}
surface particle size, max.	0.1 µm	
particles per wafer, max.	5	

The electrical resistivity of the wafer is determined during the crystal growing process by the material producer. It can be varied over an extremely wide range by adjusting the dopant concentration, but its value is tightly specified by the device manufacturer. Typical specified ranges from maximum to minimum value of the wafer center-point resistivity at present are about 5 percent of the nominal value. Currently available Standard Reference Materials (SRMs) from NIST have an uncertainty of 1.58 percent. These are typically used to calibrate other working standards, which thus are somewhat less precise. These secondary standards are used in turn to calibrate test instruments (having added uncertainties of their own) used to determine the resistivity independently by the buyer and the seller. One can readily see that if the errors introduced by each party are in the wrong directions, the buyer may believe that the material does not meet his specification while the seller believes it does. NIST is working to reduce the uncertainty of the SRMs, but a better way must be found to deliver a calibrated measurement to the points of use without additional uncertainties creeping into the process.

Oxygen is present in most silicon up to about 40 parts per million (ppm). Its presence at a controlled level of 20 to 30 ppm is useful for reducing wafer distortion during processing and for removing unwanted impurities from the active regions of the wafer. Its concentration is measured by the absorption of infrared light as measured by a spectrophotometer. No NIST SRMs exist for this measurement, though they are in preparation. The commercial situation is similar to that of resistivity; the maximum acceptable uncertainty in oxygen content is very close to the uncertainty of the measurement, given that different spectrophotometers vary somewhat in their calibration. Without suitable SRMs, this variation cannot be removed from the measurement.

Measurements needed for compound semiconductors are listed in Table 34. Wafers of compound semiconductors must be as flat as silicon wafers. In addition, numerous different, specialized measurements are required. Stoichiometry refers to the atomic ratios of the constituent elements, which can be varied at will in compound materials but must be controlled to produce materials with predictable characteristics. Silicon is easy to produce with few, if any, structural defects. This is almost impossible to achieve with compounds, though steady progress is being made. Impurity control is much more difficult in compounds as well, largely because the starting materials are not available in such pure form as those for silicon. Many devices made from compound semiconductors require wafers having extremely complex layered structures. These are made, starting with a uniform substrate, by growing layers having differing composition, doping, or thicknesses. The growth process is capable of changing composition on an atomic layer by layer basis if necessary. Consequently, the control of these three variables is crucial.

Table 34
Selected Measurement Needs for Compound Semiconductors

Property	Range	Uncertainty
stoichiometry	0-100%	0.5%
defects	wide	2%
impurities	wide	5%
layers		
thickness	atomic scale, up to few µm	1%
composition	0-100%	1%
electrical properties	wide	5%
uniformity		2%

Gases used in the semiconductor-manufacturing process need the measurements listed in Table 35. There are two large groupings of gases, those used in large volumes (nitrogen, oxygen, argon, hydrogen) and a much greater variety of specialized gases used in much smaller quantities. All share these basic measurement needs, being required to be free of particles and moisture and to meet varying specifications for composition and purity. Currently available measurements for all but composition fall short of tomorrow's needs by factors of ten to a thousand. For many special gas mixtures, composition (assay) measurements are not sufficiently accurate as well. Nitrogen and oxygen may economically be produced at the user's plant if the volume warrants it. The process involves the liquefaction and subsequent distillation of air in a facility operated by the gas supplier.

Table 35
Measurement Needs for Gases

Property	Range	Uncertainty
particles	down to killer defect* size	<5%
moisture	<1 part per billion	<10%
composition	wide	<1%
impurities	<1 part per billion	<10%

*See Table 30.

The list in Table 36 describes the measurements needed for liquid chemicals and water. Large quantities of high-purity water are used in semiconductor manufacturing for rinsing and cleaning. It must meet the same requirements for particles and impurity content as any other chemical. Special requirements apply to biological contaminants (bacteria and organic materials) in addition. Solvents, acids, and photoresists each have unique analytical and physical properties that must be measured. The majority of these measurements are beyond the state of the art today.

Table 36
Measurement Needs for Liquids

Particles	smaller than killer defect* size number, size distribution needed new techniques needed below 0.2 µm
Assay	principal constituents, to 1% of amount
Impurities	unwanted materials, below ppb levels example for water: iron<0.01 ppb silica<1 ppb

*See Table 30.

Solid materials, such as metal deposition sources and packaging materials, also require measurements. Metals are used to form the conducting patterns that interconnect the individual devices that comprise the integrated circuit. These are deposited either by physical processes such as evaporation (mostly for pure metals, not alloys), by sputtering, or by chemical vapor deposition. In the latter case, the metal is incorporated in some chemical compound that is decomposed to produce the metal film. These films, produced by either physical or chemical means, are subsequently patterned using a lithography process. Regardless of the method of film formation, the important variables to control are the metal purity and composition (if it is an alloy), film thickness and electrical conduction properties, and the grain size of the crystallites in the film. All of these must be measured. The properties of the source materials, such as solid metals for physical deposition and gases or liquids for chemical deposition, have direct effects on the properties of the deposited films. Measures for the essential characteristics of the starting materials and of the deposition processes themselves all need to be developed.

Metals are also used for bonding wires to connect the integrated circuit to the parts of the package that form the external connections, and for parts of the package assembly. Bonding wire is tailored to specific processes and equipment, and must have closely controlled composition, diameter, temper, surface finish, and cleanliness. These wires typically range in diameter from 25 to 75 micrometers (0.001 to 0.003 inch). For plastic packages, a metal stamping called a lead frame is the other metallic part. This rather intricate stamping includes a pad on which the integrated-circuit chip is mounted, slender fingers to which the bonding wires from the chip are attached, and an external frame (hence the name “lead frame”) that holds everything together until the package is molded. Then the frame is trimmed away. Lead frames have less demanding requirements for purity, though they must be clean. The important attributes of lead frames are their composition, dimensions, flatness, and surface finish. Analytical measurements for bonding wires are critical, but some attributes are not well understood. It is conjectured that some minor constituents of bonding wire reside in the grain boundaries of the metal, where they affect the bonding properties. All that really is known is that different batches of wire having apparently the same composition do not always have the same handling and bonding properties.

Plastic molding compounds are specialized for semiconductor use, having especially low levels of ionic constituents. If present, such materials contribute to corrosion of the lead frame and the metal patterns on the chip, particularly in moist environments. In addition, the molding properties, such as the viscosity of the material while being molded, must be closely controlled. If the compound is too viscous, the positions of the fine bonding wires can be disturbed and, in extreme cases, the wires can be broken. Finally, the molding compound must have good thermal conduction properties. Plastics at best are poor thermal conductors as compared with metals, but because the plastic part of a package has a far larger surface area than the metal parts and is in more intimate contact with the chip, a substantial part of the heat generated in the chip passes through the plastic to the external environment.

All of these materials must also have extremely low levels of radioactive impurities. Traces of radioactive elements pervade everything, but at such low levels that they are usually of little concern. But the regions of chips that store information are extremely small, and the number of electrons representing the information is also minuscule. Alpha particles, the products of certain radioactive decay processes, create large numbers of electrical charge carriers if they travel into the silicon. These cause loss of bits of stored information by filling the storage locations with unexpected and unwanted electrical charges. The damage is not permanent, but the changes are real and often have serious consequences. Radioactive elements do not ordinarily enter the chip-making processes because of the chemistry and physics of the processes, but if present in the surrounding materials they can cause these transient upsets of stored information.

Processes

Most processes require a much more detailed list of measurements than can be given here. An example of what must be known about a plasma etching process to provide improved process control is given in Table 37.[38] This information would be provided to a computer that models this process and predicts the listed output parameters. At the present, too little is known about plasma processes to implement such a system. The measurement tools would, in most cases, be specially designed to fit into the process chamber and not be damaged by the effects of the process.

Table 37
Plasma Process Information Requirements

Input Variables

Material Attributes
wafer resistivity
wafer reflectivity
wafer weight
gas ionization
etch parameters (at right)

Environmental Variables
temperature

Process Variables
plasma parameters (at right)
machine parameters (at right)
wafer temperature value
wafer temperature uniformity
species concentration in the plasma
pressure
flow

Etch Parameters
anisotropy
selectivity
etch rate
etch rate uniformity

Plasma Parameters
electron temperature
electron density
ion temperature
neutral density
sheath potential

Machine Parameters
power
pressure/vacuum
wafer temperature
wafer potential
flow rates
residence time
electrode spacing
DC self bias
film thickness
end point detection
electrode etch

Output Variables
film thickness
critical dimension
etch depth
wafer weight

Instrumentation[1]
vacuum diagnostic system
RF monitor
optical emission spectroscope
Langmuir probe
bulk ion temperature and residual gas analyzer
forward looking IR
laser-induced fluorescence
voltmeter, ammeter
pressure gauges
flowmeters
interferometer
reflectometer
weighing instrument

[1] Listed instruments do not relate to the variables immediately to their left.

Similar process information requirements, but substantially different in their details, exist for many other deposition and removal processes, such as epitaxial layer growth, chemical vapor deposition (which comes in several distinct varieties), plasma deposition, reactive ion beam etching, etc. Obtaining the basic chemical reaction and physical properties data is difficult, requiring the use of the most advanced scientific measurement and diagnostic tools. Most of these processes have been developed about as far as possible on an empirical basis. Further progress depends on substantially improved knowledge of these process details and the development of computer models that can translate the measurements made in-situ into process control actions that improve productivity.

Lithography (the process of making patterns), already a demanding process, will require pressing optical tools to their limits by the end of the decade, as shown in Table 38. The measurements available now are unsatisfactory and inadequate to meet future needs. Measurements of feature sizes on wafers cannot be made to the desired accuracy. The industry is making do with instruments that are moderately repeatable but of uncertain accuracy. Reaching the goals in the table will be extremely difficult. To appreciate the meaning of the numbers, one has to recognize that the spacing between silicon atoms is 0.0004 micrometer, or 0.4 nanometers. Thus the machines that produce the patterns must be able to position them with an accuracy of 125 atomic spacings in both x and y dimensions within a tenth of a second, level the wafer to fit within the focal plane, focus to well within a micrometer of the right place, expose the wafer, and repeat this cycle every second for 16 hours a day in a production environment. The demands on the automatic metrology systems that must be built into these machines to control their operation reliably are truly staggering.

Table 38
Measurement Needs for Lithography in Year 2000

Property	Target Value	Uncertainty
Patterns		
smallest dimension	0.15 μm	±0.03 μm
overlay error	zero	±0.05 μm
killer defect size	0.03 μm	
Radiometry		
light intensity	varies	±5%
wavelength (for lasers)	248 nm	0.0001 nm

Instruments used to measure the light intensity and exposure time of today's lithography exposure tools are inaccurate. This has been shown by interlaboratory experiments with a major semiconductor manufacturer and with SEMATECH. An ultraviolet radiometry calibration service is being developed by NIST for the light wavelengths now in use. Mercury arc light sources are universally used now, but future exposure tools will need shorter wavelength, brighter sources than mercury arcs. Various types of lasers are under development for this purpose. One promising candidate emits light having a nominal wavelength of 248 nanometers (1 nanometer is 10^{-9} meter). Basic optical theory shows that the focal length of lenses varies in such a way with the wavelength of the light that some means of wavelength stabilization will be necessary if the lens is to have a known focal length. This means that the wavelength must be measured with considerable accuracy. As new exposure tools for finer geometries are developed, these measurements will need extension to cover shorter wavelengths as well.

Film deposition and removal processes will need many improved measurements, as Table 39 shows. Existing equipment measures some of these variables, such as deposition rate, but indirectly. To reduce process variations, it is much preferable to measure as many of these properties as possible in real time. Film composition, thickness uniformity, and step coverage (maintenance of uniform film thickness and integrity at places where the underlying surface changes its height) are all measured outside the process machine now. Knowledge of surface chemistry and physics will allow key measurable properties of the process to be identified, which will allow more of these measurements ultimately to be made in real time. In film removal, it is necessary to take away parts of an existing

Table 39
Measurement Needs for Film Deposition and Removal Processes

Film Formation	Film Removal	Film Thickness
deposition rate	removal rate	single
composition	uniformity	multiple
surface reactions	differential etch rates	interlayers
thickness control	end-point detection	properties
uniformity		
step coverage		

film completely and cleanly without eroding the underlying surface. Nearly vertical edges should remain to define the islands of remaining film that were not removed. The process should proceed at the same rate across the entire wafer. To avoid etching the underlying layer, the chemistry of the process is chosen such as to react readily with the material to be etched and not with the underlying film, as far as that is possible. That is the differential etch rate.

More sophisticated off-line measurements are used to verify that the process produces the desired films, to test their electrical and optical properties, to examine the properties of the interfaces between layers, and to make measurements on more complex structures. Such information is used to establish accurate process performance data for use in process modeling and improvement programs.

The connection between the silicon and the metal patterns that interconnect individual circuit elements is crucial for proper integrated-circuit performance. These connections are typically made through holes in insulating overlayers. The holes are of about the same diameter as the smallest pattern dimension and may be several times that deep. Special processes are often used to form metal plugs in these holes that are then interconnected in a later step. A chip will contain several million of these holes that must *all* be successfully plugged with a metal that makes good electrical contact with the silicon. This process will become more demanding as dimensions continue to shrink.

General Purpose Techniques

Some measurements must be made in many places throughout a semiconductor plant. Particle detection, counting, and sizing is one such measurement. An example of the desired characteristics of one type of particle detection system is given in Table 40.[39] This one would be optimized for measurements inside equipment in a vacuum environment. Different types of particle detection instruments are used for measurements in gases or liquids, or on wafer surfaces. Each of these applications constrains the design of the equipment differently, including imposing different limits on the minimum size particles that can be detected. The size limit today is 0.2 micrometer if the particle dimension must be known. Smaller particles can be detected but without knowledge of their size.

Another process measurement technique involves test structures that assess how well various process steps were executed, as shown in Table 41. Test structures are tiny device-like measuring tools that are made by the same processes that make integrated circuits. They are inserted in small spaces on the wafer, such as between the integrated-circuit chips in the track to be taken by the diamond saw when the chips are separated. The test structures are designed to measure such parameters as the

Table 40
In-Situ Particle Detection Objectives

Large area detection	>100 cm^2 desirable, >10 cm^2 a good start
Detection limit	<0.05 µm desirable, certainly <0.1 µm
Sizing capability	nice feature, but not essential
Real-time detection	essential for real-time control
Inside the chamber	needed to locate particle sources
In pumping line	acceptable, but not as good as inside the chamber
On-wafer monitor	ideal, probably too expensive and <0.1 µm not possible
Optical detection	nice, but <0.1 µm not possible
Low cost	$5K - $15K for real-time process control

electrical properties of films, electrical contact between various layers, width of conducting lines, the properties of individual transistors having the same geometry as those in the integrated circuit, and many other process-dependent characteristics that can be evaluated by electrical means. Since the test structures are processed in the same way as the rest of the wafer, these measurements are a reliable way of characterizing individual process steps. They do not, however, deliver the information until relatively late in the process after the wafer has been given its metal contacts, because it is not possible to make electrical connections to the test structures before then.

Test structures can also be used to evaluate the performance of heavily stressed parts of the circuit to determine if they are likely to fail in service or not. The physical limits of performance of metals for conductors on the chip and of insulating layers such as gate oxides are being approached. As devices become smaller, other parts of the circuit will no doubt be exposed to such extreme stresses. Test structures for characterizing the electromigration failure mechanism of metals have been developed along with standard tests using them to forecast the reliability of the chip. New test structures can be devised for other such measurements as they become needed. The objective is to assure that the integrated circuit was made right from the beginning and will not fail. The approach to reliability assurance has been to test large numbers of parts under extreme conditions designed to accelerate failure mechanisms. This is time consuming and expensive, and not entirely satisfactory for technical reasons as well. Predictive testing has the potential to change this situation greatly for the better.

Many process variables are common to a large number of kinds of manufacturing tools. The temperature of the wafer during processing must be accurately known because of its influence on the rate at which chemical and physical events occur. There are serious limitations in measuring temperature accurately, because the temperature sensing device cannot be allowed to contact the wafer for fear of contamination or physical damage. Errors of 200°C at temperatures of 800°C are common in rapid thermal processing equipment that rapidly raises the temperature of a single wafer to high temperatures, holds a temperature briefly, and then cools the wafer quickly. The purpose is to

Table 41
Measurement Needs for General Process Evaluation

Process Evaluation	Process Variables	Process Information
test structures	temperature	data transmission and management
in-situ measurement tools	flow	data screening methods
reliability prediction	pressure	
	residual gas analysis	

accomplish a thermal treatment with a minimum time at high temperature, to limit undesired diffusion effects. The high rate of change of temperature greatly complicates the measurement problem. Even in equipment that operates at constant temperature the errors in temperature measurement can be intolerably large.

Flow rates of gases and liquids also offer difficulties. Calibrations are done in terms of standard fluids such as air and water. When the application is for metering other fluids, particularly gases, errors are common. Mass flow controllers use the thermal properties of the gas to measure flow, and these vary significantly. Calibrations using the actual gas to be measured are not practical because there are so many different kinds, some of which are either corrosive, toxic, or both. Errors are reported up to 20 percent.

Measurements of pressure in vacuum systems and the analysis of the gases present in vacuum systems are strongly affected by the presence of reactive chemical species. These problems were uncovered in collaborative work between NIST and Intel and are probably not widely appreciated. The conditions that lead to these errors are common in semiconductor processing environments. Ways to solve the problem are unclear at present.

Getting the right information to the right person or computer is a growing problem; the exploding volume of data only exacerbates the issue. (See Table 32.) A wide variety of computers, in tools, in instruments, and in networks, must communicate. Ways to do this that are independent of the kind of computer have been developed, for instance, in the NIST Automated Manufacturing Research Facility. Integration of all of the computer-aided operations from design and engineering through manufacturing and test and on to management and marketing is the goal. In addition, the need to *communicate* must be kept in mind. Simply delivering data is not enough. The right information must be delivered to the right person or computer or machine in ways that make it possible for each of these information users to do his, her, or its job right. This means enough information in the right form and on time. It also means accurate information; erroneous information from poor measurements or other noisy sources is worse than useless.

Device design and verification problems and those related to packaging are listed in Table 42. The device-related issues provide the foundation for integrated-circuit design. At numerous times in the past, when new devices of smaller scale were applied in integrated circuits, corrections to the computer codes for device models have been necessary. Reasonable physical and mathematical assumptions that were made to keep the computer code simpler have turned out not to be reasonable at the smaller scale. Such problems are discovered when the predictions of device performance do not match measurements of performance. Reasons for the discrepancies must then be discovered and the model corrected. This is not a trivial task, usually involving advances in the theory of device operation and sometimes significant new understanding of fundamental solid-state physics.

Table 42
Measurement Needs for Devices and Packages

Devices	Packaging
shrinking scale raises new problems in	physical properties
physics of device operation	materials and interfaces
basic theory, in some cases	polymers, metals, ceramics
use of models for device design	electrical properties
need to develop	dielectric characteristics
theory as needed	metal conductivity
corrections to existing models	chemical properties
changes need verification	polymer degradation mechanisms
deliver results to industry	metal corrosion
	thermal properties
	materials
	interfaces
	anisotropy
	failure mechanisms

As integrated circuits have become more complex and powerful, demands on the package have increased. More connections to the chip are required in many cases, leading to the need for more than 1000 external connections to a single-logic integrated circuit in the latter half of the decade. Not all chips will require this many connections, but quite a few will. Larger chips also generate more heat, adding increased thermal dissipation burdens to large packages. More subtly, differences in thermal expansion between different materials will cause problems. Silicon has a remarkably small thermal expansion coefficient, about a fifth of that of most metals. Attaching a large silicon chip to a metal heat-conducting member will mean the adoption of techniques previously needed for large power devices, but in a flat package configuration that is different from that used for power devices. Plastics generally expand even more than metals. New package configurations and low-expansion plastics will be needed to prevent the differences in thermal expansion from placing excessive stresses on any part of the assembly. Measurements of materials properties, how well plastics, metals, and semiconductors bond to one another, and what causes package materials to fail are all issues. These data are needed for new and larger packages to be designed and manufactured successfully.

To realize the performance advantages of new chips, new packages with appropriately high performance must be developed in parallel with new chip designs. The knowledge base needed to design such packages approaches that needed for designing the chips. Improvements in packaging require extensive new knowledge of the properties of dozens of materials and material interfaces. Without this knowledge, package performance and reliability will limit the abilities of advanced devices to achieve their intended potential.

ENDNOTES

1. A more extensive description of semiconductor technology for the interested lay person is contained in Robert I. Scace, *Semiconductor Technology for the Non-Technologist*, *second edition*, NIST IR-90-4414, published by the National Institute of Standards and Technology (September 1990).

2. Semiconductor materials are so called because their electrical conducting properties are not like metals (conductors) or insulators (non-conductors), but somewhere between. Their important property is that small amounts of certain elements, if present in the semiconductor, dramatically affect the electrical conduction of the semiconductor. This allows the electrical properties of the material to be controlled over an extremely wide range.

3. To get an idea of what this means, consider that there are 5 billion people on earth. At an average weight of 150 pounds, the world's population weighs 0.75 trillion pounds. If you diet and lose 7.5 pounds, you will have lost 10 parts per trillion of the weight of the entire population of the world.

4. This description is of a silicon integrated circuit process. Other kinds of devices are made using similar tools and techniques, though the details vary considerably.

5. Most solids are composed of many tiny crystals, within which the atoms are arranged in a periodic, three-dimensional array. These crystals are not normally observable except under a microscope. Large crystals of many materials occur naturally, and are often quite beautiful. They can also be grown artificially by several methods. Silicon is used in the form of a single crystal because the boundaries between adjoining crystals have different electrical and processing behavior that would interfere with successful integrated circuit manufacture.

6. Transistors are the elementary building blocks of nearly all electronic circuits, including especially integrated circuits. Transistors function as switches (either they are *on* or they are *off*), or as amplifiers, making small electrical signals larger.

7. *1991 Electronic Market Data Book*, Marketing Services Department, Electronic Industries Association, pp. 4 and 113 (1991).

8. *Directions in Engineering Research -- An Assessment of Opportunities and Needs*, p. 248 (Washington, DC: National Academy Press; 1987).

9. *Directions in Engineering Research -- An Assessment of Opportunities and Needs*, p. 182 (Washington, DC: National Academy Press; 1987).

10. Data from VLSI Research (November 1990).

11. Data from this point to the end of the paragraph come from *VLSI Manufacturing Outlook*, VLSI Research, Section 1.9.1, p. 1, and Section 1.9.3, pp. 1, 5 (1990). The data include production by both captive and merchant manufacturers.

12. Mark Moore, "U.S. Chip Makers Bounce Back", *PC Week*, p. 107 (February 1, 1993). The article is based on data from VLSI Research, Inc. and addresses world market share of production for the merchant semiconductor market only.

13. *The VLSI Manufacturing Outlook*, VLSI Research, Inc., San Jose, CA (1990).

14. Data from Dataquest and based on estimates of May 1988.

15. Data from Dataquest.

16. P. K. Chatterjee in *Electronic Engineering Times* (January 29, 1990), and G. B. Larrabee, both of Texas Instruments.

17. Data from Dataquest (June 1987).

18. Data from Dataquest (March 1989).

19. *Technology Trends Assessment*, SRC Summer Study, Wild Dunes, South Carolina (September 6-8, 1989). Further data from G. B. Larrabee, Texas Instruments.

20. SEMATECH is now devoting about half of its expenditures on external contracts, concentrated on four “major thrust” areas -- lithography, metrology, multilevel metallization, and manufacturing methods and processes. See *SEMATECH 1990*, a report to Congress by the Advisory Council on Federal Participation in SEMATECH, p. ES-4 (May 1990).

21. Data from Graydon B. Larrabee, Texas Instruments.

22. More such measurements are commonly made in Japan than in the U.S. (Private communication, G. Dan Hutcheson, VLSI Research.

23. K. Sato, Toshiba; quoted in *Technology Trends Assessment,* James F. Freedman, editor, SRC Summer Study, Wild Dunes, South Carolina, p. 93 (September 6-8, 1989).

24. Successful engineering development of 0.3-µm imagery using standard mercury i-line lithography tools in connection with phase-shift masks was reported at the 1990 International Electron Devices Meeting (San Francisco, December 1990) by more than one Japanese company. Some speakers projected that 0.15-micrometer dimensions will be possible using optical means.

25. Basic physical principles state that the minimum feature size in an optical image cannot be smaller than the wavelength of the light used to produce it. Practical reasons limit resolution of lenses to a performance level somewhat poorer than this limit.

26. Filters do not remove all particles larger than the specified size, nor do they pass all particles smaller than that size. Because unfiltered air contains very many more small particles than large ones, so does the filtered air even though the filters remove a large fraction of the particles smaller than the specified size.

27. DARPA JTECH panel report on computer integrated manufacturing and computer aided design for the semiconductor industry in Japan.

28. R.W. Anderson at a seminar sponsored by the Electronics Industry Association of Japan in Washington, DC (March 25, 1980).

29. *Semiconductor Technology: Workshop Conclusions,* Semiconductor Industry Association, p. 19 (November 17-19, 1992). Contains views generated by 179 of the country’s key semiconductor technologists, drawn principally from U.S. industry and its suppliers and customers, but including some individuals from academia, government agencies, and the National Laboratories.

30. The Semiconductor Research Corporation is a consortium funded by 31 member companies and six government agencies to fund university research on semiconductor-related subjects. It funds about $30 million in research (including about $10 million from SEMATECH) at nearly 50 universities.

31. The Semiconductor Equipment and Materials International is an international industry association of suppliers to the semiconductor device makers and is a large international standards organization.

32. The American Society for Testing and Materials is the largest private-sector standards-producing organization in the U.S. Its Committee F-1 on Electronics is the primary producer of measurement standards for semiconductor materials and processes in the world.

33. See the end of the "Summary" at the beginning of this chapter for a listing of these reports. Distribution of SRC reports is limited to conference attendees, employees of SRC member companies, and faculty members of SRC sponsored research programs.

34. *Semiconductor Technology: Workshop Conclusions*, Semiconductor Industry Association, pp. 1, 31 (to be published, March 1993).

35. A *reproducible* measurement is one that gives the same result day after day. The result may be wrong, but it is always wrong by the same amount and in the same direction, too high or too low.

36. An *accurate* measurement is not only reproducible, but it is free from error to within a stated uncertainty. *Accurate* measurements made at various times and places by different instruments having different operators will agree within the accuracy limitations of the instruments being used.

37. Data are principally from *Technical Trend of Large Diameter Silicon Wafer II*, Masaharu Watanabe, Chairman, Large Diameter Wafer Technology Trend Subcommittee of Japan Electronic Industry Development Association (1990).

38. Richmond B. Clover, in *Proceedings of SRC Topical Research Conference on Metrology for Semiconductor Manufacturing*, Santa Fe, New Mexico, p. 62 (February 21-22, 1990).

39. R. A. Bowling and G. B. Larrabee, in *Proceedings of SRC Topical Research Conference on Metrology for Semiconductor Manufacturing*, Santa Fe, New Mexico, p. 475 (February 21-22, 1990).

CHAPTER 5

MAGNETICS

September 1992

Chapter 5

MAGNETICS

SUMMARY

Magnetic technology serves a very wide range of applications. It is second only to semiconductor technology in its importance to the broad variety of electronic and electrical products. The world market exceeds $95 billion (1990) for products that rely on magnetic materials in performing their primary functions. These products fall principally into four market segments: magnetic materials, electronic components, magnetic information-storage equipment, and electrical equipment. This enormous world market is influenced strongly by technical innovation on all fronts.

Magnetic information storage is the largest segment, about $52 billion per year worldwide for both computer and consumer applications (1990). For computer applications, the U.S. is strongly competitive in two of four areas (rigid disk drives and tape drives) and is weaker in the remaining two (floppy disk drives and magneto-optic storage). The largest market is for rigid disk drives which account for $25.7 billion (1990). The U.S. is in a position of technical leadership in this field and holds 77 percent of the world market (1990). However, this position is being strongly challenged by Japan which is highly advanced in the vertical recording techniques used to achieve very high information densities. For tape drives, the world market is $2 billion (1990). The U.S. holds a strong position with a market share of 91 percent (1990). However, Japan's dominant position in helical-scan technology for consumer applications foreshadows strong competition in the emerging area of very-high-speed tape drives. For flexible disk drives, the world market is $2.4 billion (1990). Japan has captured the major part; the U.S. holds only a 3.3 percent share (1990). For optical information-storage systems, the world market is $0.9 billion (1990), of which $0.3 billion is based on magneto-optical technology; the other technologies are non-magnetic. Japan dominates the overall optical market; the U.S. holds only a 7.9 percent market share (1990). For consumer applications of magnetic information storage, the world market is about $18 billion (1990). Japan dominates and is likely to continue to do so with the introduction of new products, such as high-definition television (HDTV) which will employ magnetic information-storage capacity at unprecedented levels. Japan is strongly positioned to assume a dominant role in the emerging market for erasable magneto-optical storage systems for consumer applications.

The U.S. market for the various types of electrical equipment that rely on magnetic materials to perform their primary functions is at least $10.6 billion (1990). Included in this figure are motors, electric generators, and large power transformers. The world market is likely twice as large, or about $21 billion. The principal competitors for the U.S. market are Japan, Mexico, and Canada, in that order.

The U.S. market for electronic components that rely on magnetic materials to perform their primary functions is at least $7.3 billion (1990), which represents only six groups of components for which data are available. The world market is likely twice this level, or about $14.6 billion. There are many other components for which data are not readily available.

The world market for magnetic materials is about $7 billion per year. The largest part is magnetically soft steels of the type used in the cores of power transformers. Those steels account for about $5 billion per year. The high-performance end of this market is top grade, grain-oriented silicon steels. Two Japanese companies manufacture half of the world's supply of these products. The second largest part of the world market for magnetic materials is high-performance permanent magnets, at $2 billion (1990). As of 1990, Japan held the largest share of this market, 54 percent, while the U.S. and Europe held 20 and 17 percent, respectively (based on 1988 figures). The largest part of this market is ferrites, which accounts for 59 percent (1990). Permanent magnets are used in small electric motors and actuators. They have made possible a revolution in the engineering of small, nimble magnetic drives used in disk and tape drives, printers, etc. The best of the newest materials for permanent magnets is neodymium-iron-boron (NdFeB), which enjoys a small but spectacularly growing world market. This market of $340 million in 1990 and is expected to reach $900 million by 1995 and $1.5 billion by 2000. Because of very effective patent protection, the U.S. and Japan share this market. The U.S. holds nearly one third and Japan holds nearly two thirds. A very active and well coordinated search for a superior substitute is underway in the European Economic Community and could further erode the U.S. position, if successful. Ferrites serve in many applications requiring a wide variety of magnetic properties. Ferrites are critical elements in loudspeakers, microwave components, electromagnetic shielding for diverse electronic products, and "low-observable" (stealth) technologies.

To maintain competitiveness across this broad spectrum of products based on magnetic materials, U.S. industry must pursue key goals in product performance while achieving attractive levels of cost, quality, and reliability. For magnetic information storage, important performance goals include increased information density and speed. For electrical equipment, important performance goals include reduced energy loss for transformers, which may require amorphous magnetic materials, and increased efficiency for motors and actuators, which will require stronger permanent magnets. For electronic components, an important performance goal is improved frequency response to gigahertz frequencies to support new microwave systems for communications and radar. Common to these goals is the need for the development of superior magnetic materials, a fiercely competitive area.

Successful pursuit of these goals by U.S. industry is highly sensitive to the level of measurement capability available to industry for developing new materials, new products based on those materials, and new manufacturing processes to make both. At present, there are major shortfalls in the measurement capability needed for the pursuit of magnetic products with improved competitiveness.

Improved measurement capability is needed for several magnetic quantities, such as coercivity, permeability, and magnetization. There are several different types of improvements that are needed, depending on the particular measured quantity involved and the application: improved accuracy (by up to a factor of ten); improved sensitivity (by typically a factor of ten); broader frequency coverage, (from 0 to at least 25 gigahertz); finer spatial resolution (down to 10 nanometers); accommodation of larger sample sizes (up to 1 meter); and accommodation of more diverse sample shapes. The single most important need is for *improved accuracy*, and the most important means of achieving it is through the development of *measurement reference standards* that can assure the accuracy of measurement methods. For some magnetic quantities, the needed improvements can be accomplished by using existing measurement methods better. For other magnetic quantities new measurement methods are needed with higher capabilities. In both cases, improved documentation, promulgation, and adaptation of measurement methods to forms realizable in modern instrumentation will be needed.

Finally, new reference data on the fundamental magnetic properties of selected critical materials are needed to guide the design of products and manufacturing processes.

INTRODUCTION TO MAGNETICS

Magnetism and electricity are two inseparable parts of the same phenomenon. Magnetic fields are created by the motion of electric charge, including the flow of electric current in a conductor or the motion of electrons in atoms. Magnetic fields can exert strong forces on magnetic materials and on electric currents flowing in conductors. These interactions are the driving forces of electrical machinery. When the magnetic field linking a coiled conductor changes, a voltage appears that drives electric current through the coil and any circuit connected to it. The voltage produced increases with the rate of change of the field. This is how electrical power is generated.

Describing Magnetic Fields

Two distinct quantities are used to describe a magnetic field: the *magnetic field intensity*, H, and the *magnetic flux density*, B. The magnetic field intensity can be readily calculated from a knowledge of the sources of the field, particularly if it is derived from electric currents in wires and magnetic materials are absent. The magnetic field intensity is often called the *applied magnetic field* because it is applied to magnetic materials to elicit a magnetic response from them; that terminology will be used here. The magnetic response of a material to an applied magnetic field is reflected by its *magnetization*. It indicates the degree to which the material has become a secondary source of magnetic field. The magnetic flux density is a combination of the applied magnetic field and the magnetic field produced by the magnetization. It is the quantity that determines the force on a conductor carrying an electric current or the voltage induced in a moving conductor.

Confusion of magnetic flux density and applied magnetic field is a trap into which most students of physics or electrical engineering have fallen at one time or another. It is important to maintain the distinction between these two quantities in any discussion of magnetic measurements.

The response of a magnetic material to an applied magnetic field, as indicated by the resulting magnetic flux density, is a function of the material's *permeability*. Materials with a large permeability have a strong magnetic response to an applied magnetic field; in essence they act to amplify the applied magnetic field. The magnetic flux density and the applied magnetic field can have a complicated relationship. That relationship depends on the magnetic nature of the materials present. For some materials, the relationship also depends on the history of the levels of applied magnetic field to which the material has been previously exposed.

Types of Magnetic Materials

The great wealth of practical applications of magnetics arises from the existence of a great variety of magnetic materials. Magnetic materials can be divided into classes according to the nature of the magnetization that they produce in response to the applied magnetic field. There are five principal classes, as shown in Table 43, along with typical values of their permeability.[1] Diamagnetic properties arise from the orbital motion of electrons in individual atoms. Paramagnetic properties arise principally from the spin of electrons in individual atoms. Ferromagnetic, antiferromagnetic, and ferrimagnetic properties reflect the collective behavior of the spins of electrons across many atoms. All five types of magnetic materials are discussed below, but the ferromagnetic and ferrimagnetic materials are by far the most important for practical applications. They are the only magnetic materials whose response to an applied magnetic field is strong enough to be useful for a broad spectrum of applications.

Table 43
Types of Magnetic Materials

Type of Material	Typical Permeability (relative)
diamagnetic	0.99999
paramagnetic	1.00001 to 1.001
ferromagnetic (soft)	5000 to 200,000
antiferromagnetic	1.01
ferrimagnetic (soft)	2 to 100

Diamagnetic and Paramagnetic Materials

Diamagnetic materials respond to the applied magnetic field by generating a magnetization that opposes the applied field. The opposing field is very weak, reducing the magnetic flux density to typically 0.99999 times what it would be due to the applied magnetic field alone, as shown by the permeability value in Table 43. For practical purposes diamagnetic materials are non-magnetic; they are often used as structural materials in magnet systems. Many plastics and glasses are diamagnetic. Superconductors exhibit much stronger diamagnetic properties. They are not considered *magnetic materials* in the usual sense of the phrase. Superconductors are discussed separately in Chapter 6.

In paramagnetic materials the electron spins are essentially uncoupled and free to orient themselves with the applied magnetic field. They typically increase the resulting magnetic flux density slightly to between 1.00001 and 1.001 times what it would be due to the applied magnetic field alone. At normal temperatures their response is weak enough to be neglected for most practical purposes; as the temperature is lowered, their response increases inversely with temperature. Like diamagnetic materials, paramagnetic materials serve as structural materials. Many metals and many colored inorganic materials are paramagnetic.

Ferromagnetic Materials

Ferromagnetic materials are capable of generating spontaneous magnetic fields without the need for external stimulation. Their electron spins couple from atom to atom to point in a common direction. However, in the absence of exposure to an applied magnetic field, the spontaneous magnetization is confined to small regions known as *magnetic domains* that are oriented in opposing directions. The domain fields tend to cancel each other so that the magnetization of the material as a whole is comparatively small. However, in the presence of an applied magnetic field of sufficient strength, the magnetization from all regions can be aligned. The result is called the *saturation magnetization*. The importance of ferromagnetic materials comes from the fact that the saturation magnetization is so large. The magnetic flux density that remains after the removal of the applied magnetic field is called the *remanence*. It may also be very large in ferromagnetic materials. The remanence determines the strength of a permanent magnet. The remanence also determines the strength of the signal induced in a read head by small magnetized domains on magnetic recording media. They may be less than a micrometer (40 millionths of an inch) in size.

Most ferromagnetic materials are crystalline, and the spins of some of them are coupled to the axes of symmetry of the crystals and line up along certain preferred directions. That is, the magnetization exhibits *anisotropy*, or different properties for different orientations of the applied magnetic field.

Most ferromagnetic materials are metals or alloys from the iron group, the rare earth group, or the actinide group of the Periodic Table of Elements. Common examples of ferromagnetic materials are iron, nickel, and cobalt. At temperatures generally well above room temperature, thermal agitation breaks up the highly ordered state of the spins required for ferromagnetism, and these materials revert to paramagnetism. As the temperature is lowered, the highly ordered ferromagnetic state returns spontaneously. The well defined temperature at which this transition occurs is known as the Curie temperature.

The magnetization of a ferromagnetic material may be reversed by applying a sufficiently strong magnetic field in the reverse direction. The value of the applied magnetic field that drives the magnetization down to a level where its magnetic effects are exactly cancelled by the applied magnetic field (that is, where the magnetic flux density B is reduced to zero) is known as the *coercivity*. Coercivity is an important design parameter for magnetic recording media because it determines the strength of signal required to erase and write information on a disk or tape. Recording media are designed so that the coercivity is high enough to prevent unintended erasure but not so high as to require an excessively strong signal for writing.

Ferromagnetic materials with low coercivity are referred to as *soft ferromagnets*. The direction of their magnetization is easily reversed by reversing the applied magnetic field. These materials are used to intensify and concentrate magnetic fields in machines such as electric motors, generators, and transformers. They may increase the resulting magnetic flux density to 5000 to 200,000 times the level produced by the applied magnetic field alone. For these applications, even small levels of coercivity are a source of power dissipation, which is highly undesirable. Such dissipation can be calculated from the material’s *hysteresis loop*. It is a graph of the relationship of the magnetic flux density to the applied magnetic field as the latter is varied.

At the other extreme are the ferromagnetic materials with very high coercivity. They are referred to as *hard ferromagnets*. They are used as permanent magnets. They increase the flux density from an applied magnetic field somewhat less than the soft ferromagnets, but they retain their magnetization much better; that is, they have a much higher remanence. For many years Alnico magnets (iron-cobalt-nickel alloys with minor additions of aluminum and copper) were used for this purpose. Now they are being replaced in many applications by materials such as samarium-cobalt (SmCo) and neodymium-iron-boron (NdFeB) that have superior magnetic characteristics but are more costly and more brittle. All of these materials have much higher remanence than steels do.

Antiferromagnetic and Ferrimagnetic Materials

In antiferromagnetic materials the electron spins are coupled from atom to atom in such a way that neighboring spins tend to line up in opposite directions and cancel each other, in the absence of an applied magnetic field. Like ferromagnetic materials, the antiferromagnetic materials are paramagnetic at high temperature. As the temperature is reduced their paramagnetism is quenched at a well defined temperature called the Néel point, and they become antiferromagnetic. Antiferromagnetic materials have a relatively weak response to an applied magnetic field. They increase the flux density by a factor of only about 1.01 times the value that would be produced by the applied magnetic field alone. Some common antiferromagnetic materials are the oxides and halides of transition elements of the Periodic Table.

A closely related class of materials, called ferrimagnetic materials, have more complex structures. Neighboring electron spins oppose each other, but they are not of equal strength, producing a net magnetization of the material. Ferrimagnetic materials, like ferromagnetic materials, exhibit a transition at a Curie temperature. Above that temperature they are paramagnetic, and below it they are ferrimagnetic. Ferrimagnetic materials are important to practical applications because their magnetic properties are strong *and* because they are electrical insulators. Their low conductivity enables them to be used at high frequencies where conducting ferromagnetic materials are prone to undesirable circulating electrical currents. These currents cause electrical losses and interfere with desirable magnetic properties. Like ferromagnetic materials, ferrimagnetic materials come in both soft and hard forms. Soft ferrimagnetic materials increase the magnetic flux density by a factor of 2 to 100 times that due to the applied magnetic field alone. Hard ferrimagnetic materials provide a much greater increase. In general, the response of ferrimagnetic materials to an applied magnetic field is somewhat weaker than that of ferromagnetic materials and considerably stronger than that of diamagnetic, paramagnetic, and antiferromagnetic materials. Many ferrimagnetic materials are oxides, such as ferrites or garnets. Ferrites are the dominant materials used in permanent magnets because of their low cost. However, the strength (remanence) of ferrite permanent magnets is about one-third to one-half that of permanent magnets made from ferromagnetic materials.[2] Ferrites are essential to microwave devices, and garnets (which are transparent) are important to optical devices. The antennas of portable AM radios use ferrite rods to concentrate the field of an incoming radio signal.

Like ferromagnetic materials, both antiferromagnetic materials and ferrimagnetic materials exhibit anisotropy.

Other Important Magnetic Properties

The magnetization of a material is not the only observable effect of applying a magnetic field. *Magnetoresistance* is the variation of electrical resistivity with applied magnetic field. Magnetoresistance is important for new read-only heads for magnetic information storage. There is a corresponding effect on the thermal conductivity. The *Hall effect* is closely related to the magnetoresistance. It is manifested by a voltage that is generated in a direction perpendicular to the flow of electric current and the magnetic field. *Magnetostriction* is the variation of the physical dimensions of a material with magnetization. It is the source of the characteristic noise of large transformers. It enables magnetic materials to be used to create sound waves in water for sonar. Each of the above effects finds numerous applications in the broad class of sensors. The magneto-optic *Faraday effect* is observed in transparent magnetic materials. It is the rotation of the plane of polarization of light propagating along the direction of the magnetic field. The magnetic-optic *Kerr effect* is closely related. It is the rotation of the plane of polarization of light reflected obliquely from the surface of a magnetized material. Many of these phenomena are used as the basis for magnetic sensors and transducers in instruments such as gaussmeters, ammeters, and magnetic-disk reading heads.

Another phenomenon that has led to some interesting applications is *magnetic resonance*. Spinning electrons in magnetic materials have angular momentum, in addition to their magnetism. As a result, when a magnetic field is applied, they do not immediately line up with it, but instead precess around the direction of the field like little spinning tops. The result that can be observed in some conditions is a component of the magnetization of the material, in a direction transverse to the applied field, that rotates around it at a characteristic frequency that is proportional to the field strength. In ferrites and garnets this rotating magnetization can be strong enough to be used as the basis for devices such as microwave isolators and circulators. These devices respond to microwaves that enter from different directions differently, changing the path of some or blocking others altogether.

A related phenomenon is *nuclear magnetic resonance*. The nuclei of some chemical elements spin and precess in a way similar to electrons, but the nuclei precess at a lower frequency because they are heavier. The frequency of precession is affected to a small degree by the chemical nature of the molecule in which the spinning nucleus resides. This phenomenon is the basis for a technique of chemical analysis which can identify certain molecules by this so-called "chemical shift" of frequency. The technique is most commonly applied to complex organic molecules, which all contain many hydrogen atoms in different positions with different chemical shifts. This phenomenon is also the basis for magnetic resonance imaging, which has proved to be a very valuable technique for medical diagnosis. The patient is placed in a very uniform magnetic field (provided usually by a large superconducting magnet); and the location of hydrogen is mapped out by making small, controlled variations of the field while monitoring the magnetic resonance. The immediate environment of the hydrogen determines the exact response. The resulting map provides a striking and detailed image of the soft organs in the patient’s body; abnormalities can be readily seen. In comparison, X-ray imaging is not sensitive to soft tissues and employs radiation that can be harmful to body tissues.

APPLICATIONS OF MAGNETICS

Magnetic materials are employed in a wide variety of individual electronic components, in nearly all electronic equipment, and in over half of all electrical equipment. However, some products have a more direct dependence on magnetic materials than others. For them, the *primary function* performed by the product is dependent on the special properties of the magnetic materials they employ. For example, magnetic information-storage equipment depends directly on the performance of magnetic materials to perform its primary function -- the storage of data. In contrast, some other electronic products employ magnetic materials in supportive roles only. For example, they may use magnetic materials in the transformers within their power supplies.

The discussion below focuses on products that rely on magnetic materials for performing their primary function. Of course, this criterion may apply to some types of electronic components, like small power transformers, without automatically applying to electronic equipment incorporating those components. Thus, a distinction is made between the primary function of the component as a product, and the primary function of the equipment as a product.

The array of products meeting this *primary-function* criterion is so vast that not all of them can be discussed here. Even identifying them all would be difficult. However, the major products can be addressed, in several categories: electronic components, magnetic information storage, other electronic equipment, and electrical equipment. *Magnetic information-storage equipment* constitutes such a large market that it has been separated from *other electronic equipment* for special attention.

Each of these categories is discussed below. Magnetic materials relevant to a given application area are addressed as part of that area.

Electronic Components

A wide variety of electronic components depends on magnetic materials for performing primary functions. Both permanent magnets (hard magnetic materials) and soft magnetic materials are employed. Included are small power and signal transformers, inductors, relays, meters, acoustic and other transducers (speakers, headphones, microphones, and phono cartridges), sensors (inductive, Hall effect, magnetoresistive, and temperature sensitive), microwave components (circulators, isolators, filters, and tubes), and cathode-ray tubes (television tubes and computer-monitor tubes). Cathode-ray tubes are included because they depend on magnetic materials for focusing and for correction of geometric distortion (pin cushion), both of which are fundamental to the forming of a visual image. Microwave tubes are included because those used in high power applications (traveling-wave tubes, magnetrons, and gyrotrons) depend on magnetic materials for performing their primary functions of generating or amplifying microwave power.[3]

Magnetic Information Storage

Magnetic information-storage equipment depends directly on magnetic materials to perform its primary function. Magnetic information storage is of great importance to the computer industry and to all industries using computers. Information is stored as a pattern of magnetic domains on a thin layer of ferromagnetic material on the surface of a disk or tape. To meet the demands of the computer, the recorded information must have very high density (that is, each bit must occupy a very small area); and reading and writing must be done at very high speed. Both density and speed are the subjects of intense and highly competitive development efforts. The highest densities available commercially are 20 megabits per square centimeter (Mbit/cm^2) for tape and 10 Mbit/cm^2 for rigid disks.[4] Very high-density experimental rigid-disk drives have been reported, one with 180 Mbit/cm^2 and from the U.S.[5] and one with 310 Mbit/cm from Japan^2.[6] The theoretical limit to magnetic recording density is very high, about 16 gigabits per square centimeter (Gbit/cm^2)[7] for media based on iron. However, other factors will likely prevent reaching this limit. They include constraints on the design of recording heads, on track separation, and on the flying height of heads above the disks. The highest reading and writing speeds available commercially are 34 megabits per second (Mbit/s) for rigid disks[8] and 150 Mbit/s in helical-scan tape drives.[9] To achieve this latter speed, the heads travel at 51.5 meters per second relative to the tape, or about 115 miles per hour.[10] The best head access times for rigid disks are somewhat under 12 milliseconds.[11] Head access time is the average time required for the head to move to the location of the desired data on a disk and to begin reading. The *speed* of a rigid disk is a combination of its reading and writing speeds and its head access time.

This is all a tremendous challenge to the designers of recording heads. To write, these heads must apply a magnetic field strong enough to switch the direction of magnetization of the recording medium, but the field must be confined to an area that does not overlap with neighboring recorded bits. This requirement implies that the recording head must be very small and must be placed very close to the fast-moving surface of the tape or disk. Typically the active area of a recording head for a rigid disk drive is about 6 square micrometers and it is placed only 150 nanometers from the disk. It rides on a film of fast-moving air carried along by the motion of the disk. An experimental drive with heads flying as low as 25 nanometers on a liquid interface has been reported.[12]

Another new recording technology is magneto-optic recording, which uses disks that are magnetized perpendicular to the surface in an initial direction. To write, a uniform field of the reverse direction is applied to a large area of the disk. Individual bits are recorded by using the fine light beam of a laser to heat a very small spot of the magnetic layer instantaneously. When the temperature rises high enough to weaken the coercivity, the magnetization switches under the influence of the reverse field. Reading is done by detecting the Kerr rotation of the polarization of a weaker laser beam reflected from the magnetized layer.

Each recording technology -- magnetic and magneto-optic -- has its advantages. However, both are advancing at such a pace that comparisons are valid only temporarily. Comparing information density, present commercial magneto-optical drives are providing typically ten times higher density than magnetic drives, although high-end products show smaller differences.[13] However, the theoretical limit for magneto-optic recording is not as dense as for magnetic recording. This limit is set by the fact that light can be only be focused to a spot whose size is comparable to the wavelength. This leads to a maximum information density of roughly 200 Mbit/cm^2 for the wavelengths presently available, or about 80 times below the intrinsic capability of iron magnetic recording media.[14] Magneto-optical drives are already approaching this limit, so only small increases in information density can be expected in the future. The experimental magnetic drives noted on page 102 are already functioning at such densities. Also, magnetic drives as a group have lower error rates (1 error in 10^9 bits) than magneto-optical drives (1 error in 10^5 bits). Both work reliably by employing sophisticated error-correction techniques, but the magneto-optical drive must devote a larger fraction of its data capacity (25 percent) to error-correction information, reducing the space available for the data.[15] Comparing speed, commercial magnetic drives offer information transfer rates and head access times that are typically 4 times faster than those for magneto-optical drives.[16] This difference is significant because high speed continues to be the single most important characteristic for the vast majority of applications. Comparing data security, magneto-optical recording provides superior protection of the disks by virtually eliminating the possibility of accidental head contact. In addition, magneto-optical disks are removable, further improving security by enabling easy protective storage. The removable disks provide other advantages, too; they effectively increase the data accessible from a single drive, and they enable ready distribution of data in a highly compact form.

There are other types of magnetic information-storage products in wide use. They include the broad diversity of magnetic reading devices for credit cards, for identification cards, and for magnetic inks on bank checks. The disposable "fare cards" used on some public transportation systems undergo both read and write cycles during their short lifetimes. Toners and techniques of electrophotography (xerography) and the less well known magnetography also fall in this category.

Other Electronic Equipment

A diversity of electronic equipment, beyond information-storage equipment, depends on magnetic materials for performing primary functions. Both permanent magnets and soft magnetic materials are employed. Examples include watt-hour meters, electronic balances, ore separation equipment, and new magnetic-resonance-imaging equipment employing permanent magnets to substitute for superconducting magnets. Also included is research equipment such as particle accelerators, synchrotron radiation sources, free-electron lasers, and mass spectrometers.[17]

There is a second class of equipment whose dependency on magnetic materials is considerable but which falls on the margin of what is usually thought of as equipment dependent on magnetic materials for performing the primary function. Examples includes televisions sets and computer monitors, by virtue of their dependence on cathode-ray tubes. Also included is electronic equipment containing speakers, if audio output is a primary function, such as radios, stereo equipment, television sets again, and public-address systems. For these several examples, the primary functions are the production of sound, light, or both; and magnetic materials are directly involved in performing those functions.

Finally, there is a class of electronic equipment -- instrumentation -- that may or may not employ magnetic materials to perform the primary function, but that is focused on exploiting the magnetic properties of other materials to probe them for information. Some of these products, such as magnetometers and susceptometers, are designed to measure well defined magnetic characteristics of materials. Others use measurements of secondary magnetic characteristics to diagnose related conditions. For example, the remanent magnetization of steel is used as an indicator of its metallurgical condition; the inductance of a coil magnetically coupled to a work-piece is used as an indicator of cracks; the nuclear magnetic response of the human body is used by magnetic-resonance-imaging equipment to detect abnormalities; and the magnetic anomalies associated with submarines in the ocean are used by the Navy to track them. The oldest instrument in this class is the magnetic compass, which has given sterling service for many centuries.

Electrical Equipment

Electrical equipment has been a prime user of magnetic materials for a long time. Soft ferromagnetic steels (very low coercivity) were at the heart of the electrical machines that drove the heroic industrial age.[18] Today, large motors, generators, and transformers contain millions of tons of this material; and it is still being used to build new machines. Industrial controls, switchgear (including circuit breakers), and related apparatus also employ magnetic materials.

Today, new materials are beginning to replace the early steels used in many of these applications. Foremost among these are the amorphous magnetic materials for those applications requiring a soft magnetic material with very low losses. They are metals that are created by a process that suppresses the crystalline nature that nearly all metals normally have. As a result, the materials have extremely low coercivity and make very efficient machines. They will be more widely used if the major problems can be successfully resolved: (1) high costs relative to ferromagnetic steels; (2) difficulties in manufacture of sheets in usefully large sizes; (3) loss of desirable magnetic properties in the presence of heat; and (4) low resistance to corrosion.[19] The last two of these are serious issues only in specific applications.

Small motors and actuators also employ magnetic materials. Permanent magnets and soft magnetic materials are both used. Computer peripherals employ small motors in printers and in tape and disk drives. They offer extremely nimble performance to match the demands of the computers. The disk drives also employ magnetic actuators (steppers and voice coils) to position read/write heads with great accuracy. The motors for the drives are a very active field of mechanical engineering and their success contributes to the success of the computer industry. Small motors and actuators are also important to the automobile industry. A modern automobile contains up to 40 electric motors. Given that 6.6 million automobiles were manufactured in the U.S. in 1990, the number of motors involved is significant.[20]

Other Components and Equipment

There are a variety of other components and equipment that employ magnetic materials and that fall outside of the usual definitions of electronic or electrical equipment. They include contacting, holding, lifting, and traction devices, magnetic bearings, magnetic suspensions, magnetic locks,[21] and magnetic seals (employing ferrofluidic materials). Since the focus here is on electronic and electrical applications, these will not be addressed further, but their importance should be noted. An interesting application of magnetic suspensions that has been demonstrated but that remains somewhat futuristic is magnetically levitated trains; some approaches use magnetic materials, while others use superconductors.

Relation to Superconductors

The advent of practical superconductors introduced a new dimension to magnetic engineering. Coils of superconducting wire can generate magnetic fields up to six times as strong as the highest that can be generated using magnetic materials, and over much larger volumes.[22] This opens up prospects for compact, efficient electrical machines and for new applications requiring very strong magnetic fields. Examples include magnetic resonance imaging, magnetic ore separation, magnetically levitated trains, and controlled thermonuclear power generation from an extremely hot plasma contained and compressed by magnetic fields. These are mostly developments for the future, except for magnetic resonance imaging and magnetic ore separation, which are already commercial technologies.

ECONOMIC SIGNIFICANCE AND U.S. COMPETITIVENESS

As noted above, the products that rely on magnetic materials to perform their primary functions are highly diverse. There are significant markets for materials, electronic components, electronic equipment, and electrical equipment. Estimating the market sizes presents significant difficulties. The array of products is so vast that not all can be identified. Further, for those identified, market data are not always available. When available, the data may refer to a year earlier than wanted, or the data may be limited to just one of the following: U.S. shipments, U.S. market, or world market. To cope with this latter difficulty, the following observation has been employed: For a broad spectrum of electronic products, the world market is often 2 to 2.5 times larger than the U.S. market alone. Based on this observation, a factor of 2 has been used to provide a conservative estimate of the world market, based on data for the U.S. market. All such estimates are starred (*) in the tables that follow. In one case, the U.S. market has been estimated from the world market with the same factor of 2, also starred (*). The endnotes describe the approaches taken to accommodate other shortfalls in available data.

The discussion that follows sets lower limits on the U.S. and world markets for products relying on magnetic materials to perform their primary functions. The products are discussed by market segment, but the category *other electronic equipment*, discussed above, has not been included below. One reason is the lack of good economic data for many of the products in that category. Another reason is that the category contains products, like television receivers and computer monitors, that many may feel are on the margin of those qualifying for inclusion here. The estimated world markets for the remaining qualifying products in the various categories are summarized in Table 44. Clearly, the totals for the U.S. market ($48 billion) and for the world market ($95 billion) suggest that the lower limits are substantial. The individual market segments in the table, and the origins of the data for the individual segments, are discussed separately in the following several sections with a focus on market size and competitiveness information.

Table 44
World Markets for Electronic and Electrical Products (1990)

	U.S. Market ($billions)	World Market ($billions)
Magnetic materials	3.5 *	7.0
Electronic components	7.3	14.6 *
Magnetic information storage	27.0	51.7
Electrical equipment	10.6	21.2 *
Total	48.4	94.5

Magnetic Materials

The world market for magnetic materials is at least $7 billion per year. The largest segment for which data are available is soft ferromagnetic steels, about $5 billion per year. These steels are used in electric power equipment and in many electronic components. About $1 billion per year of the world market is top grade so-called grain-oriented silicon steel, half of which is produced by just two Japanese companies.[23] The second largest segment for which data are available is permanent magnets at $2 billion (1990).[24] Japan holds the largest share of this market, 54 percent (1990), as shown in Table 45.[25] Unfortunately, a breakdown for the U.S. and European shares is not available for 1990; however, the relative percentages may be inferred from the 1988 data in the table. Data are not available for a third important market segment, soft ferrimagnetic materials.

Table 45
World Market Shares for Permanent Magnets by Country/Region

	World Market Share (percent)	
Country/Region	1988[25(a)]	1990[25(b)]
Japan	54	54
United States	20	46
Europe	17	
Others	9	
Total	100	100

Four types of permanent magnets account for virtually all of the world market for permanent magnets, as shown in Table 46.[26] One type is a hard ferrimagnetic material (the ferrites), and the other three are hard ferromagnetic materials. The ferrites are the largest segment by both dollar value and total mass (97 percent as of 1988).[27] Their low cost has led to wide application despite the fact that hard ferromagnetic materials make stronger magnets. Neodymium-iron-boron (NdFeB) constitutes the fastest growing segment of the market for permanent magnets. The world market for this material increased a factor of two from 1988 to 1990 to reach $330 million and is expected to reach $900 million by 1995 and $1.5 billion by 2000.[28]

Permanent magnets are used in electric motors and actuators with power ratings less than one-third horsepower. The very small, high-performance drivers of computer peripherals have pressed the performance of materials to the limit. They have been made possible by the availability of relatively new materials such as samarium-cobalt (SmCo) and NdFeB. NdFeB in many major respects outperforms all other known permanent-magnet materials.

Table 46
World Market for Permanent Magnets by Type (1990)

Type	World Market ($billions)	Market Share (percent)
ferrites	1.17	59
neodymium-iron-boron	.33	16
samarium-cobalt	.31	16
Alnico	.19	9
Total	2.00	100

NdFeB is also unusual in its effective protection by patents. Two competing processes for manufacturing it are covered by patents held by General Motors (U.S.) and Sumitomo (Japan). General Motors dominates the U.S. share of the market. Sumitomo holds 42 of Japan's 64 percent share of the world market (1990).[29] The U.S. market share is sustained partly by patent protection for the production process of NdFeB. The U.S. position could change drastically if a superior substitute material is found. The motivation to find a substitute is strong because NdFeB has a low Curie temperature and is very susceptible to corrosion. The European Economic Community has organized an effort known as the Concerted European Action on Magnetics (CEAM) to search for new rare-earth permanent-magnet materials. This organization has already discovered a samarium-iron-nitride (SmFeN) material with a higher Curie temperature than NdFeB and other superior magnetic properties that appear highly promising.

Ferrites in both soft and hard forms are evolving rapidly into broader product lines and are critical to the function of many expensive products. Ferrites are used for microwave components, for loudspeakers, for filtering and blocking electromagnetic interference in electronic products, and for absorbing electromagnetic radiation for low-observable technologies like stealth aircraft, among other applications.

Electronic Components

The U.S. market for electronic components whose primary function is dependent on magnetic materials is also significant. Selected components for which market data are available are shown in Table 47.[30] U.S. market data are not available for the category *other microwave components* (those based on ferrites) in the table, so U.S. shipments have been used as an estimator of the U.S. market where indicated (**). World market data are not available for any of the components, so the factor of 2 discussed above has been used to estimate the world market (*). There are many other electronic components which should appear in the table but for which market data are not readily available. They include headphones, microphones, sensors, cathode-ray tubes other than television picture tubes (such as those for computer monitors), and many others.

Table 47
U.S. and World Markets for Electronic Components (1990)

Component	U.S. Market ($billions)	World Market ($billions)
Transformers, coils[30(a)]	2.7	5.4 *
Relays[30(b)]	1.1	2.2 *
Loudspeakers[30(c)]	1.4	2.8 *
Television picture tubes[30(d)]	1.4	2.8 *
Microwave tubes and amplifiers[30(e)]	0.4	0.8 *
Other microwave components[30(f)]	0.3 **	0.6 *
Total	7.3	14.6 *

Magnetic Information Storage

The market for magnetic recording products is the largest of the markets for all types of magnetic products, about $52 billion in 1990. The major elements of the U.S. and world markets are shown in Table 48.[31] For product categories for which world market data are missing, an estimate has been made based on U.S. market data using the factor of 2 discussed above (*). Magnetic information-storage equipment serves both computer and consumer applications.

Table 48
U.S. and World Markets for Magnetic Information-Storage Products (1990)

Computer Peripherals	U.S. Market ($billions)	World Market ($billions)
Rigid Disk Drives[31(a)]	14.3	25.7
Flexible Disk Drives[31(b)]	0.9	2.4
Digital Tape Drives[31(c)]	1.0	2.0
Magnetic Storage Media[31(d)]	1.7	3.4 *
Subtotal	17.9	33.5
Consumer Electronics		
Video Cassette Recorders[31(e)]	2.4	4.8 *
Camcorders[31(f)]	2.3	4.6 *
Audio Tape Players/Recorders[31(g)]	3.1	6.2 *
Blank Cassettes[31(h)]	1.3	2.6 *
Subtotal	9.1	18.2 *
Total	27.0	51.7

For computer applications the largest segment is rigid disk drives. They are the subject of intensely competitive technical development. Both information density and speed of access double every two or three years as a result of new technology developments. At present the U.S. is in a position of leadership, and holds 77 percent of the world market of $25.7 billion (1990).[32] A coordinated and well funded national effort in Japan is rising to this challenge. Japan is particularly strong in the futuristic vertical recording technology, which has the potential for achieving very high information densities.[33] Commercial rigid disk drives are now available with total storage capacity of over 1 gigabyte, and half the market revenue is for drives with capacity over 500 megabytes.[34] These products are the high end of the market.

In the number of units sold, flexible disk drives for computers are a larger part of the market (41.6 million in 1990, versus 26.7 million rigid disk drives), but because of their much lower price, the total revenue is much less ($2.4 billion in 1990, worldwide).[35] U.S. manufacturers have almost completely lost this part of the market to Japan. The U.S. market share is only 3.3 percent (1990).[36] It would be difficult to win this market back. U.S. industry may conclude that it would be more effective to develop a competing technology, such as small, high-density, removable disk cartridges.

The U.S. competitive position is presently much stronger for magnetic tape drives for computers. The world market is $2.0 billion (1990), and the U.S. holds a 91.3 percent share.[37] However, the outstanding work on tape drives for the most demanding applications, high-definition television (HDTV), has been done by Japan which already has fielded a commercial HDTV recorder.[38] This recorder employs helical-scan technology to obtain very high data rates. The helical-scan technology, similar to that employed in video cassette recorders and new digital audio tapes, is now being applied to information-storage products for computers. Japan's strong lead in helical-scan technology for audio and video applications positions it well to compete in the marketplace for helical-scan tape drives for computers.

In the closely related market for optical disk drives for computers, Japan dominates. The U.S. holds only 7.9 percent of the world market (1990). That market is $0.9 billion (1990), of which $0.3 billion is for magneto-optical disk drives. The other optical drives employ non-magnetic technologies. This market is growing very rapidly and is expected to nearly double by 1992.[39] These drives employ either all-optical technology for read-only or write-once capability or a combination of magnetic and optical technology for multiple read and write (*erasable*) capability. This market has not been included in Table 48.

The other major segment of the magnetic recording market is in consumer electronics. Video cassette recorders, camcorders, audio equipment, and blank cassette tape together constitute a domestic market of $9.1 billion (1990), mostly in imported equipment. (U.S. production of video cassette recorders accounted for less than 10 percent of total U.S. consumption of video cassette recorders in 1990.[40]) The world market is likely $18 billion or so, using the factor of 2 discussed above. The Japanese talent for producing attractive, high-quality consumer products at a competitive price has been a major factor in this field. Japan's leading position in the development of high-definition television will further this trend.

Electrical Equipment

The U.S. and world markets for selected pieces of electrical equipment (transformers, motors, and generators) whose primary function is dependent on magnetic materials are shown in Table 49.[41] The major international customers for U.S. products are Canada and Mexico; and the major suppliers are Canada, Mexico, and Japan. A two percent real compound annual growth rate is forecast for U.S. shipments to this market for 1991 through 1995.[42] While data for the world market are lacking, a conservative estimate would be $21 billion, using the factor of 2 discussed above (*). In addition, there are markets for other electrical equipment (industrial controls, switchgear, and related apparatus) which depend on magnetic materials for their primary function. The dependence of these product lines on magnetic materials is significant because of the magnetic coils and motors that operate so many of them, but it is difficult to sort out which should be included here. For this reason no entries for these product lines are shown in the table. Nevertheless, U.S. shipments give a measure of the significance of these product lines: about $5.8 billion for controls and $5.3 billion for switchgear and related apparatus (1990).[43]

Table 49
U.S. and World Markets for Electrical Equipment (1990)

	U.S. Market ($billions)	World Market ($billions)
Transformers[41(a)]	3.8	7.6 *
Motors and generators[41(b)]	6.8	13.6 *
Total	10.6	21.2 *

The electrical utility industry is a major user of electrical equipment. This industry sold electricity valued at $176 billion in 1990, the year of most recent data.[44] The electric power industry puts a high premium on efficiency and reliability. At present about 9 percent of all electricity generated for the U.S., or about $15 billion in revenue, is lost in various forms of dissipation during transmission and distribution.[45] Losses in power transformers alone may account for 5 percent. The development of amorphous metal transformer cores promises to improve efficiency without sacrificing reliability, provided that the problems with amorphous materials, noted on page 104, can be resolved.[46] The introduction of amorphous materials may tip the balance of the market in the favor of the nation that leads this innovation. At the consumption end, about 60 percent of the electricity consumed in end uses in the U.S. goes to electric motors, so even small improvements in their efficiency can have major economic impact.[47]

Education and Research

Every part of the commercial field of magnetics is likely to be the subject of international competition for at least the next decade. The nature of the technology is such that technical superiority will be a major factor in winning the competition. The U.S. industry will need to mount an aggressive research effort, which in turn will require trained researchers. A move has been made to meet this need, with the establishment of five organizations: the Magnetics Technology Center at Carnegie Mellon University, the Center for Magnetic Recording Research at the University of California at San Diego, the Center for Magnetics and Information Technologies at the University of Minnesota, the Institute for Information Storage Technology at Santa Clara University in California, and the Center for Materials for Information Technology at the University of Alabama at Tuscaloosa. In addition, the U.S. is served by two key national magnet facilities: the Francis Bitter National Magnet Laboratory at Massachusetts Institute of Technology and the new High-Field Magnet Facility at Florida State University.

EVOLUTION OF THE TECHNOLOGY

The fastest moving part of magnetic technology is recording. The two prime measures of performance -- density of information storage and speed of access -- are both doubling every two or three years. The advances are not achieved by a steady process of refinement, but rather by introducing radical changes in materials and design, always with the constraint that this is a commercial technology and the cost of fabrication is all-important.

The magnetic media that established the magnetic recording industry consisted of small ferric oxide particles dispersed in a non-magnetic, plastic binder in a thin layer on the surface of a tape or disk. The recording heads were tiny coils of wire mounted on ferrite yokes to concentrate the field in a very small region. The recorded information was in the form of local reversals of the magnetization in the plane of the surface. Reading was done with the same head by the inverse of this process; the passing magnetization pattern induced a voltage in the head. This technology is inexpensive to manufacture and adequate for many uses, such as audio (analog) recording. It is still in high-volume production. To advance to higher levels of information density and speed, improvements in both recording media and recording heads are necessary.

Improvements in the media have focused on developing finer-grained, high coercivity materials. The first improvement was to use smaller ferric oxide particles with the addition of cobalt to increase the coercivity. Then finely divided metal particles were used (mostly iron-cobalt alloys). The next step was continuous metal films followed by multifilm media. The films must have high enough electric resistance to suppress eddy currents during the recording process, so various alloys are commonly used. Some of these alloys, such as cobalt-chromium, tend to have a direction of easy magnetization that lies perpendicular to the surface of the film. These materials opened the field of *vertical recording*. In vertical recording, the magnetized regions stand vertically inside the media (perpendicular to its surface) rather than lying flat in the media (parallel to its surface). The vertical orientation lends itself naturally to high information density because of reduced interaction among neighboring domains. However, this technology requires a magnetically soft underlayer for flux closure, and such a layer is very sensitive to ambient magnetic fields.

Improvements in recording heads have to keep pace with improvements in media. The key focus is on increased resolution to exploit the higher information density of new media. Higher resolution requires intensifying the field produced by recording heads and confining it to smaller regions on recording media. For the ferrite heads, the limitation was the saturation magnetization of the ferrite core. Some metallic materials have higher saturation magnetization, but they must have high electrical resistance to suppress eddy currents that would interfere with the recording process. Thus in high-performance recording heads, the very tips of the ferrite core are covered with a thin film of an alloy with high saturation magnetization and high electrical resistivity. The common choice is Sendust, an alloy of aluminum, iron, and silicon. A more sophisticated head contains a coil fabricated as a thin film with an iron-nickel (FeN) core (a soft ferromagnetic material); its high resolution comes from its high saturation magnetization which allows higher coercivity media with its resulting smaller bit size. Thin-film heads are increasingly used in recording systems with high information density. Both ferrite and thin-film heads can read and write.

As the size of each recorded bit is reduced, the signal received by the reading head becomes smaller. The need for greater sensitivity is motivating the development of magnetoresistive reading heads. They sense the magnetic signal as a variation in the magnetoresistance of a small strip of thin metal film. Their sensitivity is independent of the speed of scanning, unlike heads using coils, a further advantage. To achieve even higher performance levels, a search for new materials with higher magnetoresistance has been undertaken. The most promising materials at present are *superlattices*. They are sandwiches of alternating very thin layers of different metals only a few nanometers thick. This structure affords a high level of control of materials properties. Promising results have been obtained with alternating layers of iron and chromium and with alternating layers of cobalt and copper. Magnetoresistive reading heads have the potential advantage that they can be made very compact. However, they cannot write.

The materials used for magneto-optic recording media are chosen to have vertical magnetization and a low temperature at which they can switch the direction of magnetization under the influence of an applied magnetic field. They are usually amorphous alloys of rare earth metals with iron or cobalt. A recently discovered material with great promise is a superlattice of cobalt and platinum layers.

There is a constant and very fruitful search for new magnetic materials to satisfy special requirements for saturation magnetization, coercivity, anisotropy, ease of fabrication, corrosion resistance, etc. The variety that nature provides is being augmented by artificial materials such as the superlattices. But the whole industry is acutely aware of the tradeoffs between performance and price. The established materials have set a high standard in this respect.

New magnetic instrumentation is constantly appearing; even the magnetic compass may be challenged by a magneto-optical sensor. The most significant trends are to great sensitivity, taking advantage of superconducting sensors, and to very high spatial resolution, to serve the needs of research on advanced recording media. For great sensitivity, the supreme magnetic sensor is the Superconducting Quantum Interference Device (SQUID). This is commercially available in various systems to measure the magnetic characteristics of materials. Perhaps the most interesting potential use of this device is to form an image of the electric currents flowing in the human brain, as a tool for medical diagnosis. This has been demonstrated using an array of SQUID sensors, but must await research on the interpretation of the images before it becomes useful to medical practice. Other new commercial devices, such as the Alternating Gradient Force Magnetometer, offer similar sensitivity in some applications without the requirement for low temperature operation. For very fine (submicrometer) spatial resolution, two recent instruments have shown outstanding performance. The first is a Scanning Electron Microscope with Polarization Analysis (SEMPA), which forms images of the magnetic polarization of the surface of a material.[48] The second is the Magnetic Force Microscope (MFM), which forms images of the distribution of the magnetic flux density, usually near the surface of a material specimen.[49] These instruments are examples from a formidable array of newly developed experimental techniques that use electrons, X rays, neutrons, polarized light, etc. to probe the magnetic state of materials. These techniques constitute a field of research with great practical value.[50]

GOALS OF U.S. INDUSTRY FOR COMPETITIVENESS

U.S. industry's goals for improved international competitiveness in magnetic products include reduced product cost and improved product performance, quality, and reliability. Here are some important examples. For magnetic information storage, key performance goals include increased information density and increased speed of writing and reading for magnetic information storage. For electrical equipment, a key goal is reducing energy loss while maintaining high reliability. For electronic components, key performance goals include extending frequency response and achieving higher energy efficiency. For all types of magnetic products, reduced product size is important. For example, smaller magnetic disk drives can accommodate more efficient packaging in desktop computers and can support the exploding market for laptop computers. Smaller electrical equipment can live better within the size limitations of machines that can be delivered. Smaller electrical equipment also reduces materials requirements during manufacturing and space requirements when installed. In some cases the reduction in size for a given power level enables more powerful machines to be built to provide additional economic benefits.

To pursue these goals successfully, advances in the development of magnetic materials and in the design of products and manufacturing processes will be needed. Key examples of the need for improved magnetic materials come from each of the affected market sectors. For magnetic information storage, development of the special materials required for heads capable of high-speed reading of very dense information on magnetic media presents special challenges. For electrical equipment, the development of amorphous materials or superior soft steels is necessary for further progress in loss reduction. Similarly, the development of superior permanent-magnet materials is a key requirement for more efficient motors and actuators. For electronic components, the development and characterization of ferrites with excellent high-frequency properties is critical to new microwave components that can open higher microwave frequencies to emerging communications systems. Improvements (higher permeabilities) are also needed in the performance of ferrite materials operating at lower radio frequencies.

Pursuit of all of these goals is highly sensitive to the level of measurement capability used for developing materials and for designing products and manufacturing processes. At present, there are major shortfalls in the measurement capability needed for the pursuit of magnetic products with improved competitiveness.

EVIDENCE OF MEASUREMENT NEEDS

The most systematic survey of the needs for improved magnetic measurement methods and measurement reference standards was conducted in 1984 by F.R. Fickett of NIST.[51] A six-page questionnaire was sent to members of the IEEE Magnetics Society, the Magnetic Materials Producers Association (MMPA), and the appropriate ASTM and ANSI committees; to manufacturers of magnetic instrumentation; to standards laboratories; and to numerous consultants. About 2000 copies of the questionnaire were sent out, and 487 completed responses were received. (About another 20 incomplete responses carried some useful information.) About 40 percent of the respondents (the largest group) were in the computer industry. Producers and manufacturers of materials, electronic components, and instrumentation were also well represented.

The questionnaire listed 18 magnetic quantities and asked questions about the importance of measuring each quantity, the range of accuracy required, and the need for national measurement reference standards. The responses to the survey indicated that all of these quantities were of practical importance, as well as several others that were specifically related to magnetic recording and to the microwave properties of materials. The information developed in the survey, updated to the present, is reflected in the measurement needs described in the next section. Of all the organizations responding to the survey, 33 percent indicated a need for traceability of measurements to national measurement reference standards. Of these, 73 percent cited the need for improved consistency of measurement while 23 percent cited the requirement for traceability in military contracts as the principal motivation.

Many comments received with responses to the survey indicated the importance of improved measurement capability. The following, from R.J. Parker, at that time Chairman of the Engineering and Standards Committee of the Magnetic Material Producers Association (MMPA), is typical:

> "It would appear that our numbers are of little use until we do something about measurement accuracy. Our own round robins and those of others indicate we have in our industry a very serious problem. We do not agree among ourselves, we do not agree with our customers, we do not agree with other countries."

Since the survey was completed, there has been little progress towards solving the problems of standards and quality control in magnetic measurements.

The survey found a need for greater accuracy and consistency in measurements for a wide range of magnetic quantities. This need has since been illustrated by an interlaboratory measurement intercomparison for coercivity conducted by NIST in 1987. The coercivity of magnetic recording tape was measured by 56 participating tape manufacturers on carefully selected and calibrated specimens. Coercivity determines the field strength required to record or erase information, and must be controlled and specified within narrow limits in high-performance recording media. Before the intercomparison was made, there was widespread belief that measurements made by each of the two principal measurement techniques (vibrating sample magnetometer and B-H looper) were self-consistent but were not consistent with measurements made by the other technique. The opposite proved true. The results showed a wide spread of values for each technique but close mean values. Specifically, the spread of values for both techniques was 44 percent while the means differed by only 2 percent. The wide spread of values made the proximity of the means not significant. The spread suggested the need for improved consistency in carrying out the methods, a better understanding of time and temperature effects, and the need for measurement reference standards to assure both consistency and accuracy.

The need for more consistent and accurate measurements, supported by national measurement reference standards, has been voiced by several standards organizations that have committees for magnetic measurements. These organizations include the International Disk Equipment and Materials Association (IDEMA), the IEEE Magnetics Society, the American National Standards Institute (ANSI), the American Society for Testing and Materials (ASTM), and the Magnetic Materials Producers Association (MMPA). These organizations do what they can in the absence of the supporting standards.

A new organization, the National Storage Industry Consortium, has been formed by industry and academia to focus on magnetic recording issues. This group has undertaken a number of cost-shared programs in support of recording issues with the Government through the Advanced Technology Program at the National Institute of Standards and Technology and the Defense Advanced Research Projects Agency. The Consortium has indicated an interest in many of the issues raised here and might serve as a forum for further discussions in this area.

MEASUREMENT NEEDS

Critical measurement-related needs arise in all major industry segments: magnetic materials, electronic components, magnetic information-storage equipment, other electronic equipment, and electrical equipment. However, the nature of these needs differs somewhat from one industry segment to another and from one measured quantity to another. The needs take the form of measurement methods, measurement reference standards to assure the accuracy of the measurement methods, and measured materials reference data. Table 50 summarizes the types of improvements needed in measurement methods in the most general terms. They include accuracy, sensitivity, frequency coverage, spatial resolution, and ability to accommodate larger sample sizes and more diverse sample shapes. The single most important type of improvement needed across all industry segments is *measurement accuracy*. The most important means of achieving that higher accuracy is development of new *measurement reference standards* that can assure the accuracy of the measurement methods. The principal measurement reference standards needed are in the form of standard reference materials. They provide accurately known, stable values of specific magnetic properties for use in calibrating measurement equipment.

Table 50
Types of Improvements Needed in Magnetic Measurements

Improvement	Specifics
higher accuracy	factors of 10
higher sensitivity	factors of 10
broader frequency coverage	from 0 to 25 gigahertz
finer spatial resolution	down to 10 nanometers
larger sample sizes	up to 1 meter
more diverse sample shapes	

Achieving improved measurement accuracy requires using existing measurement methods better, so that their intrinsic accuracy can be realized in practice, and developing new measurement methods when higher levels of accuracy are necessary. In both cases, more effective standardization and promulgation of measurement methods are required. Important, also, is the adaption of measurement methods to forms realizable by industry in modern instrumentation. These needs arise even for common quantities, such as magnetization, coercivity, and noise determination. A handbook or monograph on each measurement method has been suggested as a key step in the solution.

Driving Forces by Industry Segment

Summarized below are the key driving forces behind the measurement needs of the various industry segments. The needs associated with the magnetic materials used by a given segment are discussed with that segment rather than separately. The most important general types of improvements needed, from Table 50, are highlighted for each industry segment. In a later section, each quantity requiring improved measurement support is discussed individually.

Electronic Components

Improved measurement capability is needed to support magnetic components and the key classes of magnetic materials used in their manufacture: soft and hard ferromagnetic materials, and soft and hard ferrimagnetic materials particularly the ferrites. The principal driving forces are several. Improved hard magnetic materials, both ferromagnetic and ferrimagnetic, are needed to provide stronger permanent magnets for more efficient and smaller components. Improved soft ferrimagnetic materials are needed to provide higher frequency response for microwave applications. Improved soft ferromagnetic and ferrimagnetic materials are needed to provide higher efficiency or signal fidelity in a variety of components, such as small transformers and coils.

In general terms, the most important types of improvements needed are higher accuracy, broader frequency coverage, and accommodation of more diverse sample shapes. To achieve higher accuracy, new measurement reference standards will be required for some quantities. Also, new materials reference data will be needed for important materials to support the design of products and manufacturing processes.

Magnetic Information Storage

This is the magnetic technology that poses the greatest measurement challenge. All forms of magnetic information storage share a common dependence on precise and rapid manipulation of magnetic domains and their boundaries. Thus the improvements in measurement technique discussed here will serve rigid disk drives, flexible disk drives, tape drives, magnetic bubble memories (a technology with a niche market), and vertical-Bloch-line memories (a potential future technology).

Common to these technologies is the importance of small structures, and consequently, of measurements on very fine dimensional scales. These small structures are needed to achieve high levels of information density in magnetic storage media and to achieve high levels of performance (particularly speed) for write and read heads. At noted above, measurement of key magnetic quantities with spatial resolution down to 10 nanometers (0.4 microinches) is critical to near-term advances in the field. Since the spatial resolution of good optical instruments is limited to approximately the wavelength of the light used (500 nanometers for yellow visible light), all the sophisticated variations on optical microscopy that have served so well in the early development of many small-scale technologies will not be usable. Instead, new magnetic techniques based on scanning electron microscopy and scanning tunneling microscopy will be required and must deliver five to ten times finer resolution than they now provide. An almost equal challenge is to confine to very small regions the stimulating fields that are used in measuring the response of magnetic recording media. This is similar to the challenge of designing a recording head for high-density recording.

Further, industry must fabricate special magnetic structures with control down to atomic sizes (just below 1 nanometer). They are especially important to the development of magneto-resistive heads with very high sensitivity. Such magnetic microstructures can be very complex. Some are *superlattices* described above, containing many thin alternating layers of material of different composition. The regularity of the interfaces will affect the overall magnetic properties almost as much as the composition of the layers. Fabrication of these materials requires special measurement methods, including reflection high-energy electron diffraction (RHEED) techniques and scanning transmission electron microscopy (STEM). These methods are used to determine the structures of the layers being produced. RHEED can be used while processing is underway to provide real-time control. Present measurement capabilities employing these methods are adequate to current and near-term needs.

Another feature common to all magnetic recording and memory technologies is the need for ever higher frequency response. Recording speeds are doubling every two or three years. For the development of new recording equipment and media, new measurement capability will be required to determine the variation of magnetic properties with frequencies up to at least 10 gigahertz. In particular, special attention will be needed for signal-to-noise measurements since signal-to-noise problems increase significantly at the higher frequencies.

Improvements in measurement techniques are needed now to support the development of both advanced recording heads and advanced media. As the new developments become products, the same measurement techniques will have to be refined and supported by measurement reference standards to provide for quality control and performance specification. The need for tests for quality control of production components will become more critical as the next generation of higher recording densities arrives. Controlling magnetic structure on the microscopic scale at the production line will be necessary. This will require magnetic measurement techniques with sufficient sensitivity to allow determination of the effect of a wide range of process variables on the final thin film or multilayer film device or medium.

In general terms, the most important types of improvements needed are higher accuracy, higher sensitivity, broader frequency coverage, finer spatial resolution, and more diverse sample shapes. This is a very broad spectrum of needs that reflects the high demands of magnetic information storage of supporting measurement capability. New measurement reference standards will be needed for some quantities to support higher accuracy. New materials reference data will be needed for some critical materials.

Other Electronic Equipment

The measurement needs for electronic equipment, other than magnetic information-storage equipment, are the same as those described for electronic components above, since virtually all of these needs arise from the components.

Electrical Equipment

Improved measurement capability is needed to achieve higher energy efficiency levels while maintaining high reliability. This requires soft magnetic materials with lower energy loss (for large transformer and motors) and hard magnetic materials that can form stronger permanent magnets (for small motors). In particular, new measurement capability must be able to cope with the large sizes of some components (like large power transformers) so that measurements can be made right on the production line. Present methods require cutting out and machining small samples of the bulk materials for measurement.

In general terms, the most important types of improvements needed are higher accuracy, accommodation of larger sample sizes, and accommodation of more diverse sample shapes. New measurement reference standards are needed to maintain the consistency of these measurement methods. In addition, some measured materials reference data will be needed.

Specific Measurement Needs

The specific magnetic quantities for which improved measurement capability is needed are summarized in Table 51 and are discussed individually below. There are critical needs for new measurement methods, for new measurement reference standards to support the accuracy of the measurement methods, and for new measured reference materials data on the properties of key magnetic materials.

Referring to Table 50, the types of measurement improvements needed for different quantities may differ. Even for a given quantity, the improvements needed may differ from one application to another. The discussion below associates the improvements with specific applications where appropriate. On the other hand, some of the needed improvements share common characteristics. This commonality arises principally from the close relationships among some of the quantities. Those relationships are also be noted below. One common factor among many magnetic measurements made in very small or very fast systems is the increased care that must be taken to accommodate both thermal and temporal effects, even when using very good measurement methods.

Table 51
Magnetic Measurement Needs of U.S. Industry

Quantity	Measurement Method	Measurement Reference Standard	Materials Reference Data
magnetic flux density	✓		
magnetization	✓	✓	✓
permeability	✓	✓	✓
coercivity	✓	✓	
magnetoresistance	✓		
magnetostriction	✓	✓	✓
Kerr rotation		✓	✓
hysteresis loss	✓	✓	
energy product	✓	✓	✓

Magnetic Flux Density

For magnetic information storage, this quantity is measured to probe both the interfaces between domains or particles of the media and the magnetic structure of the reading head. These characteristics determine the nature of the signals presented to the reading head and the degree of interference between successive recorded bits. The main improvements needed are *finer spatial resolution* and *higher sensitivity*. Magnetic force microscopy (MFM) is a promising technique that requires further development to fill this need with the required resolution. A further improvement required is *higher frequency* capability to support high-speed recording techniques at speeds above 150 Mbit/s and possibly to 1 Gbit/s in the future. Such data rates require measurement methods with capability to about 10 gigahertz.

Magnetization

For magnetic information storage, the strength of the magnetized domains on recorded media determines the amplitude of the signals presented to the reading head. Higher amplitudes reduce the ability of random noise to produce errors in data read from the media and increase the speed with which the data can be read. Magnetization must be measured with *higher spatial resolution*. A new technique -- scanning electron microscopy with polarization analysis (SEMPA) -- is being developed and commercialized to meet this need. Its resolution must be improved about a factor of five to reach 10 nanometers.

A major challenge in magnetization measurement will be to develop techniques for measuring the *anisotropy* of the magnetization to support a variety of recording techniques ranging from those employing soft films in conventional and magnetoresistive heads to those designed for vertical recording media. Development of more sophisticated torque magnetometers can help in this area. Finally, improved methods for determining the effects of *sample shape* on the measured values of magnetization are also required.[52] For electrical equipment, improved measurement methods are needed for the saturation magnetization. It determines the largest magnetic response that a magnetic material can provide. Materials with a high saturation magnetization enable the design of more compact electrical equipment. The principal measurements needs are for on-line measurement methods applicable to *large components* and for measurement reference standards to assure the *accuracy* of those measurements. A major challenge is the measurement of the *anisotropy* of the magnetization in large components. *Anisotropy* is important in grain-oriented soft ferromagnetic steels used in power transformer cores. They are designed to maximize the magnetization in a preferred direction. For electronic components, improved measurement methods and reference standards are needed for magnetization measurements of *higher accuracy*; measurement reference standards will be required. Special attention needs to be given to measurement methods for *diverse sample shapes* and for *higher frequencies* to support the components.

Permeability

Permeability is a measure of how easily a magnetic material enhances the strength of an applied magnetic field. For electrical equipment, the principal need is for measurement methods applicable to *large components*. Also needed is selected reference materials data on the most important materials. The data are needed to support the design of products and manufacturing processes. For electronic components, permeability measurements are needed for characterizing both soft ferrimagnetic and soft ferromagnetic materials. In particular, the new measurement methods are needed to determine the homogeneity of the permeability with a *finer spatial resolution* down to about 100 micrometers. Also, new measurement methods are needed to accommodate more *diverse sample shapes*. This capability is especially important when measuring materials with high levels of permeability. The high levels complicate accommodating shape effects, especially when permeability is measured as a function of temperature. Further, measurement reference standards are needed to assure the accuracy of permeability measurements in ferrimagnetic materials over a broad frequency range of 50 megahertz to at least 25 gigahertz, and perhaps as high as 80 gigahertz. The high frequency end of this range is important for magnetic materials used in microwave systems. In addition, measurement reference standards may be needed to support the use of thin films at frequencies up to 1 gigahertz for use in high-speed recording systems.

Coercivity

For magnetic information storage, the coercivity of the medium determines the field that must be applied by the recording head to write or erase recorded data. The coercivity also determines how closely bits of data can be packed in recorded media before their opposing fields start mutually erasing each other. Coercivity is understandably a prime quantity for specifying the media. Recall that a recent interlaboratory comparison of present measurement methods showed that the methods lack adequate accuracy. Understanding time-dependent effects in recording media is of prime importance to the development of measurement techniques and standards. Coercivity measurements are important, more generally, for determining how magnetically soft or hard materials are. Thus coercivity measurements are important for developing very soft magnetic materials for transformer cores and very hard magnetic materials for permanent magnets. The range is so large (a factor of 4 million) that one would expect to need a number of reference standards and methods to cover it. The principal measurement need is for improved *accuracy*, achieved principally through the development of measurement reference standards. However, some improvements in the measurement methods will be needed as well.

Magnetoresistance

For magnetic information storage, the next generation of reading heads will probably be based on magnetoresistance. In the very small structures of reading heads, the magnetoresistance is affected strongly by the scattering of electrons at the boundaries of the material, and hence is dependent on size, shape, and magnetic structure. It is therefore important to measure the properties of materials under conditions that are realistically related to the recording process. New test structures are required to support such measurements. Also needed are new techniques capable of producing highly confined fields suitable for measurements with *finer spatial resolution*. As with magnetization and magnetic flux density measurements, measurements at *higher frequencies* are needed to support high-speed recording techniques.

Magnetostriction

The term magnetostriction has historically referred to the physical contraction and expansion of magnetic materials in response to magnetic fields. These effects have been principally important for transducers that convert variations in electric current into physical motion for applications in high-frequency audio speakers and sonar. Present measurement methods for these effects are adequate, but measurement reference standards are needed to assure the *accuracy* of the methods; and materials reference data are needed for the properties of the most important materials.

More recent usage of the term magnetostriction includes also the inverse effects, that is, the effects of elastic distortions on the magnetic properties of materials. In magnetic recording, these inverse effects can seriously degrade head performance. Increasing awareness of the importance of these inverse effects is likely to give rise to the need for new measurement methods and supporting measurement reference standards in the near future.

Kerr Rotation

For magneto-optical recording, where Kerr rotation is used to read data, accurate measurement of the saturation (maximum) Kerr rotation is essential. Present measurement methods appear adequate, but their *accuracy* needs to be established and assured by the development and use of a set of measurement reference standards. Special attention needs to be paid to thin film measurements where multiple reflections can create confusion.

Hysteresis Loss

This quantity determines the electrical power dissipated by magnetic materials. It is especially important in electrical equipment, such as power transformers, but also leads to noise in recording systems. Also, the high frequencies used in switching power supplies require that special attention be paid to domain motion and domain rotation losses in the high-frequency transformers. Hysteresis loss is not an independent quantity; a good ability to measure magnetic flux density as a function of magnetic field enables measurement of hysteresis loss. The primary measurement need is for measurement methods applicable to components too large and too small for conventional magnetometers.

Energy Product

Energy product is an important measure of the performance of permanent magnets. This is not a quantity independent of the others above. As with hysteresis loss, a good ability to measure magnetic flux density as a function of magnetic field enables determination of the energy product. In essence, energy product reflects both the coercivity and the remanence of a permanent magnet. High values of both are desirable in permanent magnets. Present measurement methods are adequate for permanent magnets, but measurement reference standards are needed to assure their *accuracy*. Further, materials reference data are needed for important types of permanent-magnet materials to support the design of products employing those materials.

Tape and Flexible-Disk Measurement Reference Standards

In addition to the needs described above for measurements of specific magnetic quantities, there is a continuing need for special measurement reference standards in the form of computer tapes and flexible disks. These do not address the accuracy of measurements of fundamental magnetic quantities, as do the other measurement reference standards discussed above. Rather, they are comparative standards designed to support determination of practical performance characteristics[53] of the media relative to mean values of performance for products of similar type in the marketplace. These special measurement reference standards are used by industry to verify that the properties of commercial tape and disk media conform to the specifications contained in voluntary industry standards. The voluntary industry standards are designed to assure the proper functioning of magnetic media on commercial recording equipment. The emergence of new recording technologies (helical scanning) and new magnetic materials (metal particles and films) for information-storage systems will require new measurement reference standards in the form of media. However, if sufficient progress can be made in developing improved measurement capability for the fundamental magnetic quantities cited above, then the need for a wide range of standards in the form of media will abate, although standards for measurements of system performance will continue to be needed.

ENDNOTES

1. Richard P. Reed and Alan F. Clark, ed., *Materials at Low Temperatures*, American Society for Metals, "Chapter 6: Magnetic Properties" by F.R. Fickett and R. B. Goldfarb, pp. 203 to 214 (1983). The permeability values were developed from information in this reference. The permeability values are useful for magnetic materials with a linear or nearly linear relationship between the magnetic flux density and the applied magnetic field. Diamagnetic, paramagnetic, antiferromagnetic, soft ferromagnetic, and soft ferrimagnetic materials meet this criterion. *Soft* as explained later in the text means *low coercivity*. Hard ferromagnetic and ferrimagnetic materials exhibit a more complicated relationship that is not readily described by a permeability value. The table actually shows the *relative permeability*. This is simply the permeability relative to free space, that is, the permeability relative to the case in which no material is present. The relative permeability of free space itself is 1.

2. Karl J. Strnat, "Modern Permanent Magnets for Applications in Electro-Technology", *Proceedings of the IEEE*, Vol. 78, No. 6, p. 928 (June 1990).

3. Karl J. Strnat, "Modern Permanent Magnets for Applications in Electro-Technology", *Proceedings of the IEEE*, Vol. 78, No. 6, p. 936 (June 1990).

4. Roger Wood, "Magnetic Megabits", *IEEE Spectrum*, p. 38 (May 1990).

5. Roger Wood, "Magnetic Megabits", *IEEE Spectrum*, pp. 33 and 38 (May 1990).

6. M. Futamoto, F. Kugiya, M. Suzuki, H. Takano, H. Fukuoka, Y. Matsuda, N. Inaba, T. Takagaki, Y. Miyamura, K. Akagi, T. Nakao, H. Sawaguchi, and T. Munemoto, "Demonstration of 2 Gb/in^2 Magnetic Recording at a Track Density of 17 KTPI", presented at the *5th Joint MMM-Intermag Conference* in Pittsburgh, Pennsylvania, p. 301 of conference abstracts (June 18-21, 1991). The technique reported employs magnetic recording but optical tracking.

7. James U. Lemke, "Magnetic Storage: Principles and Trends", *MRS Bulletin* of the Materials Research Society, p. 33 (March 1990). This is known as the *superparamagnetic limit*. It is reached when the magnetic energy of the domains becomes comparable to the thermal energy so that normal thermal agitation washes out the magnetic domains.

8. Roger Wood, "Magnetic Megabits", *IEEE Spectrum*, p. 38 (May 1990).

9. Sony HDD-1000 recorder, now a commercial product. Information provided by Sony Corporation.

10. *High Definition Digital VTR System*, Sony Corporation, the product data sheet for HDD-1000 Digital VTR (year not specified, but probably 1989).

11. *1990 Disk/Trend Report: Rigid Disk Drives*, Disk/Trend, Inc., p. SUM-34 (October 1990).

12. J. Lemke, Visqus Corporation, "Rigid Disks at 1 Microinch with Inductive Heads", presented at the *5th Joint MMM-Intermag Conference* in Pittsburgh, Pennsylvania, p. 260 of conference abstracts (June 18-21, 1991).

13. The information density per unit area is the product of the linear density of data along each circumferential "track" of data on the disk and the spacing between adjacent tracks. Typical linear density along the tracks is about the same in magnetic and magneto-optical recording (25,000 bits per inch, or 10,000 bits per

centimeter), although high-end products for magnetic recording tend to go somewhat higher (50,000 bits per inch or 20,000 bits per centimeter) than high-end products for optical recording (30,000 bits per inch or 12,000 bits per centimeter). Typical track-to-track spacing, however, is about ten times higher for magneto-optical recording (15,000 tracks per inch, or 6000 tracks per centimeter) than for magnetic recording (1500 tracks per inch, or 600 tracks per centimeter). Thus, the typical information density per unit area is about 60 $\mathrm{Mbit/cm^2}$ for magneto-optical recording and about 6 $\mathrm{Mbit/cm^2}$ for magnetic recording. These observations are based on product data in the *SPEC* pages in the *1990 Disk/Trend Report: Optical Disk Drives*, Disk/Trend, Inc. (July 1990) and the *1990 Disk/Trend Report: Rigid Disk Drives*, Disk/Trend, Inc. (October 1990). On the other hand, it is really volume density that counts in mass storage devices of this type and the much smaller heads in magnetic drives allow closer disk spacing in multi-disk drives that allowed by magneto-optical systems.

14. Laser diodes in optical storage systems operate at a wavelength of about 0.7×10^{-4} centimeters. If each recorded bit measured one wavelength on a side, it would occupy 0.5×10^{-8} square centimeters; and the resulting maximum bit density would be 2×10^{8} per square centimeter, or 200 $\mathrm{Mbit/cm^2}$. Commercially available magneto-optical drives achieve about 70 $\mathrm{Mbit/cm^2}$.

15. Roger Wood, "Magnetic Megabits", *IEEE Spectrum*, p. 33 (May 1990).

16. Most magneto-optical drives have data transfer rates falling in the range of 1.2 to 12 Mbit/s, with the vast majority well under 8 Mbit/s. In comparison, most magnetic drives have data transfer rates in the range of 5 to 24 Mbit/s with a heavy presence throughout the range of 8 to 24 Mbit/s. Most magnetic drives have head access times in the relatively narrow range of 20 to 40 milliseconds, while most magneto-optical drives have head access times in the relatively wide range of 70 to 700 milliseconds, with common values from 110 to 150 milliseconds. These observations are based on product data in the *SPEC* pages in the *1990 Disk/Trend Report: Optical Disk Drives*, Disk/Trend, Inc. (July 1990) and the *1990 Disk/Trend Report: Rigid Disk Drives*, Disk/Trend, Inc. (October 1990).

17. Karl J. Strnat, "Modern Permanent Magnets for Applications in Electro-Technology", *Proceedings of the IEEE*, Vol. 78, No. 6, pp. 936 and 940 (June 1990).

18. Silicon was placed in iron to increase its electrical resistance so that it could be used without excessive electrical losses from the currents induced by the alternating magnetic fields employed.

19. Amorphous magnetic materials are those without the high level of atomic order that characterizes crystalline materials. However, when heated sufficiently, amorphous materials crystallize. This behavior limits the temperatures at which they can be used.

20. *U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 37-5 (January 1990).

21. Karl J. Strnat, "Modern Permanent Magnets for Applications in Electro-Technology", *Proceedings of the IEEE*, Vol. 78, No. 6, pp. 935-936 (June 1990).

22. Superconducting magnets presently generate fields up to about 15 tesla, while magnets employing magnetic materials are limited to about 2.5 tesla where the best ferromagnetic materials saturate.

23. Earl Callen, "Japan, Magnetism, and Technological Competition", an abstract of a talk as described in *Journal of Applied Physics*, Vol. 69, No. 8, p. 5816 (April 15, 1991).

24. Masaru Yokokura, "Updated Japanese Market Situation in NdFeB Magnet Applications", a talk presented at the Gorham Advanced Materials Institute (February 26-28, 1989). Mr. Yokokura is from Sumitomo Special Metals America, Inc.

25. The superscripts in the table refer to the notes that follow: (a) Karl J. Strnat, "Modern Permanent Magnets for Applications in Electro-Technology", *Proceedings of the IEEE*, Vol. 78, No. 6, p. 926 (June 1990). (b) Masaru Yokokura, "Updated Japanese Market Situation in NdFeB Magnet Applications", a talk presented at the Gorham Advanced Materials Institute (February 26-28, 1989). Mr. Yokokura is from Sumitomo Special Metals America, Inc.

26. Masaru Yokokura, "Updated Japanese Market Situation in NdFeB Magnet Applications", a talk presented at the Gorham Advanced Materials Institute (February 26-28, 1989). Mr. Yokokura is from Sumitomo Special Metals America, Inc.

27. Karl J. Strnat, "Modern Permanent Magnets for Applications in Electro-Technology", *Proceedings of the IEEE*, Vol. 78, No. 6, p. 927 (June 1990).

28. Karl J. Strnat, "Modern Permanent Magnets for Applications in Electro-Technology", *Proceedings of the IEEE*, Vol. 78, No. 6, p. 927 (June 1990). This reference contains a 1988 level of $158 million (expressed as 9.0 percent of the world market of $1755 million). This is in very close agreement to the 1988 level of $162 million provided by Masaru Yokokura, "Updated Japanese Market Situation in NdFeB Magnet Applications", a talk presented at the Gorham Advanced Materials Institute (February 26-28, 1989). Mr. Yokokura is from Sumitomo Special Metals America, Inc. Mr. Yokokura also provided the 1990, 1995, and 2000 levels.

29. Masaru Yokokura, "Updated Japanese Market Situation in NdFeB Magnet Applications", a talk presented at the Gorham Advanced Materials Institute (February 26-28, 1989). Mr. Yokokura is from Sumitomo Special Metals America, Inc.

30. The superscripts in the table refer to the notes that follow: (a) The U.S. market was determined from U.S. shipments by subtracting the fraction exported (64 percent) and adding the amount imported (excess of imports over exports, or $1.6 billion, plus the amount exported). *1991 Electronic Market Data Book*, Electronic Industries Association, pp. 83-84 (1991). (b) Virtually all relays rely on magnetic materials. "Switchgear, Switchboard Apparatus, Relays, and Industrial Controls, *Current Industrial Reports*, MA36A(89)-1, Bureau of the Census, U.S. Department of Commerce, p. 6 (September 1990). *1991 Electronic Market Data Book*, p. 68. (c) The 1989 U.S. market has been used as an estimator for the 1990 level. "Radio and Television Receivers, Phonographs, and Related Equipment", *Current Industrial Reports*, MA36M(89)-1, pp. 3-4 (October 1990). (d) and (e) *1991 Electronic Market Data Book*, pp. 70-73. (f) "Semiconductors, Printed Circuit Boards, and Other Electronic Components", *Current Industrial Reports*, MA36Q(89)-1, Bureau of Census, U.S. Department of Commerce, p. 6 (September 1990).

31. The superscripts in the table refer to the notes that follow: (a) The data reflect revenues at the point of first public sale and include both captive and merchant markets. *1990 Disk/Trend Report: Rigid Disk Drives*, Disk/Trend, Inc., p. SUM-3 (October 1990). (b) The data again reflect revenues at the point of first public sale and include both captive and merchant markets. *1990 Disk/Trend Report: Flexible Disk Drives*, Disk/Trend, Inc., p. SUM-3 (November 1990). (c) These numbers are based on factory revenues at the point of first sale but associated with the region of consumption and include both captive and merchant markets. The figure for the U.S. market is 90% of the Dataquest estimate for the North American market, which Dataquest (in response to a telephoned question) estimates as the U.S. share of the North American market for magnetic recording products generally. *Market Statistics -- Tape Drives, Dataquest*, pp. 2-3 (May 1991). Remaining data in the table, except as noted below, are based on the *1991 Electronic Market Data Book*, Electronic Industries Association (1991) from the pages cited. For (d) through (h), no world market data has been located, so the U.S. market data are repeated in the world market column to enable establishing a minimum total for the world market, since the world market will be greater than the U.S. market alone. (d) These two numbers are U.S. factory shipments and are used as estimators of the U.S. market in the absence of U.S. market data (p. 52). (e) p. 5. (f) p. 5. (g) Includes portable audio equipment ($1.64 billion dominated by units containing tape

players), plus 20% (undocumented estimate of tape-equipment content) of packaged or compact audio systems ($1.27 billion), of separate audio components ($1.94 billion), and of car audio equipment ($4.29 billion) (pp. 5, 19, and 21). (h) Blank audio and video cassettes (p. 5). Not shown in the table is commercial broadcasting and studio equipment for which the magnetic equipment component has not been separated. U.S. shipments of all audio and video equipment (including magnetic recording equipment) for broadcasting and studios was $0.96 billion in 1990 (p. 36).

32. *1990 Disk/Trend Report: Rigid Disk Drives*, Disk Trend, Inc., p. SUM-5 (October 1990).

33. Earl Callen, "Japan, Magnetism, and Technological Competition", an abstract of a talk as described in *Journal of Applied Physics*, Vol. 69, No. 8, p. 5816 (April 15, 1991).

34. *1990 Disk/Trend Report: Rigid Disk Drives*, Disk/Trend, Inc., p. SUM-9 (October 1990).

35. *1990 Disk/Trend Report: Flexible Disk Drives*, Disk/Trend, Inc., p. SUM-8 (November 1990). *1990 Disk/Trend Report: Rigid Disk Drives*, Disk/Trend, Inc., p. SUM-11 (October 1990).

36. *1990 Disk/Trend Report: Flexible Disk Drives*, Disk/Trend, Inc., p. SUM-5 (November 1990).

37. *Market Statistics -- Tape Drives*, Dataquest, p. 89 (May 1991).

38. *High Definition Digital VTR System*, Sony Corporation, the product data sheet for HDD-1000 Digital VTR (year not specified, but probably 1989).

39. The 7.9 percent market share comes from the *1990 Disk/Trend Report: Optical Disk Drives*, Disk/Trend, Inc. p. SUM-10 (July 1990). The $0.9 billion figure come from Disk/Trend, Inc., as quoted from its *1991 Disk/Trend Report: Optical Disk Drives*, p. SUM-14 (July 1991). The $0.3 billion figures was calculated by Disk/Trend for this document.

40. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 38-13 (January 1991).

41. The superscripts in the table refer to the notes that follow: (a) The U.S. market for transformers was calculated from U.S. shipments by adding imports and subtracting exports. *U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, pp. 21-1 to 21-2 (January 1991). (b) For motors and generators, export and import levels for 1990 were not available, so the values for 1989 were used as estimators of 1990 to calculate the 1990 U.S. market. *U.S. Industrial Outlook*, pp. 21-1 to 21-2.

42. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 21-4 (January 1991).

43. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, pp. 21-1 (January 1991). The 1990 figure for "Relays and Controls" of $6.6 billion has been reduced by the 1989 value of U.S. shipments of relays, $0.8 billion, as the best available estimator of the 1990 level, to provide an estimate of U.S. shipments of controls alone. The U.S. market for relays has been accounted for separately under electronic components. The figure for U.S. shipments of relays for 1989 comes from "Switchgear, Switchboard Apparatus, Relays, and Industrial Controls", *Current Industrial Reports*, Bureau of the Census, U.S. Department of Commerce, p. 6 (September 1990).

44. Preliminary estimate for 1990 from the Edison Electric Institute, as of December 9, 1991.

45. *Annual Energy Review 1990*, Energy Information Administration, U.S. Department of Energy, p. 313 (1990).

46. Core losses are particularly significant relative to other types of losses, such as losses due to resistance in the windings of the transformers, for transformers which operate at low loads much of the time. Such transformers include the distribution transformers so common on every street in the nation.

47. *Electric Motors and Drives: Markets, Trends and Opportunities: Phase 1 - Motors*, Resource Dynamics Corporation, p. 4 (January 1991 for "Draft Final" version). This report was prepared for the Electric Power Research Institute.

48. J. Unguris, M. R. Scheinfein, R. J. Celotta, and D. T. Pierce, "Scanning Electron Microscopy with Polarization Analysis: Studies of Magnetic Microstructure", Chapter 11 in *Chemistry and Physics of Solid Surfaces VIII*, Ralf Vanselow and Russell Howe, editors, pp. 240-265 (Heidelberg, Germany: Springer-Verlag; 1990).

49. D. Rugar, H. J. Mamin, P. Guethner, S. E. Lambert, J. E. Stern, I. McFadyen, and T. Yogi, "Magnetic Force Microscopy: General Principles and Application to Longitudinal Recording Media", *Journal of Applied Physics*, Vol. 68, No. 3, pp. 1169-1183 (August 1, 1991).

50. Samuel D. Bader, "Thin Film Magnetism", *Proceedings of the IEEE*, Vol. 78, No. 6, pp. 909-922 (June 1990).

51. F.R. Fickett, *Magnetic Measurements, Calibrations, and Standards: Report on a Survey*, NBSIR 84-3018 (October 1984).

52. The effects of sample shape are commonly dealt with through the use of so-called demagnetizing factors. These factors have historically been calculated for only very limited geometries. NIST recently determined the factors for a wide range of cylindrical shapes (in the reference following). Many additional shapes remain to be addressed, including cubes, ribbons, powders, etc. The calculations are complex. D-X Chen, J. A. Brug, and R. B. Goldfarb, "Demagnetizing Factors for Cylinders", *IEEE Transactions on Magnetics*, Vol. 27, pp. 3601-3619 (1991).

53. The tape characteristics supported by the standard reference materials for tape are *typical current* (the current required to reach a specified point on the tape's saturation curve), *signal amplitude* (the read signal strength resulting from writing with a specified level of current), *resolution* (the information density of the tape), *overwrite* (the degree to which a previously written signal is successfully erased by a subsequent write to the same location on the tape), and *peak shift* (the apparent movement of the location of recorded bits at junctions between regions on the tape with different bit patterns).

CHAPTER 6

SUPERCONDUCTORS

May 1991

Chapter 6

SUPERCONDUCTORS

SUMMARY

The technology of low-temperature superconductivity (LTS) has been under serious development since the discoveries of the high-field superconductors in the late 1950s and the Josephson effect in the early 1960s. It has already produced major engineering successes, among them the very fast 4-bit microprocessors developed by Fujitsu and Hitachi, the Josephson-junction-array voltage measuring system developed by NIST, and the many large magnets for bubble chambers, for controlled fusion experiments, for accelerators such as the Tevatron at the Fermi Laboratory, for the experimental magnetically levitated train developed by the Japan National Railways, and for magnetic resonance imaging equipment widely used in medicine.

These and several other very promising applications of superconductivity have all been studied and developed with low-temperature superconductors. The discovery of high-temperature superconductors has stimulated much new exploration. All these ideas are receiving fresh attention, and new ones are being added. The future of this field is difficult to predict but holds great promise.

The growth of the superconductor market forecast by World Business Publications predicts a total of \$1.3 billion by 2003. This represents the market for both low-temperature superconductors, which account for all the present market, and high-temperature superconductors, which are expected to dominate the market in the later years. Note that already in 1988 there was a market of \$172 million for low-temperature superconductor projects.

Japanese forecasters take a generally more optimistic view than those in the U.S. The Nomura Research Institute forecasts a total world market of \$43 billion in 2000 and \$123 billion in 2010. This view is encouraged by the higher level of industrial research on superconductivity in Japan than in the U.S.

The main barrier to the creation of a large market for LTS is that it requires liquid-helium refrigeration, which is expensive and widely distrusted, largely because it has never been developed on a commercial scale. High-temperature superconductivity (HTS) promises to eliminate this barrier by substituting at least liquid-nitrogen refrigeration, which is cheap and well established on a commercial scale. This would then reduce the overall cost of the technology if other costs (such as that of the conductors themselves) do not turn out to be greater.

The basic components of a practical HTS technology do not exist yet, but steady progress is being made in their development. The current-carrying capacity of thin films, suitable for integrated circuits, has reached a useful level. The current-carrying capacity of bulk material, suitable for making wires, is still 10 to 100 times smaller than desired for some practical applications at low magnetic fields, and deteriorates further in high magnetic fields. Single-layer microcircuits can be fabricated by methods similar to those used for semiconductor "chips", and the first multilayer circuits

have been reported. A large effort has not yet produced a controllable technique to fabricate active electronic devices, but several prototypes have been demonstrated in the laboratory.

All the remaining problems may be overcome by clever ideas, just as similarly difficult problems in LTS were overcome. But they create great uncertainty in the future of HTS technology, especially in the time scale for practical, commercial development. A prudent plan is to continue the development of LTS technology in parallel with early efforts in HTS technology until the latter demonstrates clear superiority in each application.

The U.S. is a formidable competitor in the developing superconductor market. The U.S. has the advantages of experience gained from several large magnet projects in the Department of Energy and a large domestic military market that will accept new products of high performance with less regard to cost than the commercial market at large. Japan has the advantage of national cooperative programs with stable, long-term funding that the U.S. is attempting to match with experiments in organization of industrial consortia.

Success in developing superconductor technology and in realizing competitive products from that technology will require continued progress in the supporting measurement capability. Many of the unique measurement problems of superconductor technology have already been encountered in the development of LTS, and good progress has been made with their solution. HTS brings many of the same problems, but often in a modified or exaggerated form. It also has unique problems of its own.

Some of the measurement problems have been discussed at length at conferences, workshops, and press conferences in the U.S., and also by international committees such as the Versailles Project on Advanced Materials and Standards (VAMAS) and the International Electrotechnical Commission (IEC). Physical evidence of inadequate measurement techniques and standards has been demonstrated by "round robin" intercomparison of measurements among laboratories.

The principal areas in which improved measurement methods and standards, or improved measured reference data, are needed are these:

(1) electrical and magnetic characteristics
(2) composition and structure analysis
(3) reference data for design and control of fabrication processes

INTRODUCTION TO SUPERCONDUCTIVITY

Zero Resistance

Superconductors are a class of materials that make a transition at a well defined temperature (known as the critical temperature), below which they enter a state in which they conduct electric current with no resistance. This phenomenon raises the possibility of constructing electrical machines and electronic devices with very high performance because of the removal of this major source of energy dissipation. At a temperature above the critical temperature, superconducting materials have no striking common characteristics.

The vanishing of electrical resistance below the critical temperature is strictly true only for direct current. There is a dissipation of energy from alternating current that depends upon the condition of the material (particularly its surface), but can be very small, even at frequencies up to the microwave range. There is also an upper limit of current (the critical current) and magnetic field (the critical field), above which the material reverts to the normal state. The critical current and the critical field depend strongly on the temperature, and are also affected by mechanical strain.

Superconductors come in a variety of forms with a wide range of critical temperatures. Some superconductors are metals or alloys, some are ceramics, and some are carbon-cluster compounds called *fullerenes*, which are doped C_{60} and its relatives. The first superconductors to be discovered were metals or alloys with low critical temperatures of less than 23 kelvins or -250°C. The new class of ceramic superconductors was discovered with critical temperatures up to 130 kelvins or -143°C. This discovery has stimulated a resurgence of interest in superconducting technology. The reason for this is that it allows liquid nitrogen to be used as a refrigerant in place of the liquid helium that must be used with low-temperature superconductors. This makes a profound difference, partly because liquid-nitrogen refrigeration is well established on a commercial scale (liquid-helium refrigeration is not), and partly because of the inherent advantages of nitrogen shown in Table 52.

Table 52
Comparison of Coolants

	Liquid Helium	Liquid Nitrogen
temperature	4.2 kelvins	77 kelvins
cost per liter	$5.00	$0.20
availability	limited (natural gas wells)	unlimited (air)
handling	difficult	easy
cooling power	low	high (70 times liquid helium)

Meissner Effect

Another aspect of the phenomenon of superconductivity is the Meissner Effect. On entering the superconducting state in the presence of a magnetic field, a superconductor generates spontaneous circulating currents that expel the field from its interior. In an ideal case expulsion would be complete, but in practice only partial expulsion is usually observed. The Meissner Effect is a sure indicator of superconductivity, and is the basis of the popular floating magnet demonstration.

Flux Quantization

A related phenomenon is the quantization of magnetic flux. The magnetic flux linking a superconducting loop is quantized in units of h/2e (h is Planck's constant and e is the charge on the electron). This is about $2x10^{-15}$ weber. This phenomenon affects the manner of penetration of a magnetic field into a superconductor. In most practical superconductors, magnetic field penetrates in the form of densely packed parallel threads of magnetic flux. Each contains a single flux quantum and is separated from its neighbors by circulating electric currents. When in addition the superconductor carries a transport current, it exerts a strong force (the Lorentz force) on these threads of magnetic flux. If they move, they dissipate energy, and resistance to the passage of current appears. This condition is called flux flow, and it happens most readily in a material with a perfectly regular crystal structure. Irregularities in the structure (such as variations in composition, dislocations, grain boundaries, etc.) "pin" the threads of magnetic flux in place and thereby suppress flux flow. Creating a superconductor that will carry a high current, without dissipation, in the presence of a high magnetic field depends upon finding a process to damage the crystal structure in a way that is particularly effective at pinning magnetic flux. The strength of flux pinning is the major factor limiting the critical current of a superconductor at very high magnetic fields. Another major factor limiting the critical current of ceramic superconductors is the electrical contact at the interfaces between the grains of the material. This also is affected by magnetic field.

Magnetic flux quantization also affects the distribution of currents in superconducting electronic circuits, and sets a lower limit to the signal levels that can be handled by a broad class of superconducting devices. This limit is realized when a single flux quantum is exchanged from one part of a circuit to another as a result of a switching operation. This is known as single-flux-quantum logic, and it allows devices to be designed with surprisingly robust operation at very low signal levels.

Josephson Effect

Many applications of superconductivity to electronics depend upon the Josephson effect. This effect takes place at a weak contact (a Josephson junction) between two superconductors. A Josephson junction has two modes of conducting electric current. One is without dissipation (zero voltage). In the other a voltage appears and part of the current oscillates at a frequency f proportional to the voltage V according to the relationship $f = 2eV/h$. A Josephson junction can be made to switch from one mode to the other. When Josephson junctions are incorporated in superconducting circuits, their electrical characteristics are affected also by magnetic flux. Typically the effect is periodic, with period equal to the flux quantum. This has led to the creation of the Superconducting Quantum Interference Device (SQUID), which is basically a very sensitive magnetic detector. Josephson junctions are the basis for a host of ingenious devices that depend on their unique electrical characteristics.

Persistent Currents

Many applications of superconductivity use electromagnets wound from superconducting wire. They are capable of generating very high magnetic fields with relatively low currents and very low power requirements, compared to magnets wound from conventional conductors, like copper. Typical high-field superconducting magnets can be operated from power supplies with kilowatt power levels. In comparison, conventional magnets made from copper require megawatt power levels to reach comparable fields. Conventional magnets also require elaborate water cooling systems to keep them from melting from the heat generated by the resistance in their copper windings. For some applications, superconducting magnets can be operated in an even more efficient mode that takes advantage of their ability to sustain "persistent currents". To establish this mode, the field of the magnet is raised to the desired value by increasing the current delivered by the power supply to the magnet’s windings. Then the incoming and outgoing wires of the magnet are connected together by a superconducting switch (no dissipation). Once the switch has been closed, the current in the magnet continues to circulate, or persist, at a constant level without additional power input from the power supply and without heat dissipation in the magnet's windings. Once in this mode, the magnet produces a remarkably stable magnetic field.

APPLICATIONS OF SUPERCONDUCTORS

The technology of the low-temperature superconductivity of metals and alloys has been under serious development since the discoveries of the high-field superconductors in the late 1950s and the Josephson effect in the early 1960s. Successful electronic applications and successful large-scale applications have both been realized. Most electronic applications rely on Josephson junctions as the active elements, while most large-scale applications rely on superconducting magnets. Among the many successful large-scale applications, all of which employ superconducting magnets, have been bubble chambers, controlled-fusion experiments, accelerators including the Tevatron at the Fermi Laboratory, an experimental magnetically levitated train developed by the Japan National Railways, and magnetic resonance imaging (MRI) equipment. Among the successful electronic applications have been very fast 4-bit microprocessors developed by Fujitsu and Hitachi and the Josephson-junction-array voltage-measuring system developed by NIST. The broader spectrum of potential large-scale applications and electronic applications is reflected in Table 53 and Table 54, respectively. Most of these applications are discussed in detail below.

Table 53
Large-Scale Applications of Superconductors

Application	Example
research	
accelerators	Tevatron, Superconducting Super Collider
bubble chamber	Argonne facility
laboratory magnets	MIT National Magnet Laboratory
medicine	magnetic resonance imaging
transportation	
rotating motors	ship drives
levitation magnets	trains
linear electric motors	trains
direct sea water propulsion	ship propulsion
energy	
generators	rotating and magnetohydrodynamic
storage	magnetic energy storage coils
transmission	electric power transmission lines

The major commercial success to date is magnetic resonance imaging, with a market predicted to exceed $1 billion for 1990.[1] MRI is a medical diagnostic technique that requires a strong, highly uniform, and controllable magnetic field in a volume sufficient to accommodate the body of the subject. This and other medical applications of superconductivity have great market potential. Among the most promising is magnetoencephalography, or mapping of the electric current distribution in the brain using an array of SQUID magnetometers. This technique is at present in the medical research phase, and is being shown to be capable of providing diagnostic information on the brain that is not accessible by any other means.

The development of high-energy particle accelerators has led to the construction of many large and successful superconducting bending and focusing magnets, and to the solution of major engineering problems associated with the stability of large superconducting systems. A fine example is the Tevatron, at the Fermi Laboratory near Chicago, which features a ring of superconducting magnets 5 kilometers in circumference. It works as designed. The Superconducting Super Collider (SSC) will be a similar machine but ten times bigger. Liquid-hydrogen bubble chambers are used as detectors in high-energy physics experiments. They also require high magnetic fields in large volumes, to identify charged particles and measure momentum. Perhaps the most significant potential application of large superconducting magnets is to controlled thermonuclear fusion. Some very large magnets have already been constructed for experiments in magnetic confinement, and it is generally recognized that performance that can be achieved only by a superconducting magnet would be essential to the success of a magnetically confined fusion reactor.

In the electric power industry, superconducting magnetic energy storage (SMES) is seen as the most immediate application of superconductivity. Energy storage has great significance in the economics of electric power systems, which must cope with a demand that has large periodic variations. SMES are now being marketed to companies as power backups to compensate for voltage sags and short-term power failures. The supreme advantage of SMES is that is does not require energy conversion (except for the conversion of electricity from AC to DC) and can therefore be very efficient. Flywheel energy storage using HTS bearings now appears practical and is being discussed. There have been many studies of power transmission and generation with superconductors. A superconducting transmission line would indeed be capable of costing less energy than a conventional one, but the economic advantage would be largely offset by the need to lay it underground, instead of overhead which is cheaper. The main advantage of superconductors for generation is that superconducting windings would enable the machines to be made about half size for the same power rating. This would raise the limit on large machines, which is set by the maximum size that can be transported from factory to power station. This limit is an important factor in the economics of power generation.

The most exciting prospect in transportation is the magnetically levitated train (MAGLEV). An experimental train developed by the Japan National Railways is capable of travelling at 500 kilometers per hour, levitated by superconducting magnets (carried on board) above a special aluminum track. The drive is by linear induction motor. This train has carried passengers, and development is continuing with the objective of placing MAGLEV trains in commercial service.

Ship propulsion is also a very promising field. A compact electric drive that eliminates the constraints of accommodating a straight propeller shaft would permit the design of a more compact, efficient hull for a given payload. The benefit would then be compounded by needing to carry less fuel. The U.S. Navy has a program to develop a superconducting ship drive that has already produced a small prototype version for sea trials. Another possibility that is under investigation in Japan is to eliminate the propeller and use instead the magnetohydrodynamic (MHD) effect to impel seawater directly.

In electronics, a wide variety of applications is in the offing, as suggested in Table 54; and several have already been realized commercially. The market is mostly for scientific instrumentation in applications where great sensitivity is required. There is a modest commercial market of $5-10 million for magnetic measuring instruments based on the SQUID. One of these is magnetic anomaly detection (MAD), a military surveillance technique. Another commercial instrument is a sampling oscilloscope with 5-ps resolution. This outperforms all other electrical sampling systems by almost an order of magnitude, but the market for it appears to be small at present.

Table 54
Applications of Superconductor Electronics

Application	Example
computer circuits	microprocessors, digital-signal processors
measurement devices	magnetic sensors, voltage standards, sampling oscilloscopes
radio receivers	deep space listening radio telescope, compact antennas, filters
radiation detectors	infrared detectors, radiometers
other electronic circuits	counters, analog-to-digital converters
interconnections	pulse preserving transmission lines

Some of the unique features of superconductivity that can be applied with great advantage to advanced electronics are listed in Table 55. In applications where the ultimate performance is of paramount importance, the advantages offered can outweigh the trouble and expense of refrigeration.

Table 55
Special Capabilities of Superconductor Electronics

- high sensitivity
- fast switching speed
- low energy consumption
- good pulse-shape preservation
- frequency conversion capabilities
  - voltage-to-frequency conversion
  - frequency-to-voltage conversion
- magnetic flux quantization
- direct relationship to fundamental physical quantities

Several measuring devices under development at NIST offer commercial potential. These include a 30-bit binary counter that is capable of counting at rates up to 100 gigahertz, an analog-to-digital converter that can operate at rates in the gigahertz range, and a voltage standard that is part of an automated voltage measuring system that can refer voltage measurements in the range up to 12 volts directly to the international definition of the volt, with uncertainty of a few tens of parts per billion. The microcircuit at the heart of this system contains 19,000 Josephson junctions (the active elements of superconducting electronics). This is among the most complex superconducting microcircuits that have been achieved so far.

The application of superconducting electronics of the greatest commercial potential is to computers. The speed of computers will be limited ultimately by the delay in communication among the components. Achieving the highest possible computing speed will require packing a very large number of components very close together. This will most probably be limited by the need to dissipate the heat they generate, so one can define a figure of merit that depends on a combination of short switching speed and small energy dissipation on switching. By this measure superconducting devices are superior to all others, partly by virtue of operating at low temperature. Also, superconducting transmission lines may contribute to reducing communication delays. An ambitious project at IBM to develop a commercial superconducting computer was dropped because it fell behind the marketing schedule. Since then the work has continued in Japan, and recently both Fujitsu and Hitachi have announced 4-bit microprocessors with clock rates over 1 gigahertz. These are experimental devices, but the promise is great.

Another fruitful field for the application of superconductivity is radiation detection. Compact but efficient antennas can be made with superconductors that have ohmic resistance that is small compared to the radiation resistance. These are simple, one-layer devices that may provide an early field of application for high-temperature superconductors. Narrow-band filters have similar promise for early development. Millimeter-wave mixers with very low noise and good conversion efficiency are under development by a joint project between the University of California at Berkeley and NIST. The Kinetic Inductance Bolometer is also under development at NIST, both for a power standard and for a radiation detector. It is unique among radiation detectors in that it generates no thermal noise internally. Other ingenious devices are under development for radiation detection in all parts of the spectrum and for signal processing.

The next generation of superconducting electronics is visible on the horizon. The devices and circuits that have been attempted so far use the macroscopic characteristics of superconductors and Josephson junctions for active elements and are fabricated to design rules similar to those used for semiconductor microcircuits (indeed, the same equipment is used). The first step towards taking full advantage of the unique characteristics of superconductors would be to use single flux quantum logic (SFQL) which operates at extremely low signal levels. An example that has been demonstrated at NIST is a binary counter that uses SFQL and is capable of counting pulses at the attojoule (10^{-18} joule) level at rates up to 100 gigahertz. SFQL is likely to become important for smart sensors and instrumentation. Another field of great potential is that of ultra small (tens of nanometers) devices that take advantage of Bloch oscillations, the Coulomb blockade, and other mesoscale phenomena that could lead to new concepts such as a quantum current standard. Using these together with the more conventional superconducting technology, there is great opportunity to contribute to the development of sensors (predicted to become a $1 billion market worldwide in the next decade). In particular, good concepts have been proposed for infrared and millimeter wave radiometers and receivers, particle detectors, and temperature sensors. A closely related need is for amplifiers (integrated on the same chip) to extend the range of these basic sensors to make them attractive as system elements.

These very promising applications of superconductivity have all been studied and developed with low-temperature superconductors. The discovery of the high-temperature ceramic superconductors has stimulated much new exploration. Passive superconducting bearings are expected to an early HTS market.[2] Japanese companies are actively pursuing the use of high-T_c [high temperature] superconductors in a number of magnetic shielding applications."[3] All these ideas are receiving fresh attention, and new ones are being added. The future of this field is difficult to predict but holds great promise.

ECONOMIC SIGNIFICANCE AND U.S. COMPETITIVENESS

The growth of the superconductor market forecast by World Business Publications is shown in Table 56.[4] The data in this table represent the total market for both low-temperature superconductors, which account for all the present market, and high-temperature superconductors, which are expected to dominate the market in the later years. Note that already in 1988 there was a market of $172 million for low-temperature superconductor products.

Table 56
Estimated Future Markets for Oxide-Based and Niobium-Based Superconductors by Area of Application
($millions)

Sector	1988	1993	1998	2003
Electronics (approx.)	5-10	50	200	800
Electric Power	2	7	20	45
Medical MRI Magnets	130	140	165	180
Transport Systems	-	-	5	25
Magnetic Separation	3	20	60	175
High Energy Physics	30	350[4(a)]	80	100
Other Industrial	-[4(b)]	5	10	20
Total (approx.)	172	572	540	1345

Japanese forecasters take a generally more optimistic view than those in the U.S. The Nomura Research Institute forecasts a total world market of $43 billion in 2000 and $123 billion in 2010. This view is encouraged by the higher level of industrial research on superconductivity in Japan than in the U.S.

Evaluation of the superconductor market is more a matter of prediction than observation. The present very modest commercial market is in MRI systems, research magnets, and instrumentation. Historically, the market has been dominated by large government projects, both in the U.S. and abroad. These continue to be very important. For example, the Superconducting Super Collider project will create a transient, billion dollar market for superconductors.

For the future, the earliest large-scale applications of superconductivity are SMES and MAGLEV. Both are receiving much attention in Japan and elsewhere. Both will probably be implemented at first using LTS, with HTS to follow when it is ready to do so. In electronics, the most likely early applications are compact antennas and narrow band filters using HTS. These devices do not require the development of a process to fabricate HTS Josephson junctions.

The U.S. is in a strong competitive position for the commercial development of superconductivity. Overall, research-and-development funding levels in the U.S., Japan, and European Economic Community are comparable, although the balance of effort in Japan is slanted more towards commercial development. The U.S. has the advantage of long experience with large magnet projects funded by the Department of Energy. The U.S. also has the advantage of a large domestic military market that will accept the first expensive, high-performance products that would not find such a ready market in Japan. An example is the 8-bit digital signal processor recently announced by Fujitsu, the most advanced example of superconducting digital electronics in the world. Its best commercial prospect would be as part of a military surveillance system. Compact antennas and narrow-band filters are also likely to find an early military market. Japan has the advantage of national cooperative research-and-development programs with long-term funding at a level high enough to succeed. A long standing program in MAGLEV has been joined by more recently started programs in SMES, ship propulsion, and the development of HTS materials. In the U.S., several commercial consortia are seeking to gain some of the advantages of these national programs, but on a smaller scale and with less direct government involvement.

EVOLUTION OF THE TECHNOLOGY

LTS technology has received nearly three decades of development. HTS is the newcomer. The main barrier to the creation of a large market for LTS is that it requires liquid helium refrigeration, which is expensive and widely distrusted, largely because it has never been developed on a commercial scale. High-temperature superconductivity promises to eliminate this barrier by substituting at least liquid nitrogen refrigeration, which is cheap and well established on a commercial scale. This would then reduce the overall cost of the technology if other costs (such as those of the conductors themselves) do not turn out to be greater.

The basic components of a practical HTS technology do not exist yet, but steady progress is being made in their development. Critical current densities over 10^6 amperes per square centimeter have been reported in thin films, which is adequate for electronics, and over 10^4 amperes per square centimeter in bulk material in low magnetic fields. This is adequate for some, but not all large-scale applications. A recent report asserts that achieving 10^5 amperes per square centimeter would open many new applications.[5] Improvement is still needed. Single-layer microcircuits can be fabricated by photolithography, and successful insulating crossovers have been reported. Controllable techniques to fabricate Josephson junctions have recently been achieved in some HTS materials.

There are problems still to be overcome in the development of a working HTS technology. Sustaining a high critical-current density in a high magnetic field requires flux pinning, which is weakened at high temperature by thermally activated flux motion. It will be very difficult to fabricate Josephson junctions in HTS materials by the methods that have been successful with LTS. The chemistry of the oxide ceramics eliminates many candidate materials for use as insulating barriers, and other characteristics of these materials compound the problem. The attractive feature of LTS for computers is that operating at low-temperature enables low switching barriers (with the attendant low switching energy) to be used without interference by thermally induced random switching. Thus there is some question if a computer based on superconductors operating at liquid-nitrogen temperature or higher would have any advantage over a computer based on semiconductors operating at the same temperature.

All these problems may be overcome by clever ideas, just as similarly difficult problems in LTS were overcome. But they create great uncertainty in the future of HTS technology, especially in the time scale for practical, commercial development. A prudent plan is to continue the development of LTS technology in parallel with early efforts in HTS technology until the latter demonstrates clear superiority in each application. With the broad, worldwide effort to develop HTS, this may not take too long.

GOALS OF U.S. INDUSTRY FOR COMPETITIVENESS

The competition for the superconductor market is like a race with an unknown finish. Some very promising products can be foreseen now, but the real winners that will take advantage of the unique characteristics of superconductors (particularly HTS) to create a large market probably have not been invented yet. Nonetheless, the requirements to develop a practical superconductor technology are fairly clear.

Low-temperature superconductors for large electrical machines may claim to be a practical technology already, albeit an expensive one, as very large projects such as the Superconducting Super Collider attest. Superconducting wire is manufactured in multi-kilometer lengths with good quality control and predictable engineering characteristics, thanks in part to the efforts of the NIST superconductivity program. Improvements could be made by reducing AC loss and extending performance at very high magnetic fields. An ancillary development that would contribute much to the field is the establishment of commercially acceptable liquid-helium refrigeration.

In electronics, all the basic components of microcircuits using Josephson junctions and variations of the SQUID as active elements can be fabricated with good control, and several integrated circuits of fair complexity (tens of thousand active elements) have been demonstrated. However, these follow rather conservative design rules. The potential applications of Single Flux Quantum Logic and phenomena that occur in very small structures are largely unexplored as yet. They offer the opportunity to operate at very low signal levels. Attojoule (10^{-18} joule) pulses would be typical working levels. Attention to refrigeration at these very low heat loads required would also be of great benefit.

High-temperature superconductors have a greater hope of reaching a large market because of the simplicity of liquid nitrogen refrigeration and the possibility of discovery of a superconductor that operates at ambient temperature. However, much more remains to be done to develop a practical technology.

First, we need a practical conductor, that can be manufactured in multi-kilometer lengths with good quality control and reasonable cost. It should sustain a current density approaching 10^6 amperes per square centimeter in a field over 10 tesla. It should be electrically and mechanically stable, strong enough to withstand the Lorentz force in a large magnet, and resistant to corrosion. It should have acceptably low AC loss. Such a conductor would satisfy the requirements for a wide range of applications and would probably enjoy a large market.

In electronics, many of the components of a practical microcircuit remain to be developed, but some are in hand already. Processes have been established to fabricate thin films with high critical temperature and critical current density (over 10^6 amperes per square centimeter). Definition of circuits by photolithography without significantly degrading the superconductor's performance has been demonstrated. Surface quality can be made good enough for the microwave surface resistance to be lower than that of the normal metals used in passive microwave components, but several orders of magnitude of further improvement would be theoretically possible. The best films are obtained on only a limited range of substrate materials, which restricts microwave applications (for which the dielectric properties of the substrate are important) and the fabrication of multilayer circuits. A wider range of substrate materials must be made available, either by finding new materials with desirable dielectric properties or by developing processes to fabricate better films on the substrates we have already. Insulating crossovers can be fabricated now, but improvement would be desirable. Low resistance contacts to normal metals and semiconductors need further development. The greatest challenge of all is to develop a process for the controlled fabrication of Josephson junctions. Much ingenuity has been expended on this problem already with partial success, but we are a long way from a process that can be applied with good quality control, high yield, and low cost. Finally, a three terminal device with gain, whether based on LTS or HTS, would open up many new applications for superconducting microcircuits. A few such devices have been demonstrated, but improvement is needed to make them commercially interesting.

EVIDENCE OF MEASUREMENT NEEDS

In the early days of LTS technology, unanticipated measurement needs were demonstrated clearly by technical failures. At first there was uncertainty even in the definition of quantities to be measured, which became apparent particularly with respect to scale-up from small sample measurements to engineering designs. The fledgling industry making laboratory magnets suffered chronic difficulty with meeting performance specifications that were close to the state of the art, and a few large systems also failed to perform as designed. Since then the major problems have been solved (with significant contributions from NIST) and measurement needs have been moved to the level of engineering refinement. Much remains to be done for projects that challenge the state of the art (such as the Superconducting Super Collider), but there is also a body of established technology from which no surprises are expected.

The systematic method to test the adequacy of measurement practice is by round-robin intercomparison. A set of specimens is sent to all participating laboratories who perform nominally identical measurements using their own equipment and methods and then compare results. A series of such tests is being organized under the Versailles Project on Advanced Materials and Standards (VAMAS), among laboratories in the U.S., Japan, and the European Economic Community. The first test was of the measurement of critical current of a composite niobium-tin and copper wire.[6] Participants reported deviations up to 40%, which is not acceptable for engineering purposes; and work is in progress to find the cause. A second test, of the measurement of AC loss, is still in progress.

The field of measurement of the characteristics of high-temperature superconductors is wilder country, partly because of the lack of stability and uniformity of the specimens and partly because of the special characteristics of these materials such as low critical current density and a gradual transition at the critical field. Public comments to this effect were reported at a press conference with researchers from AT&T, IBM, and Harvard University.[7] The quantities that are particularly pressing problems are critical temperature (to support claims of the discovery of new superconductors) and critical current density (the basic measure of engineering performance). These were discussed at length at a HTS measurements workshop organized by the Defense Applied Research Projects Agency (DARPA) for its contractors and followed up by a round robin intercomparison of measurements of critical current density. One of the general conclusions is that variations in definition of the end point and variations in measurement systems can lead at present to variations of over an order of magnitude in the reported value of critical current density. This makes nonsense of the scientific literature and any attempt at systematic engineering development. A close analogy would be to a household fuse with an unknown rating somewhere in the range from 100 amperes to 1000 amperes. This would be of little use for the purpose of enabling appliances to be used while protecting the wiring of the house from fire.

Systematic studies of the needs for measurements of the characteristics of superconductors have been made in Japan by a committee of experts who reported[8] a list of 46 quantities relating to LTS and 12 relating to HTS that will need standardized measurement methods within 10 years. The International Electrotechnical Commission has formed a Technical Committee (TC-90) to address these needs. Some of the more prominent needs are addressed in the following section.

MEASUREMENT NEEDS

Improved measurement methods or new reference materials data are needed in three principal areas: electrical and magnetic characteristics, composition and structural analysis, and fabrication-process design and control.

Electrical and Magnetic Characteristics

Meetings organized by DOE, DARPA and the American Physical Society have all concluded emphatically that the published literature in HTS is in chaos because of inconsistency in test and measurement methods and reporting of results. The worst situation arises in measurement of critical current. The measurement of critical current is much more complicated than commonly perceived. Standards based on LTS materials are relevant, but not sufficient for consistent and accurate measurements of HTS materials. At high fields, differences in definitions and test conditions can change results by over an order of magnitude. In low fields, large differences can result from self-field effects (the field created by the current used for the measurement). The critical current can also be a function of the magnetic field history, both magnitude and angle. This field hysteresis can give a critical current which is higher or lower by an order of magnitude. Things are not much better with thin films and devices, especially since the effects of strain, environment, etc. have barely been touched. Consistent electrical and magnetic tests are needed to support development of *practical superconductors* and to incorporate them into national and international standards.

Other important quantities requiring standardization of measurement practice include: AC loss, critical field, and critical temperature. Agreement needs to be reached on the necessary and sufficient criteria defining the superconducting state, to support claims of the discovery of new superconducting materials that may be observed at first in very impure or highly diluted form. Possible applications of superconducting levitation and shielding will require measurements of the magnetic hysteresis of superconducting materials in standard conditions.

Composition and Structural Analysis

Computer-generated compositional maps of the surfaces of superconductor specimens, obtained by electron microprobe analysis with micrometer resolution, have proven to be a critical diagnostic tool in the development of processes to produce homogeneous materials. However, inhomogeneities at the nanometer scale have a significant effect on the superconducting properties of these materials, and segregation has been observed on this scale. Nanometer-scale compositional mapping is technically feasible and needs to be developed to its full potential.

X-ray diffraction and x-ray fluorescence analysis techniques are needed for on-line, non-destructive, non-invasive monitoring of chemical composition, phase, crystal orientation and texture of superconducting thin films during production. These techniques are commonly used to evaluate product materials. They would be of greater value if they were available for process monitoring.

The oxygen content of Y-Ba-Cu-O superconductors is variable and has a profound effect on the superconducting characteristics. Its variation with depth near the surface of a superconductor specimen could be measured by using oxygen enriched with the isotope 0^{17} in the preparation of the specimen. The distribution of this isotope could then be determined by neutron depth profiling.

Neutron scattering gives complete information on composition and crystal structure, but it is usually too slow to be used as a diagnostic technique to guide process development. The technology to reduce turn around time to as little as 30 minutes is available and should be applied.

Fabrication-Processes Design and Control

Weak flux pinning is a major stumbling block to the commercialization of HTS, in applications from micro-circuit interconnects to large-scale magnets. Strong flux pinning is the prime requirement for high critical current density. The most significant need for a measurement technique in this area is for a method to map the distribution of magnetic flux in a superconductor with the resolution of individual flux quanta (sub-micrometer). Magnetic measurements revealing the dynamics of flux motion would also be very valuable.

Phase equilibrium diagrams provide the basic information required to design fabrication processes to produce a single desired composition from a multi-component mixture. Definitive phase equilibrium diagrams for the Y-Ba-Cu-O system have been established. The Bi-Sr-Ca-Cu-O system is under intense study. The Tl-Ca-Ba-Cu-O system remains to be done. Several compositions are superconductors, among them being the compounds with the highest of all established transition temperatures. These may be expected to be of great practical importance. Because of the toxic nature of thallium, special laboratory facilities would be needed for this work.

The mechanical characteristics of HTS materials will be of critical importance in large-scale applications. The basic materials are weak and brittle, so their use as practical conductors will require composite wires or laminated tapes. The metal matrix (or substrate) used to bind the superconductor will also serve as the stabilizer and reinforcement. The elastic properties of the conductor and the strain in the superconducting constituent could be calculated by modelling the conductor as a metal-matrix composite. The calculated response could be verified experimentally using ultrasonic and mechanical techniques on composites. The results would provide the physical basis for optimizing conductor design and for establishing the strain dependence of critical current.

The nanometer scale compositional mapping facility described in the previous section could provide the means for a significant extension of research on processes to produce homogeneous materials. A variety of organic precursors would be used to prepare the starting material for sintering and the kinetics of phase formation during heat treatment would be followed by nanometer scale compositional mapping. The information generated would be vital to the design of fabrication processes that may be required to produce controlled inhomogeneities for flux pinning or mechanical stiffening.

The enormous quantity of published results of world-wide research on HTS is at present a source of more confusion than enlightenment. A critical evaluation and consolidation would create a very valuable resource that would enable all this information to become a solid base for further progress. A data base relating composition, process conditions, microstructure, mechanical characteristics, and electrical and magnetic characteristics for the various superconducting compounds should be created. Critically evaluated data are among the prime requirements for turning free-form scientific research into systematic engineering development.

ENDNOTES

1. *1990 Electronic Market Data Book*, Electronic Industries Association, p. 71 (1990). MRI equipment employs either superconducting magnetics or permanent magnets.

2. Lee Carlson, "Bullish on Bearings", *Superconductor Industry*, pp. 22-23 (Summer, 1991).

3. Quotation from Joshua Kawai, "Superconducting Magnetic Shielding", *Superconductor Industry*, p. 32 (Fall, 1991). Further information on shielding is found in "Dowa Mining Markets Cheap High T_c Bismuth Shield", *Superconductor Week*, p. 5 (November 11, 1991).

4. The data in this table come from World Business Publications, as published in *Superconductor Week* (February 27, 1989). The superscripts in the table refer to the notes that follow: (a) includes niobium-based superconducting magnets for the Superconducting Super Collider; (b) estimated at less than $1 million.

5. David Larbalestier, "Critical Currents and Magnet Applications of High-T_c Superconductors", *Physics Today*, Vol. 44, No. 6, p. 82 (June 1991). The article references the information as coming from S. J. Dale, S. M. Wolf, and T. R. Schneider, "Energy Applications of High Temperature Superconductivity", Electric Power Research Institute Report No. ER-6682 (1991).

6. "VAMAS Intercomparison of Critical Current Measurement in Nb_3Sn Wires", K. Tachikawa and others, *IEEE Transactions in Magnetics*, Vol. 25, pp. 2368-2374 (1989).

7. *Superconductor Week* (April 3, 1989).

8. Unpublished oral report at meeting of IEC TC-90 in Tokyo (May 1989).

CHAPTER 7

MICROWAVES

October 1992

Chapter 7

MICROWAVES

SUMMARY

The wireless revolution is taking on new dimensions as demand intensifies for new services based on microwave technology. That demand comes from consumer, commercial, and Government sectors. Included, for example, are calls for personal communications services, digital audio broadcasting, high-definition television, and intelligent vehicle-highway systems. These and other new consumer and commercial services lie at the leading edge of the information age. The demand for new services is complemented by the demand for improvements in the broad range of existing services provided by microwave technology. Included are aircraft safety, satellite communications, weather sensing, and defense. Through the combination of new and existing services, microwave technology supports important national goals in economic growth, security, safety, health, environment, communications, transportation, and energy efficiency.

The markets for microwave equipment are vast. In 1992, North-American shipments of microwave equipment have been estimated to be $33 billion. World production has been estimated at $61 billion for the same year. For the new services described above, world markets from billions of dollars to tens of billions of dollars are at stake in the realization of even individual services. Currently, the U.S. dominates its own market but has made relatively small penetration of foreign markets. The U.S. has much to lose if it cannot maintain its position in its own market as sweeping advances in microwave technology give rise to new commercial applications.

This growing role for microwave technology is based largely on its *speed* and *mobility*. Microwave technology is fast in the sense that it provides the highest information capacity available from an electronic technology. Supercomputers are already designed around microwave technology, and the next generation of desktop computers will have to employ microwave technology to achieve further speed gains. Microwave technology is mobile in the sense that it is a wireless technology, capable of operating efficiently with relatively small antennas. The wireless nature of microwave technology is convenient, as well; it enables wireless transmission across a room, across a continent, or around the world (via satellites).

The capabilities of microwave technology are leading to intense international competition in the development of new products. The nations that will win the expanding worldwide markets for these products must successfully pursue three revolutionary changes in microwave electronics: integration, higher frequencies, and extraordinary performance levels.

> Integration means integrating microwave electronic devices with each other, with optoelectronic devices, and with antennas. Integration is necessary to reduce the cost, size, and weight of microwave systems and to increase their power, reliability, and versatility.

> Exploitation of the higher frequencies means opening frequencies in the region from 30 gigahertz to 1000 gigahertz to greater use, or to first use, in order to gain access to the additional spectral space and special properties that they offer.
>
> Extraordinary performance levels mean improving speed, signal quality, sensitivity, versatility, and flexibility. These improvements are needed to increase information handling capacity and to open new applications areas.

These changes in microwave technology are just as significant as those that occurred in lower-speed semiconductor technologies when integrated circuits replaced transistors and vacuum tubes. In fact, that is largely what is happening now for microwave technology. These are exciting changes. They are already leading to powerful new products, such as entire microwave systems on individual wafers several centimeters (a few inches) in diameter. Each chip replaces boxes of earlier microwave hardware. The new "microwave systems on a chip" will eventually support new applications, such as robot vision, automobile anti-collision radar, and handheld telephones with world-wide access.

In the international arena, both Japan and Europe are distinguishing themselves. Japan is aggressively pursuing the development of microwave integrated circuits, driven almost exclusively by consumer applications. Further, Japan’s preeminence in gallium-arsenide semiconductor technology complements its capabilities in silicon technology and assures Japan’s access to the higher microwave frequencies. Europe is driving hard, too, with a strong focus on integration.

U.S. competitiveness in developing new products based on microwave technology is highly dependent on the development of the required measurement infrastructure. New measurement capability is needed to accelerate industrial research and development toward new products, to improve the design of manufacturing processes and the control of quality, to provide the basis for voluntary industry standards for compatibility, to support the specification and procurement of microwave products, and to improve access of U.S. products to international markets. Industry executives, through the IEEE’s unique Committee to Promote National Microwave Standards, and Government representatives, through the National Conference of Standards Laboratories, have decried the lack of needed measurement support, have surveyed the needs, and have spelled out the actions to be taken.

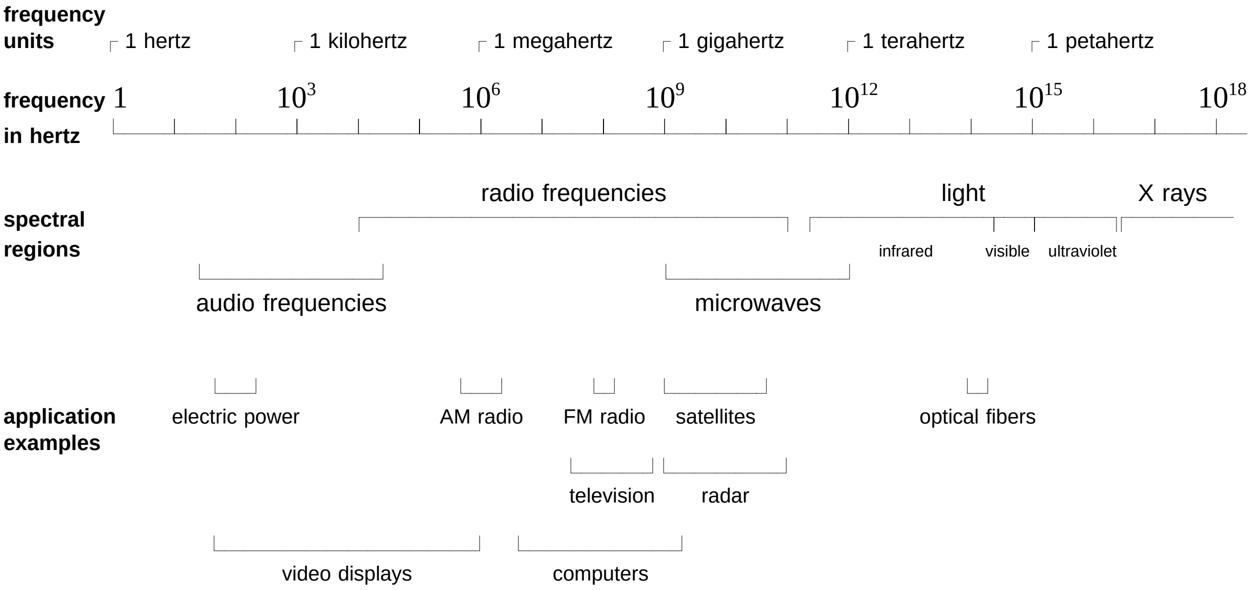
The new measurement capability is needed for three categories of microwave components: (1) individual components, including antennas, for 1 to 100 gigahertz; (2) integrated circuits and integrated antennas for 1 to 100 gigahertz; and (3) integrated circuits and integrated antennas for 100 to 1000 gigahertz. The most critical need for improved U.S. competitiveness is measurement support for the development and manufacture of integrated circuits in the frequency range of 1 to 100 gigahertz.

Among the generic types of improvements needed in measurement capability are these: higher accuracy levels, typically ten times greater; automated measurement methods suitable for integration with manufacturing processes, since manual measurement methods are not competitive; continuous frequency coverage instead of the present spot frequency coverage; and broader capability for measuring interconnections and interfaces to support integration, associated electronic "packaging" of the integrated circuits, and a merger with optoelectronic technologies. In all about 20 quantities require improved measurement support. In some cases the measurement needs can be met by refining and documenting existing measurement methods. In other cases new measurement methods are needed. New national measurement reference standards are needed to assure the accuracy of those measurement methods. Also needed are new materials reference data; the data are especially critical to the success of automated manufacturing processes for microwave integrated circuits.

DEFINITION OF MICROWAVES

Microwaves are radiated electromagnetic energy just like the energy used for radio and television broadcasts, except that microwaves are higher in frequency. Light is also radiated electromagnetic energy, but light is even higher in frequency. Microwaves have frequencies in the range of 1-1000 gigahertz, as shown in Figure 4. The figure has a logarithmic scale, providing equal space for each multiple of 10 increase in frequency.

Figure 4
Overview of the Frequency Spectrum



Most microwave systems today operate in the range of 1-30 gigahertz; but frequencies up to, and somewhat above 100 gigahertz, are increasingly used. Applications to 300 gigahertz are in the offing, and applications to 1000 gigahertz are contemplated for the future.

Microwaves get their name from the fact that they travel through the atmosphere as closely spaced waves. That is, if microwaves could be made visible, they would look like a succession of closely spaced ripples, much like the ripples that move across a pond when a rock is dropped in it. The distance from one ripple to the next in any electromagnetic wave is called the *wavelength*. The wavelength gets shorter as the frequency gets higher (inverse relationship). For microwaves the wavelength ranges from 30 centimeters at 1 gigahertz to 0.3 millimeter at 1000 gigahertz. For this reason, the different parts of the microwave frequency range are often referred to as centimeter waves, millimeter waves, and sub-millimeter waves.

Table 57 shows the frequency range corresponding to each named segment of the microwave frequency spectrum, and the corresponding wavelength range. Although the frequency range from

Table 57
Segments of the Microwave Frequency Spectrum

Segment	Frequency	Wavelength
centimeter waves	1 gigahertz	30 centimeters
	3 gigahertz	10 centimeters
	3 gigahertz	10 centimeters
	30 gigahertz	1 centimeter
millimeter waves	30 gigahertz	10 millimeters (= 1 centimeter)
	300 gigahertz	1 millimeter
sub-millimeter waves	300 gigahertz	1 millimeter
	1000 gigahertz	0.3 millimeter

1 to 3 gigahertz corresponds to wavelengths longer than 10 centimeters (30 to 10 centimeters), this range is usually lumped into the centimeter-wave category as shown in the table.

APPLICATIONS OF MICROWAVES

The applications of microwave technology are remarkably diverse. They include both information and energy functions. The information functions include radar, navigation, communications, computing, and sensing. The principal energy function is thermal processing. These functions support national goals in economic growth, security, safety, health, environment, communications, manufacturing, transportation, and energy efficiency.

Radar and Navigation

Microwave technology is already critical to navigation and radar systems used for general, commercial, and military aviation. Microwave technology is central for most advances planned for aviation, including new, more advanced automated landing systems, navigation via satellite-based systems, and new on-board windshear detection systems for aircraft that will complement present airport-based systems; these new systems are all DOT/FAA priorities.

Further, radar and navigation systems for automotive and highway applications are already under development and are a major focus of Japanese, European, and U.S. manufacturers, and of U.S. Government agencies, including DOT and DOD. The applications are broad and are often described as part of so-called intelligent vehicle-highway systems (IVHS). The goals of the IVHS are "to improve mobility and transportation productivity, enhance safety, maximize the use of existing transportation facilities and energy resources, and protect the environment."[1] The applications include near-obstacle detection, blind-spot detection, semi-automatic headway control, airbag arming, collision warning and avoidance, speed sensing, adaptive cruise control, navigation (including satellite-based systems), position tracking, and traffic management.[2] Even one of the traffic-management functions, such as traffic-light control to provide "green lights all the way", can have enormous economic impact in saving gasoline and time and in reducing air pollution.

Microwave technology will also provide the basis for local radar for guiding robot vehicles in manufacturing plants and for providing more sophisticated robots with vision.

Communications

Existing and new communications applications will serve individuals, corporations, and Government. New pocket cellular radio telephone systems, often employing digital technology, are emerging. New, wireless, handheld telephone systems with worldwide access via low-flying satellites are actively being planned.[3] They are a major part of the new move toward untethered communications systems that will provide the full range of services currently associated with wired systems. Similarly, expanded mobile communications, including direct-to-satellite systems for worldwide access, are in the offing. Mobile digital audio satellite broadcasts are planned for 1993.[4] Extensive corporate telecommunications systems are already in place,[5] and new systems are under development. Direct broadcast satellites for high-definition television (HDTV) are already in service in Japan and Europe, and FCC approval for their use in the U.S. has already been obtained. New communications systems will be a part of the new intelligent vehicle-highway systems and may offer such services as national or even international communications; weather reports; "navigation using electronic maps; route selection and guidance; information on services such as gas stations, restaurants, and hospitals; and real-time traffic information".[6]

Optical-fiber communications systems will require microwave electronics to realize the ultimate information capacity of the fibers (a data rate of about 1000 gigabits per second). Similarly, the most advanced formats proposed for HDTV will require microwave signal processing circuits to obtain sufficient speed. (A studio, or "production" quality, HDTV picture can require more than 1 gigabit per second of data.)

Computing

Special purpose high-speed computers are already employing microwave technology. The move to microwave technology will soon be required even for desktop computers which are already running at clock rates of 66 megahertz. To operate at clock rates of 100 megahertz and higher, they will have to begin employing microwave technology. Wireless local-area networks (LANs) operating at microwave frequencies are already a reality.[7]

Sensing

Continuous remote sensing of environmental changes through the Earth Observing System is a priority of NASA and DOC/NOAA. Global mapping and weather, sea, and iceberg monitoring are major applications. Detecting the effects of pollutants in the upper atmosphere and on marine life is a subject of research. Also underway is the detection of cancer (and its treatment). Measurement of the moisture content of soil, farm products (e.g., grain), and paper products is under study and shows great promise. Non-destructive testing of materials during, or between processing steps, is a promising area. Process measurements have been successfully made, including determination of the resistivity of wafers of semiconductor silicon used to make integrated circuits.[8]

Thermal Processing

Commercial and domestic cooking are familiar applications of microwave technology. Now commercial and domestic applications in clothes drying are being evaluated. They offer reductions in drying temperatures, tumbling, and energy consumption.[9] Heating in industrial processes is already a reality, including heating to accelerate chemical reactions. More selective use of microwave energy to promote specific chemical reactions, sensitive to the frequencies of the microwaves used, are presently in the research stage. Microwave heating is used for sterilization of medical devices and for a variety of medical procedures, including the treatment of cancer mentioned above.[10]

CAPABILITIES AND LIMITATIONS OF MICROWAVES

This diversity of applications is sourced in the unusual capabilities that microwaves offer. Both the capabilities and the limitations of microwaves are commonly described in terms of their performance relative to radio waves of lower frequency. Each of the following sections describes an important characteristic of microwaves. Some of these characteristics are closely interrelated.

High Information Capacity

Microwaves can carry very large amounts of information. This high capacity results from the fact that microwaves have a very high frequency. All radio waves, including microwaves, can carry information in proportion to their frequency. For example, a microwave beam with a frequency of 1 gigahertz can carry 1000 times more information than radio waves used for AM radio, which have frequencies near 1 megahertz. Put another way, one microwave signal can carry all of the program material of 1000 AM radio stations simultaneously. This high information capacity is important both to telecommunications applications and to signal-processing applications.

The space in the frequency spectrum is precious and thus the high information capacity of the microwave region is of special interest. Generally speaking, a given frequency in the spectrum cannot be used by two systems at the same time in the same location without mutual interference. The often intense competition for the use of individual frequencies within the U.S. is resolved by Government agencies, principally the Federal Communications Commission (FCC) and the National Telecommunications and Information Administration of the U.S. Department of Commerce. At the international level the World Administrative Radio Conference (WARC) plays the key role.

Line-of-Sight Transmission

When microwaves are radiated by antennas, they travel in a straight, line-of-sight manner, just as light travels. This is true because microwaves, unlike radio waves of lower frequency, pass right through the ionosphere, which is the electrically charged atmospheric layer surrounding the earth. For this reason, microwaves cannot work their way around the globe by bouncing back and forth between the earth and the ionosphere, as radio waves of lower frequency do. Instead, microwaves must be relayed along the surface of the earth from one microwave tower to another, or they must be transmitted to satellites that retransmit them downward to desired locations or that relay them to other satellites for retransmission downward. This line-of-sight characteristic is a limitation for some applications: it makes microwaves ineffective for terrestrial (earth surface-to-surface) wide-area broadcasts. For other applications, it is an advantage; it enables microwave systems serving nearby areas to employ the same frequencies with reduced mutual interference. This characteristic of microwaves improves efficiency in the use of the microwave portion of the frequency spectrum.

Sharp Focusing

Microwaves can be focused into narrow beams much more easily than radio waves of lower frequency. In this sense microwaves approach the characteristics of light; light is readily focused by reflectors like those in flashlights. Because microwaves can be focused, their energy can be delivered to small nearby receiving antennas or, by satellites, to entire geographical areas, such as entire states, with equal success. Also, because of focusing, two or more microwave beams can be directed to adjacent areas without mutual interference, even if they operate on the same frequency. This characteristic further improves efficiency in the use of the microwave portion of the frequency spectrum.

Small Antennas

The high frequencies of microwaves reduce the size of the antennas needed to achieve a given degree of focusing and efficiency in transmitting and receiving electromagnetic energy. These advantages reduce the cost of microwave systems and increase the diversity of the applications that they can serve. For example, the relatively small size of microwave antennas makes them suitable for roof-top mounting for corporate communications systems.

Both small size and energy efficiency are especially important to mobile applications. These include microwave systems for automobiles, airplanes, satellites, space vehicles, and personal, handheld use. The importance of small size for these applications is obvious. Energy efficiency is also very important because mobile systems often run on battery and solar-cell power and thus are limited to lower transmitting power levels. For example, a satellite transmitter may emit only 5 to 20 watts of power to illuminate receiving antennas about 30 kilometers (19,000 miles) below on the surface of the earth.[11] This power level compares to AM, FM, and TV stations which emit levels from 100 watts to 5 megawatts, depending on frequency and local requirements, to illuminate receiving antennas only tens of miles away.[12] The higher power levels for these terrestrial broadcast systems are necessary to support inefficient receiving antennas, like the small antenna just a few centimeters in size in a handheld AM radio. An efficient AM radio antenna would be 150 meters (about 500 feet) long.

The size reduction provided by microwave frequencies is particularly important for the design of highly focused antennas, since antennas with a high degree of focusing are larger than those with a lesser degree of focusing at a given frequency. Even very highly focused microwave antennas can be tolerable in size (several meters). Antennas with comparable focusing at lower frequencies would be impossibly large.

High Spatial Resolution

The higher frequencies of the microwave spectrum provide high spatial resolution. High resolution is particularly important for radar systems since it allows them to "see" objects more clearly and thus to identify them more accurately. In particular smaller objects can be seen than would be possible with radio waves of a lower frequency. One example is improved imaging of airfields by automated landing systems for use under conditions of zero visibility. Another example is the ability to distinguish a truck from a car in intelligent vehicle-highway systems.

Atmospheric and Materials Interactions

Microwaves, especially at the higher frequencies (millimeter waves and sub-millimeter waves), interact more strongly with the atmosphere than do radio waves of lower frequency. Selected frequencies of microwaves are highly absorbed. This interaction limits the distances over which microwaves at affected frequencies can travel through the atmosphere -- a problem for some applications. For other applications, this interaction is an asset: it enables the use of microwaves to sense weather conditions that radio waves of lower frequency cannot "see"; and it enables limiting a transmission to a controlled distance to reduce unwanted reception by nearby users or unwanted interference with them. This characteristic is particularly important when using millimeter waves for local applications such as local communications systems or local radar systems for monitoring moving objects. In free space, atmospheric interactions are not a problem, even for satellite-to-satellite communications, since satellites orbit well above the atmosphere.

Microwaves interact more strongly with some materials than do radio waves of lower frequencies. This enables microwaves to expand the range of materials than can be heated in processing applications beyond those addressable with radio waves of lower frequencies alone. In particular, microwaves of certain frequencies interact strongly with water. This interaction provides the basis for microwave cooking and microwave drying.

Optoelectronic Technology

Microwave technology can be integrated with optoelectronic technology to provide powerful capabilities that are not achievable by either technology alone. For example, advanced microwave technology is used to place information on lightwaves in optical-fiber systems to achieve very high information rates.[13] Optical fibers show promise for use in large microwave antennas. Embedded optical waveguides show promise for use in integrated microwave antennas and in microwave satellite systems generally.[14] Progress in either microwave or optoelectronic technology thus depends on progress in the other.

Preeminent Benefits

If one had to settle on two preeminent benefits of microwave technology, they would be *mobility* and *speed*. The mobility results from the untethered, wireless nature of many microwave systems and from the relatively small size of the antennas needed, relative to those required by radio waves of lower frequency. Mobility is convenient; further, wireless systems are convenient even when mobility is not required [wireless local-area networks (LANs) for computers]. The speed of microwave technology results from the high information capacity of microwaves. These benefits of mobility and speed are realized in radar, navigation, communications, and computing functions. They provide the productivity and flexibility that explain the continued growth of applications of microwave technology.

Note that microwaves are associated with *smallness* in several ways: small waves, measured in centimeters and millimeters; small antennas; and fine scales of resolution for radar. These characteristics derive from the high frequencies of microwaves.

RELATIONSHIP OF MICROWAVE AND OPTICAL-FIBER TECHNOLOGIES

The impact of optical-fiber cable technology on the markets for microwave technology is often a subject of interest. Optical-fiber technology has an impact but a very limited one. The reason is that optical-fiber technology serves only one category of applications provided by microwave technology: terrestrial communications between fixed points of generally high population density. Further, in the area of communications viewed more broadly, the demand for services is so great that all available means of transmitting data continue to be needed. Even in the short-haul terrestrial communications, including both common-carrier and private networks as well as other services, microwave technology continued to show mild growth in 1991, not a contraction,[15] in spite of the continued growth of optical-fiber communications. The continued spectacular growth in wireless, cellular telephone systems is a major indicator of the appeal of wireless technologies; U.S. subscribers increased 40 percent from 1990 to 1991.[16] For long-haul communications via satellites, microwave systems continue to expand. For example, U.S. manufacturers delivered 9 commercial satellites in 1991 and expect to deliver 10 in 1992 and 12 in 1993; they have orders for 13 more for 1994.[17] Further, the market for emerging "smallsats" appears to be exploding and would affect the market in the mid-1990s. They weigh typically less that 450 kilograms (1000) pounds and fly in low orbit 9000 kilometers (5000 nautical miles) to support low-power systems. The present 11 proposals in place for smallsat systems "translate into nearly 270 smallsats."[18] Such low-earth-orbit systems, as presently conceived, will operate at frequencies from 137 megahertz to several gigahertz.[19]

For all other categories of applications served by microwave technology, optical-fiber technology cannot compete at all. The other applications require one or more of the capabilities that cable technology cannot provide, as outlined in Table 58. These capabilities include flexibility, area coverage, mobility, and microwave energy. *Flexibility* may require clarification. It takes several forms: (1) Microwave systems, as a wireless technology, can be set up quickly without the need to resolve the difficult right-of-way problems that arise when laying cable. (2) Microwave installations are less sensitive to variations in the terrain than cable installations. (3) Satellite and terrestrial microwave systems can be reconfigured readily to serve different geographical areas and new users. (4) Satellite systems can serve vast areas of low population density just as easily as small areas of high population density. Because of this flexibility, microwave systems can serve in both permanent installations and temporary installations such as those used for emergencies or when terrestrial or undersea cable systems have been damaged.

The more relevant relationship between optical-fiber and microwave technologies is complementary or even symbiotic.[20] For example, microwave systems will provide "local-area" delivery for signals carried across the country by optical fibers, particularly when mobile subscribers are involved. Further, microwave and optical-fiber technologies will gradually merge as optical-fiber technology takes on functions in microwave systems and microwave technology takes on functions in optical-fiber systems. Together, they will provide levels of performance that neither can achieve alone.

Table 58
Microwave Capabilities That Cable Cannot Provide

Capability	Applications Exploiting the Capability
area coverage	communications, sensing, radar, navigation, guidance, electronic warfare, consumer applications
mobility/movement	communications with satellites, mobile individuals, and vehicles including automobiles, aircraft, ships, and spacecraft; radar
flexibility	communications, electronic warfare, emergencies
microwave energy	industrial, medical, and consumer applications

MICROWAVE ELECTRONICS MARKETS

U.S. Market

Locating all U.S. shipments of microwave electronic products is difficult because microwave technology is implicit in so many products that are not themselves labelled *microwave*. The discussion above indicates the broad range of those products. Table 59 shows the categories into which microwave products fall and for which U.S.-shipments data are available. These categories have been taken from Table 17 in Chapter 3. Of course, not all of the products in these categories are based on microwave technology. In addition, Table 59 lists microwave electrical appliances; they are considered outside the definitions of electronic products or electrical equipment, as described in Chapter 3.[21]

Only one estimate of microwave-equipment shipments has been found. That estimate is $33 billion for North America for 1992, a figure presumably dominated by the U.S. component.[22] The estimate includes microwave products performing information functions, such as communications, radar, and navigation, and excludes microwave products performing energy (heating) functions, specifically microwave ovens.

A very large part of U.S. microwave equipment falls in the category "radio communications and detection equipment", which focuses on information functions and which includes U.S. Standard Industrial Classification (SIC) codes 3663, 3669, 3812. U.S. shipments for this category were $55 billion in 1991 according to the Department of Commerce. Because U.S. exports and imports are relatively small fractions of U.S. shipments for this category (11 percent and 6 percent respectively, for 1991),[23] the value of the U.S. market for microwave equipment within this category is unlikely to differ much from the value of U.S. shipments.

U.S. Government Purchases

Of the $55 billion of U.S. shipments in the category "radio communications and detection equipment" for 1991, somewhat less than half represented purchases by the U.S. Government, according to the Department of Commerce.[24] The Government fraction of U.S. shipments of microwave equipment

Table 59
Microwave Electronic Products and Appliances

- Components
  - electron tubes
  - solid-state (semiconductors)
  - other components

- Communications Equipment
  - commercial, industrial, and military communications equipment
    - mobile and portable communications equipment
    - space satellite communications systems
  - search and detection, navigation and guidance systems and equipment
    - radar systems and equipment
    - missile and space vehicle equipment
    - navigation systems
    - counter measures equipment
    - special electronic and communication intelligence equipment
    - other (sonar, specialized command and control, etc.)

- Industrial Electronics
  - control and processing equipment
    - processing equipment
  - testing and measuring equipment

- Electromedical Equipment
  - other (patient monitoring, irradiation equipment, therapeutic equipment)

- Electronic-Related Products and Services
  - aerospace equipment (aircraft, missiles)

- Electrical Appliances
  - commercial microwave ovens
  - residential (consumer) microwave ovens

within this broader category may be higher because of the heavy use made by the military of such equipment.

While the DOD may continue to account for the largest fraction of the Government segment of the microwave equipment market, other Government agencies are increasingly involved in using large amounts of microwave equipment. The Department of Transportation, the Department of Commerce, and the National Aeronautics and Space Administration provide examples:

> The Department of Transportation (Federal Aviation Administration) is currently implementing its Capital Investment Plan (its expansion of the National Airspace System Plan) to provide improved aircraft management systems. The estimated cost is $16.2 billion (1991 dollars).[25] As of February 1991 "all but two of the original 90 plan projects are under contract, and over 30 percent are completed."[26] The plan includes: (1) communications systems for exchange

of information between airplanes and control towers; (2) landing systems for aircraft; (3) radar systems for monitoring weather; and (4) radar systems for monitoring the movement of airplanes in the air and on the ground.[27] As part of the expanded plan, the FAA is exploring the possibility of meeting all of the nation's civil navigation needs with the Global Position System (GPS), a 24-satellite navigation system being built by the Department of Defense at a cost of $10 billion.[28] It is slated for completion in 1993.[29] The use of the GPS receivers by aircraft will enable fuel savings by providing access to more direct routes and more efficient altitudes.[30] The U.S. GPS receiver market for a broad range of applications has been estimated at $6.7 billion for 1991 through 1996.[31]

The Department of Commerce (National Oceanic and Atmospheric Administration), the Department of Defense (Air Force), and the Department of Transportation (Federal Aviation Administration) are building an advanced weather radar system with 152 sites in the U.S. and territories at a cost of $440 million, to be completed by 1996. As of September 1992 the first site had been installed.

The National Aeronautics and Space Administration and the Department of Commerce (National Oceanic and Atmospheric Administration) are actively pursuing the development of the Earth Observing System. This system would employ satellites to monitor long-term changes in the earth's atmosphere, vegetation, magnetic field, geology, etc. The system would employ 6 satellites launched over a four-year period beginning in 1998 at a cost of $11 billion over eight years.[32] The system would rely critically on advanced microwave and optical equipment to perform its remote sensing and communications functions.

World Market and U.S. Competitiveness

Only one estimate for the world market for microwave equipment has been found. That is $61 billion for 1992.[33] Again, the information functions are included in this estimate, and the energy functions are excluded.

The U.S. balance of trade for the broad category of "radio communications and detection equipment" was favorable for 1991. However, as noted above, since both U.S. exports and U.S. imports are relatively small percentages of U.S. shipments, one can conclude that the U.S. is not penetrating foreign markets very well, and other nations are not penetrating the U.S. market very well. Early data for 1991 indicate that the major foreign markets receiving U.S. exports were Japan (10 percent), Canada (10 percent), Mexico (9 percent), and the United Kingdom (7 percent). The dominant sources of imports into the U.S. were Japan (30 percent), Canada (12 percent), Taiwan (8 percent), Mexico (7 percent), and the United Kingdom (5 percent). For 1991, imports from Japan were nearly three times those from any other country. Also, imports from Japan were roughly twice U.S. exports to Japan.[34]

Against this background, there are a number of indicators of intense competition for emerging commercial markets for the products of microwave technology. Clearly, the U.S. has much to lose if it cannot maintain its dominance in the U.S. market in the face of the sweeping technological advances in microwave technology that are enabling new commercial applications. Here are a few examples of the new applications.

High-Definition Television

Japan leads in the development of high-definition television and offers a full line of production equipment for programs. Japan is already broadcasting HDTV programs using a direct broadcast satellite on a regular basis (8 hours per day) to about 200 public viewing places.[35] Europe follows in HDTV development. The U.S. is at an earlier stage still and is presently formulating and testing standards for HDTV formats. The world markets at stake are enormous. Based on present markets for video products, the HDTV market will likely exceed many $10s of billions and will include a significant amount of microwave technology.

Digital Audio Broadcasting

Broadcasting of digital audio signals with the quality of compact discs appears inevitable. Both satellite and terrestrial broadcast options are being considered, as are both microwave frequencies up to 3 gigahertz and submicrowave frequencies down to the present FM-radio band located at 100 megahertz.[36] Both mobile and fixed receivers would be served. The audience for such "radio" broadcasts is worldwide; the potential markets for new equipment are vast.

Personal Communications Services

The potential market for wireless personal communications services (PCS) is also vast. Recent developments in this area have led the Department of Commerce to conclude that "tremendous potential exists for generating strong new demand for radio communications equipment over the next 5-10 years."[37] PCS is also called Future Public Land-Mobile Telecommunications System (FPLMTS).[38]

The demand has been so strong that the Federal Communications Commission has proposed a major spectrum reallocation of the frequencies in the 1.85 to 2.2 gigahertz band to permit rapid commercialization of emerging telecommunications technologies, such as personal communications services.[39] This band is low enough in frequency to enable the use of relatively inexpensive silicon semiconductor technology but high enough to provide adequate information capacity. [Above about 2.5 gigahertz, more expensive gallium-arsenide semiconductor technology is presently required.] Further, "a number of countries will be pushing for a worldwide spectrum allocation for PCS at the WARC [World Administrative Radio Conference]".[40]

Personal communications services may eventually serve half of the present 130 million U.S. subscribers served by wired systems. Worldwide, 150 to 200 million subscribers are estimated. However, the U.S. appears to be lagging: "While European and Japanese personal communications spectrum, technology, and standards issues are taking shape, the U.S. lags far behind in all aspects of personal communications."[41]

Intelligent Vehicle-Highway Systems (IVHS)

The estimated cost of implementing IVHS cannot be stated with precision, partly because the range of potential services is so broad, and partly because those that will be implemented are not yet clear. However, what is clear is that even a minimal national system for the U.S. will require a total investment of many tens of billions of dollars. As a result of IVHS, the world market for electronic equipment in automobiles may rise to $28 billion by the year 2000, according to the National

Advisory Committee on Semiconductors. One manufacturer, Motorola, estimates that "as many as 300,000 autos or 10 percent of 1994 models prices over $20,000 will carry such equipment [an advanced driver information system], adding $3000 to the list price."[42] Much of the equipment for IVHS will depend critically on microwave technology (communications, navigation, and radar).

Japan, Europe, and the U.S. are all highly active in IVHS. Europe and Japan have been preparing the infrastructure for such new programs for some time. The U.S. is just beginning, but the momentum is increasing.[43] In general, the U.S. is relying on independent demonstration projects, "while Europe and Japan are moving forward with major funding in place and scores of coordinated projects under way."[44]

GOALS OF U.S. INDUSTRY FOR COMPETITIVENESS

To meet the demand for expanded and new microwave services and to compete internationally, U.S. industry will pursue three goals for microwave systems:

(1) Reduce cost, size, and weight.
(2) Improve access to higher frequencies.
(3) Improve performance.

Reduce Cost, Size, and Weight

Clearly, there are enormous present and potential markets for the products of microwave technology. However, microwave technology is highly sophisticated and relatively costly. Reduction of costs to levels appropriate for mass commercial and consumer markets is highly dependent on the continued development of microwave integrated circuits and integrated antennas. They will enable cost reductions by factors of 100. As a result, *the first nations to achieve broad commercialization of microwave integrated circuits and antennas will gain a tremendous advantage in both the U.S. market and the international market for microwave products.* The cost reductions, in turn, will inevitably open larger markets for microwave products. Not surprisingly, size and weight reductions are particularly important to mobile, air, and space applications. The central role of integration is well realized by the competitors of the U.S. Both Japan and Europe are pushing hard to develop microwave integrated circuits.

Improve Access to Higher Frequencies

A second requirement for competitiveness is to gain access to the available spectral space and the special properties of the higher frequencies. Development of low-cost systems, using integrated circuits, with capabilities to at least 100 gigahertz is needed in the near term. Later, systems operating to frequencies of 300 gigahertz and then 1000 gigahertz are foreseeable. At these very high frequencies microwaves begin to take on quasi-optical characteristics.

A factor in competitiveness is the effort to push silicon technology to higher frequencies. Silicon technology, is discussed in Chapter 4, Semiconductors, beginning on page 53. Silicon technology is the basis for virtually all integrated circuits operating at frequencies below the microwave region, that is below 1 gigahertz. At present silicon technology provides low-cost devices with operating frequencies up to only about 2.5 gigahertz.[45] However, new silicon fabrication processes promise silicon devices with operating frequencies above 10 gigahertz.[46] The upper frequency limit will likely advance, but just how far is very difficult to guess. At the present time, to access frequencies above 2.5 gigahertz with solid-state technology, more expensive and complicated gallium-arsenide technology is used. Also, whenever significant power levels are required, vacuum tubes must be used. They are klystrons, magnetrons, gyrotrons, and traveling-wave tubes; the latter three employ magnetic materials. Vacuum tubes are likely to remain important for all applications requiring high power levels (transmission and thermal applications). Japan's excellence in silicon and gallium-arsenide semiconductor technologies, and in microwave tube technology, continues to position Japan well as a major international competitor.

Improve Performance

Improved performance will require higher information density, improved signal quality, and expanded versatility and flexibility.

Higher information density can be achieved by increasing the information capacity of existing systems and by enabling a greater number of systems to operate without mutual interference in a given region. These advances are highly important because crowding is becoming increasingly serious as additional terrestrial systems come into use. Only limited space is available in the valuable geostationary orbit.[47] To enable satellites to operate without mutual interference, high performance antennas with low side lobes must be developed.[48] Also, higher frequencies must be used to provide additional spectral space. Finally, the information capacity of microwave frequencies already in use and of existing optical-fiber systems must be increased by developing advanced microwave signal processing electronics.

Improved signal quality is needed to maintain low data error rates, especially as information capacity is increased. Improved signal quality is also needed to improve the sensitivity of microwave systems so that they can serve wider applications in communications, radar, and sensing. Such improvements will require improved fidelity in the microwave signal processing circuits for both digital and analog signals. As microwave systems are pushed to higher operating frequencies, the performance of electronic devices general degrades, so a natural competition arises between frequency and signal quality.

Expanded versatility and flexibility will also be required. They can be achieved by developing microwave systems that can operate on many different frequencies simultaneously and that can serve many different geographical areas simultaneously through electronically steered antennas.

EVIDENCE OF MEASUREMENT NEEDS

Successful pursuit by the U.S. of the three goals for improved competitiveness will require major advances in supporting measurement capability. This measurement capability is needed to support all phases of product realization and use: research and development, manufacturing, marketplace transactions (including voluntary standards, marketplace entry, product specification, and performance evaluation), and after-sales support (including installation, maintenance, and operation).

The urgency of this need has been stated in unqualified terms by U.S. industry:

> U.S. industry executives have expressed this need individually and have joined together under the auspices of the Institute of Electrical and Electronics Engineers to form the Committee to Promote National Microwave Standards. The Committee members, which include company chief executive officers and chairpersons of the board, have issued a report stressing the urgency of the measurement needs, decrying the weakening position of the U.S. compared to its competitors, spelling out the specific measurement problems, and urging resolution of them by NBS [NIST].[49]
>
> The National Conference of Standards Laboratories (NCSL), operating under the guidance of member representatives from major U.S. high technology companies and Government agencies, surveyed 400 organizations (346 industry, 61 government, and 4 universities) to determine key measurement needs. NCSL found a serious shortfall in NBS [NIST] measurement support for microwave technology with adverse effects on major sectors of U.S. industry. The report spelled out the measurement problems and urged resolution by NBS [NIST]. It also noted the negative effects of inadequate measurement support on national goals, such as productivity improvement, quality improvement, and international competitiveness, among others.[50]
>
> Industry trade articles have highlighted the urgency of the need for expanded microwave measurement support from NBS [NIST] for many years.[51] For example, one article notes: "The NBS struggles under a woefully meager budget to research measurement technology and supply measurement services to support the measurements revolution of the ’80s and the ’90s. Truly an impossible task."[52]

MEASUREMENT NEEDS

Needs Discussed in Other Chapters

The measurements needed to support new products employing microwave technology are extensive and diverse. This chapter describes a broad range of those measurement needs, but not all of them. Some of the measurement needs are described in Chapter 4, Semiconductors, beginning on page 53. That chapter addresses measurements for characterizing active electronic materials (such as silicon semiconductors and gallium-arsenide semiconductors), the internal workings of individual active electronic devices (such as transistors and diodes), dielectric materials (principally silicon dioxide) used inside the active electronic devices, and the fabrication processes for making individual electronic devices and integrated electronic circuits from semiconducting materials. Chapter 5, Magnetics, beginning on page 95, addresses the measurements required for characterizing magnetic materials, including those used in microwave power tubes and in other microwave components (such as circulators, isolators, and filters). Chapter 12, Electromagnetic Compatibility, beginning on page 381, describes the measurements required for assuring electromagnetic compatibility among the growing number and diversity of electronic systems.

Needs Discussed in This Chapter

This chapter focuses on measurements at the circuit level and on measurements of materials not covered in Chapter 4. Specifically, this chapter addresses measurements made at the inputs and outputs of individual electronic devices (whether based on semiconductor or tube technology), integrated circuits, and antennas (whether made from individual components or in integrated form). Included are measurements made at the inputs and outputs of electronic devices *internal* to integrated circuits and integrated antennas. Also included are measurements made on the interconnections between individual components and within integrated components. This chapter also addresses measurements for conducting and insulating (dielectric) materials used in interconnections, whether internal or external to components, in substrates on which microwave circuits are placed, and in antennas. The measurement needs described in this chapter are most readily classified by the types of components supported, as outlined in Table 60.

The most critical need is to provide the measurement capability required for the successful development of microwave integrated circuits in the frequency range of 1-100 gigahertz where most new systems will operate in the near term. Much of this work will be built upon measurement capability for individual components, some of which is already in place. In the future, measurement support for systems operating in the frequency range of 100-1000 gigahertz will be required, with first emphasis on the range of 100-300 gigahertz. The following sections describes the measurement needs arising for each category of components.

Table 60
Categories of Microwave Measurement Needs

- Individual Components for 1-100 Gigahertz
  - Electronic Devices
  - Antennas

- Integrated Components for 1-100 Gigahertz
  - Integrated Electronic Circuits
  - Integrated Antennas

- Integrated Components for 100-1000 Gigahertz
  - Integrated Electronic Circuits
  - Integrated Antennas

Virtually all microwave systems contain circuits operating at lower frequencies, too, such as 1 to 100 megahertz. Some supporting measurement capability will be needed at these frequencies as well. Although, that capability has not been catalogued here, it often requires measurement of many of the same measured quantities discussed here.

Generic Challenges of Microwave Measurements

The character of the measurements required to support electronic systems changes dramatically with increasing frequency. As the frequency increases from the lower radio frequencies to the microwave frequencies, the performance of electronic circuits become increasingly sensitive to the size, placement, surroundings, and materials of every electronic part. This sensitivity is reflected especially in interconnecting wiring. That wiring must increasingly be treated not just as a path for an electrical current but also as a transmission line for an electromagnetic wave associated with that current. As a result, supporting measurement methods must be increasingly sophisticated in order to (1) measure a wider range of potential interactions among circuit elements, and (2) avoid interactions between the measuring instrumentation and the circuit elements, thus distorting the measurement process.

At even higher frequencies, approaching those of light, the measurement of electrical currents becomes progressively of less interest, and the measurement of electromagnetic waves becomes necessary. In fact, the waves may be carried by materials that are insulating, such as optical fibers, and thus are not capable of carrying an electrical current at all in the normal sense.

The transition into the microwave region comes on gradually. For analog circuits, the lower boundary of the microwave region, 1 gigahertz, is a good dividing line. However, for digital circuits, such as those used in computers, a clock rate of about 100 megahertz, as noted above, is a good dividing line. At that clock rate, the electromagnetic energy present in a digital signal will contain frequencies extending into the lower microwave region near 1 gigahertz.

Individual Components for 1-100 Gigahertz

The new measurement capability needed for individual components must embody major advances over those currently available. These advances are summarized in Table 61 and are explained more fully in the following text.

Table 61
Measurement Advances Required for Individual Components

Improvement	Comment
higher accuracy	typically ten times greater
susceptibility to automation	manual measurement methods too slow for sophisticated new microwave systems
continuous frequency coverage	not just selected frequencies
broader interface compatibility	to support new interfaces for miniature components and new optoelectronic devices

The new measurement methods must be typically ten times more accurate to support emerging high performance systems. Achieving that accuracy will require new measurement reference standards. New measurement methods must also be susceptible to automation since modern systems are so complex that they cannot be efficiently characterized with today's manual measurement methods. The new measurement methods must operate over broad, continuous frequency ranges to support new systems with greater frequency flexibility; in the past, measurements at key fixed frequencies sufficed. The new methods must support a broader range of interfaces for two reasons: new miniature microwave components, with new types of interfaces, are being introduced to cut the size, cost, and weight of microwave systems; and new optoelectronic devices will be interfaced with microwave devices. For example, optical fibers will likely replace metal waveguides in key roles in satellites and antennas. The fibers are lightweight, low cost, and flexible and do not interfere with microwave beams. They may serve as signal interconnections within satellites or as controlling lines within powerful phased-array antennas, among other applications.[53]

The individual components that require improved measurement support may be divided into two groups: (1) *electronic devices* that generate, process, and transport microwaves without radiating them, such as signal sources, amplifiers, detectors, waveguides, coaxial cable, and connectors; and (2) *antennas* that transmit and receive the microwaves. The measurement needs of these two classes of components are described below.

Electronic Devices

New measurement capability is required for determining critical microwave quantities in individual electronic devices. The most critical measurement needs correspond to the two functions that a microwave system must perform to convey information from one location to another: (1) transfer microwave power efficiently, and (2) preserve signal fidelity while doing so. In addition, improved measurement methods and reference data are required for the microwave properties of materials. These materials measurements support the design and manufacture of new components. The specific quantities requiring improved measurement support are shown in Table 62.

Measurements of electronic noise serve as an example of the problems currently faced by industry. Noise is unwanted electronic energy that tends to obscure the desired signal. Noise is the single most important characteristic of microwave amplifiers and amplifying components (transistors); premiums are charged and paid for good noise performance. Noise performance affects the sensitivity and information capacity of systems. Yet present noise measurement capability is not adequate to characterize the noise performance currently being achieved in modern products, so manufacturers and buyers cope constantly with inadequate tools for product development, specification, and evaluation. New methods for noise measurement are needed that are far better than those currently available and that are suitable for adoption as standard methods industry wide. Manufacturers have repeatedly sought NIST assistance in resolving this problem.

Materials are important because the performance of microwave components is very sensitive to the microwave properties of the materials of which they are made. That sensitivity is particularly important in the design of the "substrates" (similar to the circuit boards of conventional electronics) that provide the waveguide channels that interconnect individual microwave electronic components. Unfortunately, most existing data on the microwave properties of materials are based on measurements made many years ago on conventional materials at frequencies below 10 gigahertz. Few data exist for frequencies above 10 gigahertz, and practically no data exist for newer materials.[54] The shortfall in measurement support and reference data for microwave materials has been long recognized.[55]

Antennas

For antennas, new measurement methods are needed to provide greater accuracy, improved susceptibility to automation, broader frequency coverage, and better control of the radiation pattern of antennas. Such measurement capability will enable antennas to deliver stronger signals to receiving stations, to work in close proximity to other antennas without mutual interference, and to serve applications requiring greater sensitivity and versatility. The phased-array antenna is a key example of the type of powerful modern antenna requiring improved measurement support.

The measurements required fall into two classes: (1) performance measurements, and (2) materials measurements. In particular, improved measurement support is needed for performance measurements

Table 62
Measurements for Electronic Devices for 1-100 Gigahertz

Measured Quantity	Description
Power transfer measurements	
power	determines power levels throughout a microwave system
impedance	enables components to be interconnected compatibly, with minimum power loss
attenuation	determines power loss within a component
Signal fidelity measurements	
noise	enables minimizing electronic noise to optimize information throughput and to reduce errors in transferring data in a microwave system
waveshape	enables evaluating signal quality throughout a system and determining performance of signal sources, amplifiers, modulators, detectors, and electronic switching devices
Materials measurements and data	
permittivity	determines the electrical properties of the materials
permeability	determines the magnetic properties of the materials
uniformity	determines the degree of uniformity (homogeneity) of materials properties throughout the material
anisotropy	determines the degree of uniformity of materials properties as a function of angular orientation

for 26-100 gigahertz and for materials measurements for 1-100 gigahertz. The specific quantities requiring improved measurement support are shown in Table 63.

Microwave materials measurements are needed for materials used for protective covers (lenses and radomes) for antennas. The microwave properties of these materials are not well understood, particularly above 10 gigahertz, yet those properties significantly affect the performance of antennas and the success and cost of design and manufacturing.

Integrated Components for 1-100 Gigahertz

Increasingly, microwave systems will be built in integrated form to reduce cost, size, and weight and to ease access to the higher frequencies which are not easily accessed by relatively bulking individual components. Integration will take two principal forms:

> *Integrated electronic circuits*: integration of microwave electronic devices, like signal sources and waveguides, onto a common substrate along with associated optoelectronic devices

Table 63
Measurements for Antennas for 1-100 Gigahertz

Measured Quantity	Description
Performance measurements	
gain	determines how successfully an antenna can focus its power in the forward direction, toward the intended receiving antenna
pattern	determines focusing performance of an antenna in all directions, essential for illuminating the desired receiving area and for reducing stray radiation that can interfere with adjacent systems
strength	determines strength of the microwave beam from a transmitting system
polarization	determines a special property of an antenna that allows it to send two non-interfering signals on the same frequency to double information handling capacity
bore sight	determines how accurately an antenna radiates in the expected direction
natural signal source characterization	enables use of the sun, moon, key planets, and stars as natural signal source standards for evaluating and maintaining the performance of earth terminals and satellite microwave systems while in service
Materials measurements and data	
permittivity	determines the electrical properties of the materials
permeability	determines the magnetic properties of the materials
uniformity	determines the degree of uniformity (homogeneity) of materials properties throughout the material
anisotropy	determines the degree of uniformity of materials properties as a function of angular orientation
transmissivity	determines how well the material passes microwave power
reflectivity	determines the amount of microwave power reflected by the material

> *Integrated antennas*: integration of the elements of an antenna onto a common substrate along with associated electronic and optoelectronic devices

Integration poses difficult measurement challenges for several reasons:

> *Minimal measurement ports*: Input and output ports of electronic devices within integrated circuits are not readily accessible. In individual components those input and output ports provided convenient locations for attaching measurement devices.
>
> *Measurement probe sensitivity*: Circuit elements are smaller and closer to each other, so measurement probes are more likely to disturb the circuits that they measure, especially at the high frequencies at which microwave circuits operate.
>
> *Materials sensitivity*: The performance of circuits made in integrated form is even more sensitive to materials properties than the performance of circuits made from individual components, and on a microscopic spatial scale.
>
> *Multiple technologies and interfaces*: The incorporation of optoelectronic technology in microwave integrated circuits and antennas leads to an especially complex mix of technologies and interfaces.

These differences necessitate new definitions for measured quantities, new measurement methods, and new measurement reference standards to support those measurement methods.

Integrated Electronic Circuits

The integration of microwave electronic devices is pursued for reasons similar to those that motivated the integration of conventional semiconductor devices operating at lower frequencies. However, the structures required, the materials employed, and the types of devices used may all differ. Also, microwave integrated circuits operate at 10 to 1000 times the frequency of most semiconductor integrated circuits, making the performance of the microwave circuits much more difficult to measure. In particular, they are much more sensitive to the presence of measurement probes.

The integration of microwave electronic circuits presently really takes place at two levels. The first is the development of individual microwave integrated circuits. The second is the interconnection of several individual integrated circuits within an appropriate mounting, or electronic "package". This interconnection process requires fabricating precision, short, miniature, so-called printed transmission lines. These close interconnections are necessary for several reasons: (1) to minimize the time delays associated with signal travel from one integrated circuit to another, which can easily become too long relative to the speed of these circuits; (2) to provide perfect electromagnetic interconnections, which can transmit signals without attenuating, reflecting, or distorting them; and (3) to keep overall component sizes small. One example of an electronic "package" is the multichip module (MCM) which contains an assembly of integrated circuits with necessary interconnections and external connections.

At present there is a need both for improved understanding of existing measurement methods and for the standardizing of the use of those methods. There is also a need for new measurement methods and for measurement reference standards to assure the accuracy of those methods.

The specific quantities requiring improved measurement support are shown in Table 64. Nearly all of these quantities are important to individual electronic components, too; but the measurement methods required for integrated circuits are quite different. Because of the small scale of integrated circuits and the proximity of all of their components, special measurement methods are required to access the internal workings of the integrated circuits and to prevent interference with their operation during the measurement process. If some interference occurs, it must be minimized. Corrections for any remaining interference must be made based on theoretical calculations. Special contacting electrical measurement techniques and non-contacting optical measurement techniques must be developed to accomplish these aims. A key goal is to enable testing individual chips at full operating speed prior to their being cut from the larger wafer on which they are fabricated. [The steps employed in the fabrication of integrated circuits are described in Chapter 4, Semiconductors, beginning on page 53.]

Improved measurement methods for the quantities in Table 64 will enable addressing a host of key problems, including determining the characteristics of the interconnecting transmission lines: propagation properties (pulse degradation), degree of mutual coupling (transfer of energy), and junction characteristics.

Lack of the needed measurement capability for microwave integrated circuits has hampered the design of integrated circuits and the development of fabrication processes for making them. For example, existing measurement capability does not adequately support the development of the mathematical models of microwave devices and circuits needed for efficient design. Measurement problems in microwave integrated circuits are so critical that manufacturers estimate that measurement costs continue to represent a disproportionately large fraction of product cost. In some cases, manufacturers are finding that they cannot produce highly complex microwave integrated circuits that meet all of their performance specifications with high yield.[56]

Similarly, lack of critically evaluated data on the microwave properties of materials is severely impairing the design and manufacture of integrated circuits. The development of such data will require new measurement methods for materials properties and new measurement reference standards to assure the accuracy of those methods. The materials data are essential for systems used in automated manufacturing and computer-aided design since these systems depend on materials of known and uniform properties. As noted above, most existing data on the microwave properties of materials are based on measurements made many years ago on conventional materials at frequencies below 10 gigahertz. Few data exist for frequencies above 10 gigahertz, and practically no data exist for newer materials.

The materials properties of interest must be measured as a function of temperature and humidity to simulate true conditions when in use. They must be measured as a function of composition to support design and manufacturing processes. They must be measured as a function of frequency to support wideband (frequency flexible) applications. They much be measured for spatial uniformity (homogeneity) to establish quality. As a key example, improved reference data are needed for the properties of the thin dielectrics whose performance is critical to all printed transmission lines.

Integrated Antennas

Integrated antennas are one of the most promising of emerging microwave technologies. They can be formed on a surface that is flat or curved. They contain embedded metallic radiating elements that

Table 64
Measurements for Integrated Electronic Circuits for 1-100 Gigahertz

Measured Quantity	Description
Power transfer measurements	
power	determines power levels throughout a microwave system
impedance	enables components to be interconnected compatibly, with minimum power loss
attenuation	determines power loss within a component
Signal fidelity measurements	
noise	enables minimizing electronic noise to optimize information throughput and to reduce errors in transferring data in a microwave system
waveshape	enables evaluating signal quality throughout a system and determining performance of signal sources, amplifiers, modulators, detectors, and electronic switching devices
Materials measurements and data	
permittivity	determines the electrical properties of the materials
permeability	determines the magnetic properties of the materials
uniformity	determines the degree of uniformity (homogeneity) of materials properties throughout the material
anisotropy	determines the degree of uniformity of materials properties as a function of angular orientation
transmissivity	determines how well the material passes microwave power
reflectivity	determines the amount of microwave power reflected by the material

look much like the wiring patterns in integrated circuits. Integrated antennas can contain thousands of elements, especially when configured as phased-array antennas.[57] Integrated antennas can be very large (many meters in size) or very small (a few centimeters in diameter, fabricated directly on the surface of a single semiconductor wafer). Integrated antennas can contain electronic devices built right into them. These devices serve as small transmitters and receivers. Each transmitter or receiver may be associated with individual radiating elements or with groups of elements within the antenna. The elements in an integrated antenna can be controlled by optoelectronic devices.[58] The elements can contain built-in signal processing circuits.[59] Some new antenna designs are so complicated that they cannot be economically constructed without integration.[60]

Integrated antennas offer several advantages over conventional antennas:

> *Reduced size, cost, and weight*: Embedding the elements in an insulating substrate reduces the structural complexity of an antenna considerably.
>
> *Adaptable shapes*: Integrated antennas can be shaped to conform to the skin of a vehicle, such as an airplane, satellite, or spacecraft, and thus can meet the special aerodynamic and structural requirements of those vehicles. They can also be shaped to conform to the surfaces of buildings for aesthetic or practical reasons.
>
> *Sophisticated designs*: Integrated antennas can be implemented as phased arrays, the most versatile of emerging antenna types. Phased arrays have capabilities beyond those of familiar dish antennas. Phased arrays can transmit and receive in many directions through electronic changes, without being physically moved. Further, phased arrays can deliver beams of different shapes also through electronic changes. Both of these capabilities can be exercised instantly, enabling, for example, a single phased-array satellite antenna to serve many different geographical regions on a time-shared basis by sending its signal to one location after another in rapid succession.

Integrated antennas can be driven by multiple small low-power transmitters, built right into them. They do not require the traditional single high-power transmitter. The use of multiple transmitters eliminates the need for a complex network of microwave waveguides within the antenna to carry microwave power to the antenna elements from a central transmitter. The result is a considerable reduction in weight. Also, because the transmitters for the individual elements are low power, they can employ low cost semiconductor signal sources, rather than the expensive and often short-lived microwave tubes that are required by high-power central transmitters. Integrated antennas may be controlled with optical signals fed through optical waveguides. Optical waveguides are desirable because they are non-interfering; thousands of such signals will be required to control a complex integrated antenna.

To support integrated antennas, new measurement capability is needed for overall antenna performance and for the performance and interaction of individual antenna elements within integrated antennas. New measurement methods and reference materials data are needed for the materials from which antennas are made. The specific quantities requiring improved measurement support are shown in Table 65. The new measurement approaches required will be especially challenging for several reasons:

> *No single input/output port*: Integrated antennas do not have a single point of microwave power input or output which can provide a reference for measurements. New definitions for quantities such as gain and noise will be necessary, along with new measurement approaches.
>
> *Different transmit/receive performance*: Integrated antennas will not perform exactly the same way in transmission and in reception, as conventional antennas do, so separate measurements must be made for each of these modes. The differences arise from the fact that different built-in electronics will be activated for reception versus transmission. Also, in some cases, different radiating elements within the antenna will be used for each mode.

Large size: Integrated antennas will often be physically large. They will also often be "electrically large". That is, they will be large relative to the wavelength (the distance from one ripple to another in a traveling microwave beam). Because of these "large" sizes, new measurements will be needed for determining the radiation pattern. These new methods must enable present-day antenna measurement ranges to cope with antennas that are far too complex for use of traditional measurement approaches. So-called quasi-far-field measurement techniques will be required.

Element interactions: Integrated antennas are composed of many elements, so the interactions among these elements will greatly affect performance and will require a variety of measurements sensitive to the relationships among them.

As with integrated circuits, materials measurements and data will be very important. Materials properties must be measured as a function of temperature and humidity to simulate the environment in which they will be used. They must be measured as a function of composition to support design and manufacturing processes. They must be measured as a function of frequency to support wideband applications. They must be measured for spatial uniformity (homogeneity) to determine quality.

Integrated Components for 100-1000 Gigahertz

New measurement capability must be developed for the 100-1000 gigahertz region as a longer term goal. Particularly important is the range 100-300 gigahertz in which considerable experimental development is already taking place.

The new measurement capability will support industrial research and development efforts aimed at exploiting the frequency space and special properties of the higher frequencies, with a view to discovering a broad spectrum of practical applications. Particularly attractive are the prospects for even smaller antennas, higher resolution for local radar and possible passive imaging systems, greater versatility for local communications and radar systems, and ultrafast speeds for signal processing.

Most of the measurement capability needed is for integrated components, so they are the focus of the subsections that follow. However, some measurement development for individual components will also be needed.

Integrated Electronic Circuits

The key parameters that must be measured are the same as those specified for the 1-100 gigahertz range above, but the measurement methods required will differ. The differences are caused by the following factors:

Quasi-optical behavior: The frequencies in the 100-1000 gigahertz range, particularly at the high end, take on the behavior of light, giving rise to new materials problems, new requirements for component designs, and higher levels interface complexity.

Smaller geometries: The characteristic sizes of circuit elements will be smaller at these frequencies.

Table 65
Measurements for Integrated Antennas for 1-100 Gigahertz

Measured Quantity	Description
Overall performance measurements	
gain	determines how successfully an antenna can focus its power in the forward direction, toward the intended receiving antenna
pattern	determines focusing performance of an antenna in all directions, essential for illuminating the desired receiving area and for reducing stray radiation that can interfere with adjacent systems
polarization	determines a special property of an antenna that allows it to send two non-interfering signals on the same frequency to double information handling capacity
bore sight	determines how accurately an antenna radiates in the expected direction
noise	determines the noise contributed by the integrated antenna which now contains the transmission and reception electronics and must be evaluated as a whole
Antenna element measurements	
diagnostics	determines which elements of the antenna are not working correctly
coupling	determines the effects of the coupling between closely spaced elements on antenna performance
phase	determines how the signals from the individual elements track in relation to each other
strength	determines how well the strengths of the signals from the individual elements track with each other.
Materials measurements and data	
permittivity	determines the electrical properties of the materials
permeability	determines the magnetic properties of the materials
uniformity	determines the degree of uniformity (homogeneity) of materials properties throughout the material
anisotropy	determines the degree of uniformity of materials properties as a function of angular orientation
transmissivity	determines how well the material passes microwave power
reflectivity	determines the amount of microwave power reflected by the material

Higher sensitivity to measurement probes: The sensitivity of integrated circuits to the perturbing presence of measurement probes will be greater at these frequencies.

Higher sensitivity to materials properties: The sensitivity of integrated circuits to materials properties will be even greater at these higher frequencies.

In this 100-1000 gigahertz domain of frequency, development of non-contacting optical methods, rather than contacting electrical methods, for measuring microwave parameters within integrated circuits will be especially important to prevent disturbing the circuits being measured.

The full spectrum of materials properties cited above for integrated circuits operating in the range of 1-100 gigahertz must be addressed for this new frequency range of 100-1000 gigahertz where the properties will differ considerably.

Measurements supportive of microwave signal processing electronics will be especially important to realize the promise that they offer for extraordinary processing speeds.

Integrated Antennas

New measurement capability for antennas for 100-1000 gigahertz is also needed, again with the earliest emphasis on the domain of 100-300 gigahertz. As for integrated electronic circuits above, special measurement methods will be needed to cope with quasi-optical behavior, smaller geometries, and higher sensitivity to measurement probes and to materials properties. In addition, measurements for coupling among antenna elements, and between these elements and other circuit elements, will prove very important and challenging at these frequencies.

ENDNOTES

1. Ronald K. Jurgen, "Smart Cars and Highways Go Global", *IEEE Spectrum*, p. 26 (May 1991).

2. Howard Bierman, "Personal Communications, and Motor Vehicle and Highway Automation Spark New Microwave Applications", *Microwave Journal*, pp. 36-38 (August 1991).

3. Howard Bierman, "Personal Communications, and Motor Vehicle and Highway Automation Spark New Microwave Applications", *Microwave Journal*, p. 26 (August 1991). Edward E. Reinhart, "Mobile Communications", *IEEE Spectrum*, pp. 27-29 (February 1992).

4. Howard Bierman, "Personal Communications, and Motor Vehicle and Highway Automation Spark New Microwave Applications", *Microwave Journal*, p. 28 (August 1991).

5. "Facility bypass networks" are a common corporate application. They bypass local public telephone networks to achieve one or more of these aims: "reduce costs, improve service quality, and offer greater security, reliability, and service flexibility". They are usually implemented with microwave systems since the installation of custom cable systems is too expensive and far less flexible. The Department of Commerce indicated, in a review of studies of the subject, that "16-29 percent of large-volume telephone company customers are bypassing their local telephone companies". "End-to-end facility bypass networks" bypass both local and long distance lines of public telephone networks. *1988 U.S. Industrial Outlook*, U.S. Department of Commerce, International Trade Administration, p. 33-3 (January 1988).

6. Ronald K. Jurgen, "Smart Cars and Highways Go Global", *IEEE Spectrum*, p. 28 (May 1991).

7. Howard Bierman, "Personal Communications, and Motor Vehicle and Highway Automation Spark New Microwave Applications", *Microwave Journal*, p. 32 (August 1991).

8. Howard Bierman, "Personal Communications, and Motor Vehicle and Highway Automation Spark New Microwave Applications", *Microwave Journal*, p. 26 (August 1991). Howard Bierman, "Microwave Applications Range from Under the Soil to the Stratosphere", *Microwave Journal*, p. 39 (November 1990).

9. John Kesselring, "Microwave Clothes Dryers", *EPRI Journal*, pp. 34-36 (June 1992).

10. Howard Bierman, "Personal Communications, and Motor Vehicle and Highway Automation Spark New Microwave Applications", *Microwave Journal*, p. 36 (August 1991). Howard Bierman, "Microwave Applications Range from Under the Soil to the Stratosphere", *Microwave Journal*, p. 32 (November 1990).

11. The geostationary orbit, in which satellites can maintain a stationary position over a given point on the earth, has a radius of 22,600 miles. Since the earth has a radius of 4000 miles, these satellites fly at an altitude of 18,600 miles, or 29,760 kilometers, above the surface of the earth.

12. Data on power levels was provided by the Federal Communications Commission (October 1992).

13. "Lasers", *Fiber Optic News*, page 4 (April 4, 1988). The potential for microwave electronics in fiber optic systems is further discussed by Jeff Montgomery and John Ryan, "Leap Seen for Microwave Fiber-Optics", *Microwaves & RF*, Vol. 26, No. 6, p. 53 (June 1987), and by Robert Olshansky, "Promising Merger of Microwave and Lightwave", *Lightwave*, p. 25 (January 1988).

14. M. Lawrence, "Optics Illuminate Next-Generation Radar Systems", *Microwaves & RF*, p. 165 (June 1988). Jeff Montgomery and John Ryan, "Leap Seen for Microwave Fiber-Optics", *Microwaves & RF*, p. 53 (June 1987).

15. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 30-11 (January 1992).

16. *U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 30-7 (January 1992).

17. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 30-4 (January 1992).

18. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 30-3 (January 1992).

19. Edward E. Reinhart, "Mobile Communications", *IEEE Spectrum*, p. 28 (February 1992).

20. "The Evolving Symbiosis of Optical Fiber, Satellites", *Washington Post*, pp. H-1 and H-5 (August 21, 1988). Also, George Heiter of AT&T Bell Laboratories, "Personal View: Systems Integration", *Microwaves & RF*, p. 126 (May 1988).

21. Those definitions are elaborated in Appendices 2 and 3. Instead such appliances are thought of as part of electrical applications products and thus are reflected in a representative list of such products in Table 161 on page 445 at the end of Appendix 3.

22. Microwave equipment shipments are not tracked as an independent data category in the Standard Industrial Classification (SIC) System of the U.S., so the exact levels are difficult to determine. The estimate shown is from "Commercial Microwave for the 21st Century", ElectroniCast Corporation (San Mateo, CA), page 2 and attached Table 5.2 (October 1992). That document makes no explicit reference to the use of the SIC data of the U.S. or to U.S. trade data, so its foundation in that data could not be established.

23. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, pp. 30-1 to 30-2 (January 1992).

24. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, pp. 30-1 (January 1992).

25. *Aviation System Capital Investment Plan*, Federal Aviation Administration, U.S. Department of Transportation, p. 1-0-14 (December 1991).

26. Tekla S. Perry, "Special Report: Air Traffic Control; Improving the World’s Largest, Most Advanced System", *IEEE Spectrum*, p. 23 (February 1991).

27. Information provided by the FAA. The system is described in the *National Airspace System Plan*, Federal Aviation Administration, U.S. Department of Transportation (June 1988). A discussion of the new system appears in "Why Tomorrow’s Skies will be a Whole Lot Safer", *Business Week*, pp. 99, 102 (September 5, 1988). The new "Microwave Landing System" (MLS) will replace the present "Instrument Landing System" (ILS) used since the mid-1940s. The first MLS has been installed as a demonstration unit. Compared to ILS, MLS will support landings from multiple directions at different speeds and angles of descent, at a faster rate, and with less susceptibility to surrounding terrain, weather, and radio interference. The International Civil

Aviation Organization wants microwave systems installed at 200 airports by 1998. "Microwave Landing System Installed", *Electronic Engineering Times*, p. 48 (April 17, 1989).

28. Bob Brewin, "Dorfler Plays FAA's 'Satellite Advocate'", *Federal Computer Week*, p. 22 (July 20, 1992).

29. *Sensor and Instrumentation Markets 1992-1995,* Richard K. Miller & Associates, Inc., p. 303 (1992).

30. "FAA Speeds GPS Approval, 1993 Operations Planned", *Aviation Week & Space Technology*, p. 30 (October 19, 1992). The article does not indicate if the U.S. or the world market is being characterized.

31. Information provided by Colwell-Kirtland International (Sunnyvale, California) as obtained from that organization's study *GPS Industry Analysis* (1991). The U.S. market for 1991-1996 is expected to have these segments: $2.57 billion for vehicle tracking and navigation; $1.4 billion for marine navigation; $1 billion for military (currently authorized); $0.852 billion for aviation and navigation; $0.5 billion for differential user products; and $0.395 billion for surveying and mapping.

32. Robert Langreth, "Earth Observing System: Eyes on the Earth", Popular Science, p. 44 (July 1992).

33. "Commercial Microwave for the 21st Century", ElectroniCast Corporation (San Mateo, CA), Figure 2.2 (October 1992). This is the same reference used for the value of North-American shipments in footnote 22.

34. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 30-2 (January 1992).

35. "Digital HDTV for Japan?", *Television Digest*, Vol. 32, No. 26, p. 14 (June 29, 1992).

36. Edward F. Reinhart, "Broadcasting", *IEEE Spectrum*, pp. 24-26 (February 1992).

37. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 30-3 (January 1992).

38. Edward E. Reinhart, "Mobile Communications", *IEEE Spectrum*, pp. 28-29 (February 1992).

39. Andrew D. Lipman, "Leaping from the Barricades", *IEEE Communications Magazine*, pp. 34-36 (June 1992). "In the Matter of Redevelopment of Spectrum to Encourage Innovation in the Use of New Telecommunications Technologies", Federal Communications Commission Report No. FCC 92-437, p. 1 (October 16, 1992).

40. *U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 30-10 (January 1992).

41. Jack Taylor, "Toward a Wireless World", *Electronic Engineering Times*, p. 80 (November 4, 1991).

42. Howard Bierman, "Personal Communications, and Motor Vehicle and Highway Automation Spark New Microwave Applications", *Microwave Journal*, p. 40 (August 1991).

43. Ronald K. Jurgen, "Smart Cars and Highways Go Global", *IEEE Spectrum*, p. 26 (May 1991).

44. Ronald K. Jurgen, "Smart Cars and Highways Go Global", *IEEE Spectrum*, p. 28 (May 1991).

45. Jack Taylor, "Toward a Wireless World", *Electronic Engineering Times*, p. 81 (November 4, 1991).

46. Stephanie Rynas, "Silicon MMICs Add Low-Cost Power to Wireless Systems", *Microwaves & RF*, p. 199 (May 1992). Promising performance to 8 gigahertz for silicon has already been reported. Charles Huang, "MMICs, Move Into New Marketplaces", *Microwaves & RF*, p. 136 (September 1992).

47. The geostationary orbit lies over the equator and is about 23,000 miles high. It is the only orbit which permits satellites to revolve about the earth in such as way that they remain positioned over the same point on the earth all of the time. This "stationary" position enables them to provide continuous service to the locations beneath them. To place more satellites in this orbit, the spacing between them must be reduced, increasing the problems of mutual interference with adjacent satellites.

48. Antennas with low side lobes have minimal radiation off to the sides. Such antennas reduce mutual interference between adjacent satellites in orbit or between adjacent ground stations on earth.

49. *The PNMS Report - Microwave Metrology in the U.S.A.*, Committee to Promote National Microwave Standards, an ad hoc committee sponsored by the IEEE Microwave Theory and Techniques Society (November 1987).

50. *Report by the National Measurement Requirements Committee*, National Conference of Standards Laboratories, Sections 1 and 3 (Revision 2, April 1987). The findings of the original survey were published in 1983 and were updated in the 1987 revision and again in 1989 in a report of the same name (NMRC 89-01, January 1989).

51. Jim Fitzpatrick, "NBS and Microwave Metrology -- Growing Concern within the Industry", *Microwave System News*, p. 43 (May 1984). John L. Minck, "A (Modest) Proposal to Establish a National Bureau of Microwave Standardization", *Microwave System News*, pp. 67-71 (May 1986). "Metrology Funds Needed", *Microwaves & RF*, p. 31 (April 1988). John Minck, "Some Significant Things That Have Happened to RF in the Last 10 Years", *RF Design*, pp. 29-34 (October 1988). Ron Schneiderman, "No Funds for MMIC Metrology", *Microwaves & RF*, p. 76 (December 1988).

52. John Minck, "Some Significant Things That Have Happened to RF in the Last 10 Years", *RF Design*, pp. 29-34 (October 1988).

53. Phased array antennas contain many radiating elements. By controlling the strength and phase of the signal emitted by each element, the array can be made to produce diverse and well controlled radiation patterns to serve special needs. For example, phased array antennas can be designed to illuminate receiving areas of diverse shapes, or to reduce of unwanted side lobes (radiation off to the side of the antenna) that could interfere with adjacent microwave systems. The "phase" of an element determines the moment at which the alternating microwave signal fed to it goes through its maximum value, relative to the corresponding moments for the signals fed to the other elements.

54. NIST surveyed industry’s measurement requirements for microwave materials and described the urgent needs in its "Summary Report on Measurements and Standards Requirements for Materials Used in Electromagnetic Applications: Results of a Limited Survey" (February 1985). Materials measurements affect everything from substrates, to protective antenna covers (radomes), to microwave integrated circuits. For example, better materials measurements are essential to the success of the automated design and manufacturing processes that produce microwave integrated circuits; these circuits, in turn, can cut the cost of microwave electronics by factors up to 100.

55. NIST conducted an interview survey with industry representatives in 1985.

56. Jeff D. Montgomery, "Hybrid MIC North American Markets", *Microwave Journal*, p. 32 (April 1989).

57. The integrated phased array will be a powerful type of antenna. Its radiation pattern and direction of radiation will be electronically controlled, without physical motion of the antenna. The control is accomplished by adjusting the strength and phase of the signal radiated by each element within the antenna. In some configurations, each element of the antenna will have its own transmitter and receiver. In other configurations, each element will have its own phase shifter but will share its receiver or transmitter with several other elements. In still other configurations, groups of elements will have a single phase shifter; and many groups, or even all groups, will share a common transmitter or receiver.

58. Optical waveguides embedded in a phased array integrated antenna may be used to control the relative strengths and phases of the signals transmitted by the individual elements of the antenna. Relative to metallic waveguides, which would otherwise be used for such control, the optical waveguides are lighter and will not interfere as much with the radiation pattern of the antenna.

59. For example, the information received by each element of an integrated receiving antenna may be converted to digital form by its own semiconductor electronics, located right at the antenna element. The resulting digital data from each antenna element can then be brought out of the antenna on light beams carried by optical waveguides. This approach results in a much lighter antenna than possible when microwave signals themselves must be brought out through multiple metallic waveguides.

60. The individual faces of an active array may contain as many as 5000 active transmit/receive devices, and thus active arrays may not be practical to implement in any form other than an integrated form.

CHAPTER 8

LASERS

November 1992

Chapter 8

LASERS

SUMMARY

Lasers are the single most important component for a broad range of new commercial products based on powerful lightwave (or photonic) technology. These products use light to serve both energy and information functions. In the world market, laser sales for individual commercial applications have begun to surpass laser sales for research and development, reflecting the movement of lasers from the laboratory into practical use. At the same time, U.S. competitiveness in laser production appears to be declining in the face of major laser-development efforts in Europe and Japan.

The world market for lasers for all commercial applications is projected at 50 million lasers valued at $1.2 billion for 1992. Since 1987, this market has grown at a compound average rate of 25 percent per year in unit sales and 15 percent per year in dollar sales. The world market for lasers and supporting laser-optic components for both commercial and government applications, including military applications, is expected to exceed $2 billion for 1992.

Lasers serving *energy* functions are expected to represent about 53 percent of the dollar value of all lasers sold for commercial applications worldwide in 1992. The principal applications are materials processing and medicine, in that order. These applications surpassed research and development in the dollar value of lasers consumed, in 1990 and 1991 respectively. The growth of these two applications reflects major advances in manufacturing and medical treatment.

Research and development applications are expected to consume about 17 percent of the dollar value of all lasers sold worldwide in 1992. Lasers employed in research and development serve both energy and information functions.

Lasers serving *information* functions are expected to represent about 31 percent of the dollar value of all lasers sold for commercial applications worldwide in 1992. The dominant applications are communications and optical memories, in that order. Each of the remaining applications serving an information function consumes less than half of the dollar value of lasers of either of these top two. The other applications include printers, test and measurement, barcode scanners, alignment and control, and color separation.

Viewing the worldwide commercial market for lasers based on unit sales, rather than on dollar volume as above, the applications serving information functions consume a much larger number of lasers than the applications serving energy functions. On this basis, optical memories alone consume 80 percent of all lasers sold; laser printers consume 14 percent. All other applications, serving both information and energy functions, share the remaining 6 percent.

The world market for the products made with lasers is more difficult to assess because market data are available for so few of the many applications of lasers. However, for those few applications (optical memories, laser printers, and barcode scanners), the world market for 1992 is likely in excess of $16 billion.

Unfortunately, as lasers have grown in commercial prominence, the U.S. has progressively lost market share in the world commercial market for lasers. In the five years from 1982 to 1987, the U.S. share of this market fell from 75 percent to 38 percent. Japan and Europe picked up the entire difference in roughly equal amounts. Similarly, the U.S. is losing market share in the world commercial market for supporting laser optics. From 1984 to 1988, the U.S. share of that market dropped from 64 percent to 45 percent. Laser optics include components such as lenses and mirrors that are critical to laser performance. In the world market for products made with lasers, the U.S. market share for 1990 is even lower, 26 percent, while Japan has 38 percent, Europe has 42 percent, and the rest of the world has 4 percent. The intent of other nations is clear; both Japan and Europe have launched major programs to advance their competitiveness in the laser technology required for commercial applications.

The economic implications of success in the commercial markets for lasers and laser-based products will continue to be highly significant. Laser applications in materials processing are at the leading edge of modern manufacturing in major industries such as automobiles and semiconductors. Lasers may also play increasingly important roles in the chemical industry. Laser advances in medicine already provide significant support for the largest U.S. service industry. Laser-based systems for optical storage, printing, and scanning are key products in the paperless office. Laser-based optical-fiber communications systems are providing the highest information capacity available from a cable technology. Thus lasers are intimately tied to modern manufacturing, medicine, and the information age.

Improving U.S. competitiveness in lasers and laser-based products for commercial applications will require the development of new lasers with new capabilities, higher performance levels, improved quality, and reduced cost. Needed are lasers with a greater diversity of frequencies (including the visible and ultraviolet frequencies), higher power levels, better beam control, and high levels of reliability in service. These advances are needed to meet the diverse requirements of the broad range of commercial applications that will employ lasers.

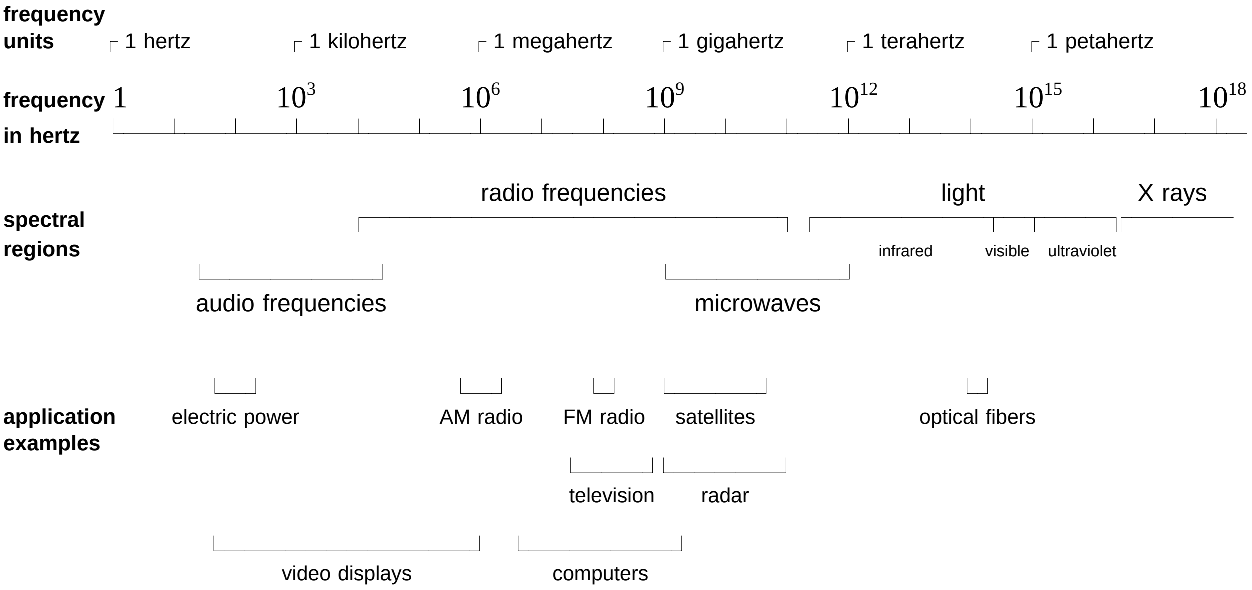
U.S. industry’s ability to achieve all of these aims for improved competitiveness is highly sensitive to the level of available measurement capability. Improved measurement capability is needed to support: (1) research and development of new lasers, laser optics, and the products that apply lasers; (2) the development of manufacturing processes for reducing product cost and for controlling quality and reliability; (3) the development of voluntary standards by industry; (4) the specification of products to facilitate market transactions including entry to the domestic and international markets; and (5) the safe use of lasers. In all, improved measurement capability, in one form or another, is needed for more than 30 critical measured quantities, all important to marketplace success. Among the needs are new measurement methods with higher accuracies, broader frequency coverage, and broader power coverage, compared to current capabilities. Also needed are new national measurement reference standards to assure the accuracy of the measurement methods. Finally, new materials reference data are required to support the design and manufacturing of lasers and of supporting optical components.

LASER LIGHT

Light is electromagnetic radiation, much like radio waves; but light is much higher in frequency. Like other radio waves, light travels like waves of water moving across a pond. Those waves exhibit alternating peaks and valleys as they pass by. Waves of light do, too, even though the peaks and valleys cannot be seen. The distance from one peak to the next is the *wavelength* of the light. The number of peaks per unit time that passes a given point is the *frequency* of the light. The frequency of the light determines its color. The human eye can see only a narrow range of frequencies that vary about a factor of two from red at the low end to violet at the high end. Light of other frequencies, both higher (ultraviolet) and lower (infrared), exists even though it is not visible. The position of light in the frequency spectrum is shown in Figure 5.

Lasers can generate a very wide range of frequencies. Some of those frequencies are thought of as light and others lie outside of that range and extend downward into the microwave region and upward into the X-ray region.

Figure 5
Overview of the Frequency Spectrum



In ordinary light, like that from the sun or from light bulbs, many different lightwaves travel along together without any special relationship to each other. Some of the lightwaves are a bit ahead of, or behind, those along side them. Laser light, in its ideal form, is different for three reasons. First, laser light has one frequency, that is, one color. Second, all of the waves in laser light can be made to travel in a single direction. Third, when travelling in a single direction, all of the waves are in lock step with those along side them, like soldiers marching in perfectly aligned rows. These three characteristics of laser light (single color, single direction, lock step nature) constitute the property of "coherence". Coherence gives laser light these special capabilities:

(1) Laser light can be emitted in a highly concentrated form and can be focused to a small point.[1]
(2) Laser light can interact with materials selectively.
(3) It can "interfere" with itself in special ways.

The first capability enables lasers to launch light into the very small core of an optical fiber. The second capability enables lasers to stimulate specific chemical reactions. The first and second capabilities together enable lasers to cut materials[2] or to remove specific unwanted materials by ripping individual atoms from them (ablation). The third capability enables lasers to detect small imperfections in the shapes of objects and to produce three-dimensional images suspended in air (holography).

PRINCIPLE OF LASER OPERATION

All lasers rely on the same basic principle for operation. That principle is embodied in the word *laser* itself, which is an acronym for light amplification by stimulated emission of radiation. Inside a laser, energy is stored in a physical medium. The energy is released in a rush when a weak beam of light of a specific color passes through the medium. As the energy is released, it produces additional light of the same color. The additional light joins with the light in the weak beam, thus amplifying it.

The release of the additional light occurs in a coherent manner. This means that all of the peaks and valleys in the wave of the original weak beam are strengthened without being disturbed from their perfect regularity. Using the analogy to marching soldiers, this process is much like adding new soldiers to each row of soldiers marching by without disturbing the perfect straightness and spacing of the successive rows. Thus, in a laser, a weak beam of light is amplified into a strong one by the stimulated emission of additional light of exactly the same nature.

The laser medium is usually in the form of a rod, a tube, or a rectangular solid. Special mirrors, or other reflecting surfaces, are placed at opposite ends of the medium.[3] The mirrors reflect laser light back and forth through the medium. On each pass through the medium, the strength of the laser light is amplified by further stimulated emission. The mirror at one end is partially transparent to allow a portion of the laser light to be released continuously for use. Sometimes the mirror is blocked from the inside by a special material that can be "switched" from a transparent state to an opaque state. This special material enables the laser to be turned on and off so that it can produce pulses instead of continuous output. Outside a laser, other precision optical materials, such as additional mirrors and lenses, are used to direct the light and to condition it for specific applications.

TYPES OF LASERS

There are several types of lasers, as shown in Table 66.[4] They can be classified by the *state* and the *structure* of the physical medium of which the laser is made. The *active part* of the medium is the part which produces light. In some lasers all of the medium is active and thus has the same structure as the medium itself. In other lasers only a part of the medium is active, and its special properties may derive from a different structure. The starred (*) entries in the table indicate the structure of the *active part* of the medium. The active part may be atomic (isolated atoms), molecular (chemically combined sets of atoms), crystalline (periodic arrays of atoms), or free electrons in a *gas*. The last column of Table 66 provides examples of specific lasers which, collectively, are important to a wide variety of applications of lasers.

Table 66
Types of Lasers

Medium			
State	Structure	Subgroup	Examples
solid state	amorphous (atomic*)		neodymium or erbium in glass
	crystalline	insulator (atomic*)	neodymium in yttrium-aluminum-garnet, titanium in sapphire
		semiconductor (crystalline*)	gallium-aluminum-arsenide
liquid state	molecular*	dye	rhodamine G
gaseous state	atomic*	neutral	helium-neon metal-vapor helium-cadmium copper
		ionized	argon
	molecular*	common	carbon-dioxide
		chemical	hydrogen-fluoride
		excimer	argon-fluoride
electron gas	electron gas*		free-electron

* structure of active part of medium

Note that industry sometimes breaks out semiconductor lasers separately from other solid-state lasers when reporting market data. Also, industry sometimes breaks out helium-cadmium lasers separately from other metal-vapor lasers when reporting market data. Both of these conventions will be followed in this chapter. Semiconductor lasers are often called *semiconductor-diode lasers* or just *diode lasers*.

The frequency of the light produced by a laser is determined by many factors. The active part of the laser medium is the most important. It determines the set of frequencies that the laser can *potentially* produce. The actual frequency produced by a laser is selected from this set by adjusting the positions and properties of the mirrors and the other optical components of the laser.

WORLD MARKET FOR LASERS AND LASER OPTICS

Available world market data for lasers and laser-based products are spotty. World market data *are* available for lasers used in commercial applications. World market data *are not* generally available for lasers used in government, including military, applications. Similarly, world market data are also not available for most laser-based products used in commercial applications. However, with the data that are available, the following discussion will show that the world market for lasers and supporting laser optics appears to be more than $2.3 billion for 1992. Further, the world market for equipment employing lasers, based on just three applications (optical memories, laser printers, and bar code scanners) is certainly over $17 billion for 1992, and likely substantially so.

Lasers

For lasers themselves, current data are available for the commercial component of the world market. The world market is estimated at $1.2 billion for 1992. The U.S. commercial market represented about 37 percent of the world commercial market in 1991.[5] Sales of lasers to the government component of the world market are more difficult to determine. However, the military component of the U.S. Government market is estimated at $412 million for 1992.[6] If the U.S. share of the world government market in 1992 is similar to the U.S. share of the world commercial market in 1991, then the world market for government applications (military plus other government) should be above $1.1 billion per year for 1992. This leads to a total world market above $2.3 billion for 1992.

The growth rate for the world market for commercial lasers is estimated at 7 percent from 1991 to 1992.[7] The growth rate for unit sales of lasers from 1991 to 1992 is estimated at 30 percent.[8] Over the five year period, the growth in unit sales of lasers for commercial applications has been 202 percent, reflecting a compound annual growth rate of 25 percent per year. The growth in dollar sales over the same period has been 103 percent, or 15 percent per year.[9] The high growth rate in unit sales reflects the rapid and broad adoption of lasers as components in other products. The substantial but lower growth rate in dollar sales reflects the emergence of low-priced lasers and the intense price competition in the world marketplace.

Laser Optics

The most important components supporting the lasers are laser optics. Laser optics create, condition, and direct laser beams. Laser optics are used both inside and outside lasers. Laser optics used inside lasers (intracavity) include mirrors, prisms, beam splitters, filters, crystals, and windows. Laser optics used outside lasers (extracavity) include lenses, mirrors, prisms, filters, and other components. The world market laser optics used inside lasers is projected at $45 million for 1992.[10] The world market for laser optics used outside lasers is projected at $161 million for 1992.[11]

U.S. INTERNATIONAL COMPETITIVENESS

Lasers

The world market shares for the consumption of lasers for 1991 are shown in Table 67. The breakdown is about the same as for 1990.[12] For 1991, the U.S. is the second largest consumer of lasers. This is a change from 1989 when the U.S. was the dominant consumer of lasers with a 45 percent share of the world market.[13]

Table 67
World Market Shares for Consumption of Lasers
(percent)

	1991
U.S.	37
Japan and Pacific Rim	23
Europe	40

The most recent data available on world market shares for laser production are shown in Table 68.[14] In the five years from 1982 to 1987, the U.S. lost half of its market share for lasers to Japan and Europe. While data on laser market shares for later years are lacking, an estimate has been made of market shares for laser systems for 1990: U.S., 26 percent; Japan, 38 percent; Europe, 32 percent; and the rest of the world, 4 percent.[15] If laser market shares track with system market shares, then these figures suggest a further erosion of the U.S. position, a major strengthening of Japan's position, and a small improvement in Europe's position. Further, these 1990 figures, when compared with the consumption data for 1991 in Table 67, suggest that the U.S. and Europe are net importers while Japan is a net exporter.

Table 68
World Market Shares for Production of Lasers
(percent)

	1982	1987
U.S.	75	38
Japan	10	27
Europe	10	30
Other	5	5
	100	100

The U.S. appears to be doing a bit better in the domestic market, at least for non-semiconductor lasers. In 1989 U.S. exports of non-semiconductor lasers were five times U.S. imports. However, "Japan supplies over 85 percent of the world market for semiconductor lasers", and the semiconductor lasers are increasingly replacing other lasers in low-power applications.[16] In optoelectronics broadly, the U.S. continues to fall behind Japan whose "optoelectronics industry grew faster than any other sector--in the country that, for a while, had the highest growth rate in the world."[17]

Japan's success is attributed to its efforts to develop its domestic market first, to reduce costs and refine its products, and then to enter foreign markets with low-priced exports.[18] Japan appears committed to maintaining its excellence in laser technology. A portion of a recent $1.6 billion Japanese program for funding basic and applied technologies was focused on lasers and electro-optics.[19] The seriousness of the Japanese challenge was recognized in 1985 in the DOC-sponsored Japanese Technology Evaluation Program (JTECH) which found that the gap between Japan and the

U.S. for the technology underlying semiconductor lasers was threatening to become permanent.[20] Comments on Japan's competitiveness in specific laser applications have been incorporated into the section "World Commercial Market by Laser Application", beginning on page 192.

Europe, too, is emphasizing laser research and the development of laser applications in its long-range plans. Through the European Economic Community, Europe is investing $400 million over five years in this field.[21] Europe is expected to have increasing success in penetrating the U.S. market for industrial lasers. In fact, U.S. consumption of European industrial lasers is expected to increase 9 percent per year through 1994.[22] U.S. exports of lasers to the European Community, as reported in June of 1991, totaled $57.5 million (presumably for the year 1990) and reflected a drop of 6 percent relative to 1989.[23]

Laser Optics

In the closely related field of laser optics, U.S. competitiveness has also been dropping rapidly. In 1984 the U.S. made 63.5 percent of high-technology laser optics. By 1988 the U.S. share had dropped to 45 percent. Quality control is cited as one of the key factors explaining the loss of U.S. competitiveness.[24]

The loss of U.S. competitiveness in laser optics is not limited to the commercial market. The military part of the government market is also affected. A key indication of this effect is that the U.S. military is turning to foreign suppliers, reluctantly but increasingly, for the electro-optic products that it needs.[25]

WORLD COMMERCIAL MARKET BY LASER TYPE

The principal types of lasers sold in the world commercial market for 1992 are shown in Table 69.[26] In some cases individual lasers are shown, such as carbon-dioxide lasers. In other cases, a whole category of lasers, such as dye lasers, is shown, in keeping with the practices of categorization common in the industry. The term *solid-state*, as used here, excludes the semiconductor-diode lasers, which are counted separately. The single most important solid-state laser is the yttrium-aluminum-garnet laser. Similarly, the term *metal-vapor*, as used here, excludes the helium-cadmium lasers, which are counted separately. The lasers are arranged in decreasing order of sales in dollars. The first four types of lasers account for 85 percent of the world commercial market on a dollar basis: carbon-dioxide, semiconductor-diode, solid-state, and ion lasers.

Expected levels of unit sales of lasers in the world commercial market for 1992 are shown in Table 70.[27] Unit sales are heavily weighted toward the two lowest cost lasers: semiconductor-diode lasers and helium-neon lasers. The semiconductor-diode lasers dominate unit sales, representing 99 percent of the nearly 50 million lasers sold worldwide for commercial purposes. Helium-neon lasers are by far the second largest seller on a unit basis, yet they represent only 1 percent of all sales.

The reduction of the semiconductor-diode laser to a commodity is one of the most important developments in laser technology in this decade. Nowhere is the impact of this development more evident than in the product category of optical memories, which in 1992 is expected to use 80.4 percent of all laser units sold in the commercial market.[28] Although the U.S. invented the semiconductor-diode laser, Japan was the first to reduce it to a cheap product (less than $10); and

Table 69
World Commercial Laser Sales in Dollars by Laser Type (1992)

	Sales ($Millions)	Percent
carbon-dioxide	323.1	27.9
semiconductor-diode	286.8	24.8
solid-state	244.4	21.1
ion	133.8	11.6
helium-neon	46.5	4.0
excimer	34.3	3.0
dye	60.7	5.3
metal-vapor	15.7	1.4
helium-cadmium	10.9	0.9
	1,156.2	100.0

Table 70
World Commercial Laser Sales in Units by Laser Type (1992)

	Sales (Units)	Percent
semiconductor-diode	49,120,820	99.12
helium-neon	402,400	0.81
ion	19,729	0.04
solid-state	5330	0.01
carbon-dioxide	4897	0.01
helium-cadmium	2895	0.01
dye	1341	0.00
excimer	550	0.00
metal-vapor	131	0.00
	49,558,093	100.00

Japan now virtually owns the consumer market for the resulting products, such as the audio compact-disk players.[29]

The growth of the market for semiconductor-diode lasers has been further accelerated by the emergence of (1) higher power diode arrays that produce watts of output versus milliwatts[30], (2) electronically steerable versions, and (3) visible diode lasers to complement earlier infrared versions[31]. In fact, visible lasers accounted for 43 percent of all diode lasers sold worldwide for commercial applications in 1991.[32] These developments have greatly broadened the potential applications of semiconductor-diode lasers.

It is noteworthy that semiconductor-diode lasers have begun substituting for helium-neon lasers in many applications, reducing product cost and broadening markets accordingly. Also, the diode lasers have become integral components in other types of lasers, particularly in atomic solid-state lasers. In the resulting so-called diode-pumped solid-state lasers, the diode lasers charge up, or pump, the active atomic element in the medium of an atomic solid-state laser, usually neodymium in yttrium-aluminum-garnet. The atomic medium then produces the light output of the laser. Relative to diode lasers themselves, the diode-pumped solid-state lasers provide a greater diversity of frequencies and a pulsed output at a higher peak power level. Compared to alternative lasers, the diode-pumped solid-state lasers offer high efficiency, small size, stable low-noise output, and desirable optical properties.[33] Thus diode lasers are improving the utility, reducing the cost, and broadening the applicability of other types of lasers, too. The diode-pumped solid-state lasers are classified under the solid-state category for the purposes of this discussion.

WORLD COMMERCIAL MARKET BY LASER APPLICATION[34]

The applications of lasers sold in the commercial market are diverse and expanding. Laser applications can be divided into two classes as shown in Table 71. Once class serves energy functions and the other serves information functions. Some applications may require elements of both classes of functions.

Table 71
Classes of Laser Applications

Classes	Functions	Examples
Energy	deliver energy	surgery, welding, drilling, lighting
Information	read information	compact-disk players, barcode readers
	generate information	medical diagnostics, environmental sensing
	transmit information	optical fiber communications, free-space laser communications

The movement of lasers from the laboratory to commercial applications is reflected in changes over the years in the dollar sales of lasers for each application compared to the dollar sales of lasers for research and development. As shown in Table 72, research and development placed first in 1989,[35] second in 1990,[36] and third in 1991.[37] Materials processing passed research and development in 1990, and medicine passed research and development in 1991. Communications and optical memories are the fourth and fifth largest applications consistently over the three-year period.

Table 72
Dominant Commercial Applications Based on Laser Sales
($millions)

1989		1990		1991	
Research and Development	216.5	Materials Processing	256.4	Materials Processing	337.6
Medicine	200.2	*Research and Development*	218.4	Medicine	233.8
Materials Processing	173.9	Medicine	209.5	*Research and Development*	190.0
Communications	76.4	Communications	112.1	Communications	111.2
Optical Memories	70.9	Optical Memories	73.8	Optical Memories	76.9

The dollar sales of lasers by application for 1992 are listed in Table 73 in descending order.[38] The percentages of the world commercial market represented are also shown. The two top applications -- materials processing and medicine -- are from the energy class, as described in Table 71. These two applications alone account for over half of dollar value of all lasers consumed worldwide for commercial applications. If we skip over research and development, then the applications from communications on downward are primarily from the information class, with the exception of entertainment which employs lasers as light sources. Within this information class, optical memories are coming on strong and, in 1992, are expected to exceed half of the sales level for research and development for the first time. Communications are expected to continue at the same level as in 1990 and 1991. After the top five applications in Table 73 -- including the two energy applications, research and development, and the two information applications -- no other application is yet half as large as these, on the basis of dollar sales.

Table 73
World Commercial Laser Sales in Dollars by Application (1992)

	Sales ($Millions)	Percent
Materials Processing	357.0	30.9
Medicine	250.5	21.7
Research and Development	191.0	16.5
Communications	111.4	9.6
Optical Memories	100.8	8.7
Printers	39.7	3.4
Test and Measurement	35.2	3.0
Barcode Scanners	26.8	2.3
Alignment and Control	24.9	2.2
Entertainment	9.9	0.9
Color Separation	9.0	0.8
	1,156.2	100.0

The applications of lasers are listed again in Table 74, this time in descending order of unit sales.[39] Optical memories are by far the dominant application on this basis, consuming 80.4 percent of the 49.6 million lasers projected to be sold in 1992. Printers are a distant second with 14.2 percent. Semiconductor-diode lasers dominate these two applications. All other applications reflect far lower unit counts. Thus information applications dominate energy applications on a unit basis.

A wide variety of types of lasers are presently employed in commercial applications, as shown in Table 75.[40] The applications are again listed in descending order of projected dollar sales of lasers for 1992, as they were in Table 73. For each application, the types of lasers employed are listed from left to right in descending order of dollar sales. It is interesting to note the principal changes that have occurred in the prominence of individual types of lasers, on a dollar-sales basis, for individual applications from 1989 to 1992. Diode lasers have become the most important type of laser for six applications, up from four in 1989. The newly included applications are barcode scanners and alignment and control; the diode lasers edged out helium-neon and helium-cadmium lasers, respectively. Also, solid-state lasers have become the most important type of laser for both medicine and research, edging out ion lasers in both cases. Improvements in the power output and reliability of solid-state lasers have been major factors in their advance.[41] Continuous power levels above 1

Table 74
World Commercial Laser Sales in Units by Application (1992)

	Sales (Units)	Percent
Optical Memories	40,001,550	80.7
Printers	7,054,614	14.2
Alignment and Control	1,035,650	2.1
Barcode Scanners	900,000	1.8
Communications	400,530	0.8
Research and Development	68,087	0.1
Test and Measurement	39,090	0.1
Medicine	31,358	0.1
Color Separation	13,820	0.0
Materials Processing	7054	0.0
Entertainment	6340	0.0
	49,558,093	99.9

kilowatt have been achieved, as has fiber-optic delivery of the beam which greatly eases the process of applying the beam to the desired locations.[42] For the remaining applications, the most important types of lasers remain unchanged from 1989.

The net result of these changes is that the number of different types of lasers most important for individual applications, on a dollar-sales basis, has been reduced from five in 1989 to four in 1990. The net change of minus one reflects the addition of solid-state lasers and the removal of helium-neon and helium-cadmium lasers, while three other types of lasers continue among the top four: carbon-dioxide, diode, and ion.

Two of the types of lasers introduced in Table 66 are not yet sufficiently significant in the commercial market to appear in market data: chemical lasers and free-electron lasers. These lasers are capable of very high output and have other special properties. They are important primarily to research activities or to military applications.

The following sections describe present and emerging commercial applications of lasers. The order of the sections reflects the order of sales of lasers in dollars in Table 73. Further comments on U.S. competitiveness, as it relates to individual applications, are also included. The world market for equipment based on lasers is available for only a few of many applications. For those few applications (optical memories, laser printers, and barcode scanners), the world market for 1992 is likely in excess of $16 billion. The details are contained in the following discussion.

Materials Processing

Materials-processing applications of lasers include semiconductor processing, materials working, and marking. Each of these is discussed below. World market data for materials-processing equipment employing lasers has not been located. However, Japan estimates that it produced $650 million of processing equipment based on lasers in 1991.[43]

Table 75
Types of Lasers Employed in Commercial Applications

Application	Types of Lasers Employed
Materials Processing	carbon-dioxide, solid-state, excimer, ion, metal-vapor, diode, helium-cadmium
Medicine	solid-state, carbon-dioxide, dye, ion, metal-vapor, excimer, helium-neon, diode, helium-cadmium
Research and Development	solid-state, ion, dye, diode, excimer, helium-neon, carbon-dioxide, metal-vapor, helium-cadmium
Communications	diode, helium-neon
Optical Memories	diode, ion, helium cadmium, helium-neon
Printers	diode, ion, helium-neon, helium-cadmium
Test and Measurement	diode, ion, solid-state, helium-neon, helium-cadmium
Barcode Scanners	diode, helium-neon
Alignment and Control	diode, helium-neon, helium-cadmium, solid-state
Entertainment	ion, helium-neon, dye
Color Separation	ion, helium-neon, diode, helium-cadmium

Semiconductor Processing

Semiconductor-processing applications include lithography, selective annealing, thin-film removal (resistor trimming), link-making and link-blowing on memory chips, vapor deposition, metal planarization, doping, hole drilling, surface inspection, etching, aligning, and photomask and screen generation, among others.[44]

Materials Working

Materials-working applications include cutting, drilling, welding,[45] and soldering. Lasers perform these tasks without contact with the material being processed and thus with no wearing of the "tool".[46] Lasers can cut cleanly with virtually no damage to surrounding material; the ablative process that they employ circumvents the drawbacks of thermal processes such as melting or vaporization.[47]

Japan has moved quickly to exploit the capabilities of lasers for manufacturing processes and, as of 1989, reportedly had five times more laser cutters in service than the U.S. Further, the Japanese auto industry has been eyeing laser welders to improve efficiency. Japan generally has been aggressively pursuing industrial applications of lasers.[48]

Europe, too, has targeted laser research for industrial applications through its Eureka Project.[49] At least one German automotive manufacturer (Volkswagen AG) has already implemented laser cutters in its automotive assembly operations.[50] European nations and the U.S. have also been looking to laser welders for their auto industries.[51]

Marking

Laser marking is applied to a broad range of products including pharmaceuticals, foods, cosmetics, circuit boards, automobiles, and many others. Laser marking is accomplished by ablating the surface layer of a material or by causing a color change in the surface layer through heating.

Laser marking has a long list of advantages over conventional methods of marking. Here are some examples: all materials can be marked; permanent, high quality marks can be produced; high resolution can be obtained; the marks can be made without contact, mechanical distortion, or contamination; and the marks can be made with high speed, low cost, and low maintenance.[52]

While laser marking is a small market, it is apparently expanding rapidly. In 1987, U.S. sales of laser marking systems were $25 million to $30 million. By 1988, fifty U.S. companies were already manufacturing marking systems.[53] More recent data have not been located.

Emerging Applications

Other applications of lasers, related to manufacturing, are being explored. They include: stereo lithography (creation of three-dimensional solid parts from liquid photopolymers without cutting or molds)[54]; surface treatments (hardening, alloying, glazing, annealing); scribing (for ceramics and glasses); laser-illuminated vision for robots, for positioning, and for inspection for quality control inspection[55]; micromachining; and laser chemistry.

The micromachining capabilities of lasers are noteworthy. Lasers can easily operate on micrometer scales.[56] For example, they can drill tiny holes through objects that are themselves as small as human hairs.

The potential of lasers for chemistry is also very important. Laser light of the proper frequency can stimulate chemical processes that otherwise would not occur or would occur too slowly to be commercially viable.

Medicine

Lasers are very prominent now in medicine, and the medical market for lasers continues to grow. Both therapeutic and diagnostic applications are involved.[57] Overall, the medical equipment and services markets that can be potentially served by lasers are enormous. The medical services industry is the largest services industry in the U.S. economy, providing an estimated $817 billion of services, supplies, and equipment in 1992.[58] Worldwide production of health-care-technology products was $70.9 billion in 1991, with just under half produced and consumed in the U.S.[59] The world market for medical equipment employing lasers is not known. However, Japan produced an estimated $40 million of such equipment in 1991, a modest amount.[60] Given that $234 million dollars of lasers were incorporated into medical equipment worldwide in 1991,[61] Japan is apparently not a major manufacturer for this application.

Diagnostic Applications

Diagnostic applications include DNA sequencing, cell separation, cell sorting (cytometry), AIDS detection, cancer detection, and others.

Therapeutic Applications

Therapeutic applications of lasers are wide ranging. Promising fields of medicine for increased use of lasers are shown in Table 76.[62] Laser techniques are attractive because they are less invasive than other techniques. Therefore, they are easier on patients, and they can shorten hospital stays. Of the 11.5 million surgical procedures performed each year in the U.S., an estimated 10 percent to 40 percent could potentially employ lasers.[63]

The economic implications of the successful implementation of laser-based therapeutic procedures are highly significant. For example, corneal sculpting,[64] if it can be realized as a practical process, has been estimated to have a potential market of billions of dollars per year for the replacement of eye glasses.[65]

Table 76
Therapeutic Applications in Medicine

Field of Medicine	Examples of Applications
ophthalmology	radical keratotomy (nearsightedness correction), corneal sculpting (near- and farsightedness correction), glaucoma, retinal disease
obstetrics	
otolaryngology	
podiatry	
gastroenterology	
dermatology	birthmark and port-wine stain removal
general surgery	
orthopedics	
plastic surgery	
urology	lithotripsy (kidney stone shattering)
cardiovascular surgery	heart treatment, photocoagulation, laser angioplasty (vaporization of fatty deposits and plaque in arteries)
colorectal surgery	
thoracic surgery	
neurology	neurosurgery
oncology	photodynamic therapy for removing cancers
gynecology	
dentistry	soft tissue removal, restorative material curing, cavities

The combining of diagnostic (information) and therapeutic (energy) functions in individual laser systems would be particularly powerful if it could be realized in the medical systems of the future. For example, in laser angioplasty, lasers could potentially identify the material that is to be vaporized, vaporize it, and monitor the progress of the vaporization -- all in one combined process.

Research and Development

Virtually every field of research uses lasers: physics, photochemistry, materials science, electronics, biology, medicine, spectroscopy, laser fusion, and many others. The U.S. share of world laser consumption for research is dropping as other countries develop their research capabilities. In 1986, the U.S. accounted for about 45 percent of the world market for research lasers, Europe for 40 percent, and Japan and the Pacific Rim countries for about 15 percent. In 1988, only two years later, Europe surpassed the U.S. and accounted for 44 percent of the world market; the U.S. share dropped to 40 percent, and Japan and the Pacific Rim countries rose to 16 percent.[66] Data for later years are not available.

Communications

Lasers are used for both optical-fiber communications and free-space communications. They bring high information capacity to both types of applications.

Optical Fibers

Low-frequency (near-infrared) semiconductor-diode lasers are the light sources for optical-fiber communications systems, particularly for long distance systems, both cross country and undersea. Optical-fiber communications are discussed in detail in Chapter 9, Optical-Fiber Communications, beginning on page 217.

Free Space

Laser systems for communicating over long distances in free space, air, and water are emerging. Lasers can link: spacecraft (including satellites) to other spacecraft; spacecraft to aircraft; spacecraft to ground[67]; and spacecraft, aircraft, and surface vessels to submarines[68]. Such laser communication systems have several advantages compared to radio frequency systems: small size and weight; high information capacity due to the high frequency of operation; and high security and high freedom from interference due to narrow beam divergence.[69] These advantages must be weighed against the disadvantages: greater weather sensitivity when penetrating the atmosphere, and higher requirements for pointing accuracy due to the very narrow beams.

Laser systems for communicating over very short distances in air are under consideration for use inside computers. Lightwaves directed from one computer circuit to another would carry information at a high rate and would replace the slower, parallel sets of conducting paths in present-day circuit boards.

Optical Memories

Optical memories are a fast growing application of lasers. Optical memories include audio compact-disk players, video players, and optical data-storage devices for computers. Based on the projected number of lasers going into optical memories in 1992 (40 million from Table 74), the market for all optical storage products is likely to be very large for 1992, consuming more laser units than any other application of lasers by a factor greater than five. Japan is preeminent in optical memories, with 1991 production estimated at $6.3 billion for optical disk equipment and $4.4 billion for optical media, as shown in Table 77.[70]

Table 77
Forecasted Production of Optical Memories by Japan (1991)

	1991 Forecast ($billions)	1990-1991 Growth (percent)
Optical Disk Equipment		
audio compact disk	4.2	17
video	1.1	-1
data storage, including:	1.0	40
compact disk read-only		
write-once		
rewritable		
	6.3	
Optical Disk Media		
audio compact disk	3.2	25
video	1.1	5
data storage (as above)	0.1	46
	4.4	
total	10.6*	18

* total does not match exactly because of effects of rounding

Audio Compact-Disk Players and Video Players

For audio compact-disk players, which are a read-only technology, Japan held more than an 80-percent share of U.S. market, on unit basis, in 1991.[71] That market was $0.8 billion in 1991.[72] The U.S. is one of the major consumers of these products. As of January 1992, an estimated 35 percent of all U.S. households owned an audio compact-disk player.[73] Based on Japan's production alone, as shown in Table 77, the world market for audio compact-disk players is at least $4.2 billion for 1991.[74]

Players combining video capability with audio compact-disk capability are now available but as of 1992 still represent a smaller market. Again, based on Japan's production alone, as shown in Table 77, the world market for video disk players is at least $1.1 billion for 1991. In the U.S., only an estimated 1 percent of all U.S. households owns them.[75]

Optical Data-Storage Devices

For optical data-storage devices for computers, including both read-only and read-write implementations, the world market was about $0.9 billion in 1990. The U.S. held only a 7.9 percent market share of the world market in that year. This market is expected to double by 1992.[76] Japan estimates its 1991 production at $1.0 billion, as shown in Table 77.

The U.S. is already a major consumer of this new technology. In 1988, the U.S. represented 73 percent of the world market for optical data-storage products. In 1993, the U.S. is expected to represent 60 percent of the world market.[77]

Optical data-storage systems offer high information density per unit area. However, compared to magnetic storage systems, the optical systems have generally had slower access times. Optical systems have not been implemented in a tape format which provides much greater area for storage than a disk format. However, optical technology offers the prospect of three dimensional storage. Optical technology will continue to compete with advances in magnetic technology. A more detailed comparison of these technologies can be found on page 103 in Chapter 5, Magnetics.

Compact-disk read-only memories (so-called CD-ROMs) were the first optical data-storage products. As their name suggests, they can be read from but not written to, so they can handle only permanently recorded data. CD-ROMs have found use in a wide diversity of applications, including those in library, financial, medical areas, aerospace, and education areas, among others. By one estimation reported in 1990, CD-ROM drives represented a world market of $0.4 billion in 1990 and are expected to reach $0.8 billion by 1993, with the U.S. consuming about half of world production on a dollar basis.[78] By another estimation reported in 1992, CD-ROM drives represent a world market of $0.3 billion in 1992 and are expected to reach $0.6 billion in 1996, with the U.S. consuming 69 percent of world production on a unit basis in 1991.[79]

Printers

Laser printers are being adopted widely. The U.S. market alone is expected to reach $7.4 billion in 1992, as which time laser printers are expected to represent just over half of the U.S. printer market.[80] As reported in 1988, Japan held an 85 percent share of world production of low-end laser printers.[81]

A related emerging application is laser scanning. Laser scanners are capable of digitizing virtually any image. The scanners are available for black-and-white and color. These systems completely replace paper copy and yet produce high quality paper output on demand.

Test and Measurement

Test and measurement applications are diverse and include: measurement of particles; environmental sensing; measurement of thickness of and voids in plastics; measurement of surface finish and distortions of many types, including those in aircraft tires; and other applications.

The opportunities in environmental sensing have significant economic potential. The capabilities of two types of laser systems are presently being tested: laser radar (lidar)[82], and laser-induced fluorescence. Laser radar can be used to detect atmospheric pollutants,[83] such as ozone[84] and aerosols,[85] or weather-related parameters, such as water-vapor concentration and wind velocity.[86] Laser radar conducted from aircraft can be used for estimating tree height and timber yield, and for biomass estimation more broadly.[87] Laser-induced fluorescence, generally conducted from aircraft, is also a powerful technique. Plants are illuminated from aircraft with a pulsed laser. In response they fluoresce (emit light). Analysis of the fluorescent light enables identification of plants and assessment of plant vigor. In particular, the analysis can provide information on drought effects, nutrient deficiencies, acid-rain effects, fungal infestations, abnormal levels of heavy metals,[88] and, possibly, presence of herbicides.[89]

Barcode Scanners

Barcode scanners have penetrated supermarkets but have only begun to penetrate other retail outlets, warehouses, and industrial and medical locations. Nevertheless, barcode scanners constituted a $2.5 billion per year industry by 1988.[90]

Alignment and Control

Alignment and control applications include surveying, construction alignment, tool alignment, agricultural applications (e.g., laser guidance for leveling of fields), pointers, and others. An emerging application, which may be loosely classified in this category, is laser radar for aircraft and automobile collision avoidance and for wind-shear detection.[91]

Entertainment

Entertainment applications include light shows and laser pointers. Applications have increased in number and sophistication. Concerns for safety have been successfully addressed. The industry itself has grown to the point where it now has its own international industry association with about 100 firms from seven countries.[92]

Color Separation

Lasers are used to separate the colors during color printing. They are also used increasingly to make the plates used for printing. Emerging applications may soon require broadening this category to include image processing (for example, high-speed Fourier transformations)[93] and laser-based displays, including heads-up displays.

GOVERNMENT APPLICATIONS OF LASERS

Government applications include nearly all commercial applications and others as well. Applications especially significant to the U.S. military are shown in Table 78.[94]

Table 78
U.S. Military Applications of Lasers

Research and Development
Targeting and Ranging
Directed Energy Weapons
Communications
Guidance (laser gyros)
Lidar

The world market for lasers and related optics for military applications is not known, but the U.S. market has been estimated at $412 million for 1992.[95]

GOALS OF U.S. INDUSTRY FOR COMPETITIVENESS

U.S. industry's major goals for improved competitiveness are shown below:

(1) Develop new lasers with additional capabilities and higher performance levels, specifically:
    more frequencies
    higher power levels
    improved beam quality
(2) Improve quality control during manufacturing
(3) Reduce product cost

In addition, U.S. industry must apply laser technology more effectively in multiple areas of laser-based products and processes to maintain its competitiveness in those areas. All of these goals are highly measurement intensive. Unfortunately, the present level of measurement capability does not optimally support industry's pursuit of these goals for improved competitiveness.

EVIDENCE OF MEASUREMENT NEEDS

The measurement needs requiring improved support were identified through interactions with industry, other Federal agencies, and the National Conference of Standards Laboratories (NCSL), composed of industry and Government-agency representatives. NCSL surveyed over 400 organizations within the U.S., including 346 industrial companies, 61 government agencies, and 4 universities, to determine the broad range of critical measurement needs in many technologies that must be addressed to serve "national interests including commerce, international competitiveness, and defense preparedness."[96] With this and other information, NIST developed a description of the measurement needs and submitted them for review by the U.S. laser industry. Industry's responses are reflected in the following analysis.

MEASUREMENT NEEDS

The principal measurement needs that must be met to support U.S. industry are described below. They can be divided into two groups: (1) measurements for lasers, and (2) measurements for laser optics used both within and outside lasers. These two groups of measurements are addressed below. In addition, measurements for characterizing the semiconductor properties are important for semiconductor-diode lasers. This last group of measurements is addressed in Chapter 4, Semiconductors, beginning on page 53.

There are certain types of generic improvements in measurement capability that are needed and that affect the laser industry on an especially broad basis. They are summarized in Table 79.

Much of the needed measurement capability is supportive of all types of lasers. However, there is a special need to assure adequate measurement support for two groups of lasers: (1) semiconductor diode lasers whose low cost and increasing versatility explains their broadening applications and their total domination of the market on a unit basis; (2) high-powered lasers, especially solid-state and carbon-dioxide lasers, which are essential for applications in materials processing and medicine and which represent the greatest part of the market on a dollar basis.

Table 79
Types of Improvements Needed in Laser Measurement Capability

additional specific measured quantities	support for all key measured quantities required for characterization of modern lasers, versus the few now supported
broader frequency ranges	support for all measured quantities over broader ranges of frequencies, versus the limited ranges now supported
broader power levels	support for all measured quantities over broader power levels, versus the limited levels now supported
higher accuracies	ten to twenty times present accuracies for the few measured quantities presently supported, and very high accuracies for additional measured quantities

Laser Measurements

Improved measurement capability is needed for supporting the development, manufacturing, marketing, and application of lasers. A common factor in the application of lasers is assuring their safe use, especially with respect to eye safety. In fact, all lasers manufactured or sold in the U.S. must be labelled for safety in compliance with regulations of the Food and Drug Administration.

The measurement needs for lasers fall into two principal categories: beam strength measurements and beam quality measurements. Beam strength measurements are fundamental to nearly all applications of lasers and thus are a very high priority across the entire U.S. laser industry. Beam quality measurements are needed because all laser beams depart from ideal behavior. Those departures must be minimized to the degree necessary for the success of lasers in given applications. The measured quantities that require improved support are shown in Table 80.

Three of the measured quantities in Table 80 are especially important to diode lasers: switching speed for ultra high speed data and communications applications; beam direction for new electronically steerable diode lasers; and phase front measurements for powerful new diode laser arrays.[97]

It is useful to see how the generic requirements in Table 79 are reflected in Table 80. Here are some examples. Measurement support for *additional specific measured quantities* is needed especially for beam quality measurements for which standard measurement methods are lacking for a number of quantities. Measurement support for *broader frequency ranges* is needed, especially to cover higher frequencies (ultraviolet) of excimer lasers and lower frequencies (down into the infrared) for medicine. Measurement support for *broader power levels* is needed for higher power levels extending to at least 10 kilowatts for materials-processing applications and to 1 megawatt for defense applications. Measurement support is also needed for lower power levels to support laser safety, remote sensing, and targeting equipment. Measurement support for *higher accuracies* is needed broadly across many measured quantities for many applications.

Table 80
Laser Measurement Needs

Measured Quantity	Description
beam strength measurements	
average power	power output from a continuous-wave laser
pulse energy	energy output from a pulsed laser
peak power	highest instantaneous output occurring during a pulse from a pulsed laser
beam quality measurements	
temporal	
continuous-wave lasers	
frequency stability	stability of frequency over time
power stability	stability of power over time
pulsed lasers	
pulse shape	evolution of pulse strength with time
pulse stability	for successive pulses over time:
timing	stability of timing
energy	stability of energy per pulse
peak power	stability of peak power per pulse
frequency	stability of frequency
switching speed	speed at which laser can be turned on and off (laser diodes)
spatial	
beam profile	field distribution across face of beam
beam divergence	angular width of beam
beam direction	direction of beam (for steerable laser diode arrays)
beam waist	minimum diameter of beam over distance (Gaussian beams)
focal spot	diameter of, location of, and intensity within focal spot
modal	
power in modes	distribution of laser power among complicated internal modes of laser operation
polarization state	linear, circular, elliptical, or random
angle	angular direction of polarization of laser beam
purity	unwanted polarization in desired polarization (extinction ratio)
coherence	
temporal	degree of coherence of laser beam over time
spatial	degree of coherence across the front of laser beam
phase front	uniformity of phase across front of beam (for laser diode arrays)
noise	unwanted variations in laser beam
phase	fluctuations in timing of instantaneous power of beam
intensity	fluctuations in power of beam (relative intensity noise, or RIN)
photon statistics	distribution of photons in time
spectral	
purity (linewidth)	degree to which laser beam is truly a single color (single frequency)
wavelength	wavelength (or frequency) of the laser

Laser Optics Measurements

Laser optics include lenses, mirrors, prisms, gratings, windows, filters, beam splitters, attenuators, and other components that are used within and outside lasers to create, condition, and control the direction of laser beams. The performance and quality of laser optics have a controlling effect on the quality of the light produced by a laser. Special measurement methods are required to improve the design of laser optics and to control their quality during manufacturing.

The quantities requiring improved measurement support are shown in Table 81. The primary need is for improved measurement methods, and often for supporting measurement reference standards to assure the accuracy of those methods. However, for some quantities, such as index of refraction, current measurement capability is adequate; and the need is for measured reference materials data. Some of the materials for which measured reference data are needed are not strictly optical materials but rather the targets of lasers, that is, the "worked" materials, such as those cut, welded, or drilled as part of manufacturing processes. For imaging quality, new definitions are needed that are appropriate for coherent systems. Existing definitions for quantities such as *modulation transfer function* or *optical transfer function* are valid only for ordinary light (incoherent) but not for laser light (coherent).

Impact of Measurements on Applications

For any given application a large number of measured quantities will be important. Here are some examples of quantities important to selected applications and to laser safety more broadly.

Materials Processing

Lasers used for cutting, drilling, and welding depend critically on precisely measured levels of beam strength at ever increasing levels of power, extending to 100 kilowatts of continuous power. Further, beam strength must be measured at many different frequencies, since different frequencies will be needed to interact optimally with different materials. For example, for cutting many materials, very high-frequency lasers (ultra violet) are often the best choice. Also important is the measurement of the power in the various modes of the lasers, since minimizing the power in unwanted modes is essential to focusing the beam to a sharp point. Temporal measurements are critical to all applications of pulsed lasers used for materials processing.

Medicine

Lasers used for surgery must have precisely known levels of beam strength and beam profile. For example, the promise of corneal sculpting for replacing eye glasses cannot be realized without accurate beam profile measurements to determine the exact distribution of laser energy across the face of the beam and thus across the cornea. Surgical applications more broadly require lasers that operate at many different frequencies to permit selection of frequencies that provide strong interaction with the tissue to be removed but minimal interaction with the tissue to be retained. Improved measurement capability and supporting national reference standards are needed at both ultraviolet and infrared frequencies, in particular, to support excimer and solid-state lasers. Measurement of the temporal properties of lasers is very important to medicine since pulsed lasers are frequently used for surgical processes.

Table 81
Laser Optics Measurement Needs

Measured Quantity	Description
transmission properties	
scattering	deflection of light by imperfections in a material
reflectivity	amount of light reflected by a material
absorption	amount of light absorbed per unit length of a material
attenuation	reduction in strength of a light beam after passing through a material
lidar cross-section	response of a target material to a laser-radar beam
damage threshold	minimum power per unit area ("irradiance") and minimum energy per unit area ("fluence") that damages a material
refractive index	factor by which light is slowed as it passes through a material
birefringence	difference in refractive index for light of different polarizations passing through a material
magneto-optic properties	
Verdet constant	variation of refractive index with magnetic field for light transmitted through a material (Faraday effect: induced circular birefringence)
Kerr rotation	rotation of polarization for light reflected by a material due to changes in refractive index with magnetic field
Kerr ellipticity	differences in refractive index for light of different polarizations passing through a material due to a magnetic field
electro-optic properties	
linear coefficients	variation of refractive index with electric field (Pockels effect)
Kerr constant	variation of refractive index with square of electric field
non-linearities	departures of optical materials from linear (or proportional) response to light
imaging quality	degree to which optical components direct light as designed
focal length	distance from a lens at which focus is obtained
spot size	diameter of light beam from a lens at the point of focus

Communications

For laser communications, switching speed, spectral purity (linewidth), and wavelength measurements are important to the most powerful implementations, whether in optical-fiber systems or in free-space applications. Also important is the measurement and minimization of noise on the beam which can interfere with the transmission of data.[98] Improved measurement capability is also needed for attenuation levels in the optical materials used in attenuators.

Test and Measurement

Measurement of coherence is particularly important to precise test and measurement processes, such as inspection of surface irregularities which depends critically on the interference properties of laser light.

Safety

Improved capability for beam-strength measurements is particularly important to laser safety, with a special emphasis on average power and pulse energy. Following closely in importance are improved capability for beam-quality measurements with an emphasis on the temporal, spatial, coherence, noise, and spectral categories in Table 80. In particular, beam-profile and beam-divergence measurements capable of easy interpretation are needed. Further, improved measurement capability for optical materials is needed. In particular, new national measurement reference standards are needed that are capable of supporting the development of improved optical filters with 100-times higher levels of attenuation for eye protection. Note that laser safety must be addressed for almost any application of lasers. Because laser light is coherent, it interacts more strongly with materials, including the retina of the eye, than ordinary (incoherent) light. For this reason, low levels of laser light are more potentially damaging to the eye than much higher levels of ordinary light.

ENDNOTES

1. In technical terms, laser light can have a high *radiance* (power per unit area of the laser surface per unit solid angle of divergence as the light leaves the laser). Laser light can be focused to produce a high *irradiance* (power per unit area).

2. George A. Shukov and Al Smith, "Micromachining With Excimer Lasers", *Lasers & Optronics*, p. 76 (September 1988).

3. This is the most common configuration for a laser. More complex configurations are also used.

4. Excimer lasers form molecules that are stable only when the participating atoms are in an electronically excited state. Common molecular lasers use molecules that are stable even when not in an excited state. Chemical molecular lasers combine atoms via a chemical reaction to produce molecules that are stable and in an excited state.

5. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, pp. 57 and 63 (January 1992).

6. W. Conard Holton, "The Writing on the Wall: Military-Based Photonics Firms Get the Message", *Photonics Spectra*, p. 62 (June 1992).

7. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 57 (January 1992).

8. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 63 (January 1992).

9. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 63 (January 1992). David Kales, "1988 Laser Economic Review and Outlook", *Laser Focus World*, p. 91 (January 1988).

10. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 139 (January 1992).

11. David Kales, "Optics Makers Battle Defense Cuts and Recession", *Laser Focus World*, p. 128 (October 1991).

12. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 57 (January 1992).

13. David Kales, "1992 World Review Forecast of Laser Markets", *Laser Focus World*, p. 83 (January 1991).

14. Len DeBenedictis, "Foreign Competition Poses Threat", *Laser Focus World*, p. 19 (January 1989).

15. David Belforte, Belforte Associates, private communication (July 1990).

16. Len DeBenedictis, "Foreign Competition Poses Threat", *Laser Focus World*, p. 19 (January 1989).

17. Ernest Sternberg, "Should We Have a Policy for Optoelectronics?", *Laser Focus World*, p. 59 (September 1992).

18. Len DeBenedictis, "Foreign Competition Poses Threat", *Laser Focus World*, p. 19 (January 1989).

19. David Kales, "1989 Laser Economic Review and Outlook", *Laser Focus World*, p. 112 (January 1989).

20. *JTECH Panel Report on Opto- & Microelectronics,* Science Applications International Corporation, p. 6-10 (May 1985). The Japanese Technology Evaluation Program (JTECH) was initiated in 1983 by the U.S. Department of Commerce.

21. David Kales, "1989 Laser Economic Review and Outlook", *Laser Focus World*, p. 112 (January 1989).

22. According to a report issued by Market Intelligence Research Company (Mountain View, California) as cited by Susan Walsh, "Industrial Lasers Continue Steady Climb", *Managing Automation*, pp. 18-19 (July 1989).

23. Mark A. Fischetti, "Exporting to Europe: A Special Report", *Photonics Spectra*, p. 112 (June 1991).

24. "Competitiveness in the U.S. Optics Industry: What's Missing", *Laser Focus World*, p. 59 (January 1989).

25. David Kales, "What's Ahead in Military Electro-Optics", *Laser Focus*, p. 91 (August 1988).

26. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, pp. 62-63 (January 1992).

27. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, pp. 62-63 (January 1992).

28. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 63 (January 1992).

29. David Lytle, "U.S. Photonics Industry: Facing Premature Death?", *Photonics Spectra*, p. 36 (February 1989).

30. Gary T. Forrest, "Market Performance Indicates Bright Outlook for Lasers in Japan", *Laser Focus World*, p. 112 (June 1989).

31. Lewis M. Holmes, "Commercial Lasers 1988-89", *Laser Focus World*, pp. 67-68 (January 1989). Donald B. Carlin, "Optical Recording Drives Diode-Laser Technology", *Laser Focus World*, p. 84 (July 1992). "Blue Diode Emits at 447 nm", *Laser Focus World*, p. 15 (September 1992).

32. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 63 (January 1992).

33. Dennis Worth, "Small Lasers With a Big Future", *Photonics Spectra*, p. 143 (April 1988).

34. Much of the material in this section comes David Kales and in "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, pp. 56-81 (January 1992). Additional examples of applications come from Lewis M. Holmes, "Commercial Lasers 1988-89", *Laser Focus World*, pp. 67-93 (January 1989), and from "1989 Trends: Lasers", *Photonics Spectra*, pp. 111-120 (January 1989). Many other articles are cited, as appropriate, in other notes.

35. David Kales, "1990 World Review and Forecast of Laser Markets", *Laser Focus World*, p. 91 (January 1990).

36. David Kales, "1991 World Review and Forecast of Laser Markets", *Laser Focus World*, p. 91 (January 1991).

37. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 63 (January 1992).

38. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 63 (January 1992).

39. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 63 (January 1992).

40. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 63 (January 1992).

41. Heather W. Messenger, Yvonne A. Carts, "Commercial Laser Technology Advances Steadily", *Laser Focus World*, p. 64 (December 1991).

42. David A. Belforte, "High-Power Nd:YAG Lasers Shine in Industrial Processing", *Laser Focus World*, p. 69 (September 1992).

43. *OITDA Activity Report, Vol. 5, For Fiscal Year Ended March 31, 1992*, Optoelectronic Industry and Technology Development Association, p. 5 (1992). The datum is for the 1991 fiscal year used in Japan (ending March 31, 1992) rather than for the calendar year 1991. The datum was published in yen and was converted to dollars by dividing by 134.59 yen per dollar. This conversion factor is the average annual exchange rate for calendar 1991, based on a Federal Reserve Statistical Release of January 24, 1992. An average annual exchange rate for Japan's fiscal year was not available.

44. Yvonne A. Carts, "Lucrative Niches Await Lasers in Semiconductor Processing", *Laser Focus World*, p. 105 (May 1989). Alan Cable, "Laser Use Grows in Linear and Mixed-Signal IC Production", *Photonics Spectra*, pp. 162-164 (November 1991). Excimer lasers, which can produce very short wavelengths (ultraviolet range) may prove especially important for semiconductor lithography when linewidth requirements drop below 0.5 micrometers which is the linewidth required for 16 megabit dynamic random access memories (DRAMs). Gary T. Forrest, "Excimer Lasers Access Submicrometer Semiconductor Lithography", *Laser Focus World*, p. 23 (May 1989). Robert F. Miracky, "Lasers Advance into Microelectronics Packaging", *Laser Focus World*, p. 85 (May 1991).

45. Harald N. Bransch, "Welding with High-Power Pulsed and CW Nd:YAG Lasers", *Photonics Spectra*, p. 107 (September 1991).

46. "Laser-Wielding Robots Tackle Automobile Production", *Photonics Spectra*, p. 139 (November 1988). Susan Walsh, "Industrial Lasers Continue Steady Climb", *Managing Automation*, p. 18 (July 1989).

47. Thomas A. Znotins, Darcy Poulin, and John Reid, "Excimer Lasers: An Emerging Technology in Materials Processing", *Laser Focus*, p. 54 (May 1987).

48. Gary T. Forrest, "Market Performance Indicates Bright Outlook for Lasers in Japan", *Laser Focus World*, p. 112 (June 1989).

49. The European Economic Community has launched the Eureka Project, which contains a Eurolaser Program, funded at $74 million, to develop three types of industrial lasers (carbon dioxide, yttrium aluminum garnet, and excimer). [Uwe Brinkmann, "The EEC develops industrial lasers", *Laser Focus World*, p. 75 (May 1989).] The U.K. Atomic Energy Authority's Culham Laboratory is leading a $22.5 million program for building excimer lasers for cold machining of the hardest materials. [Brian Dance, "Culham will lead an excimer project", *Laser Focus World*, p. 76 (May 1989).]

50. Roger Main, "Laser Robots Lend VW a Helping Hand", *Lasers & Optronics*, p. 42 (post August 1987).

51. David Kales, "1989 Laser Economic Review and Outlook", *Laser Focus World*, p. 108 (January 1989).

52. Bernhard H. Klimt, "Review of Laser Marking and Engraving", *Lasers & Optronics*, p. 61 (September 1988).

53. Bernhard H. Klimt, "Review of Laser Marking and Engraving", *Lasers & Optronics*, p. 62 (September 1988).

54. "Laser Lights Way to 3-D Prototypes", *Photonics Spectra*, p. 114 (January 1989).

55. "Europe/U.S.: Three Vision Shows Spawn New Products", *Lasers & Optronics*, p. 30 (September 1988).

56. George A. Shukov and Al Smith, "Micromachining With Excimer Lasers", *Lasers & Optronics*, p. 75 (September 1988).

57. Some of the material in this section comes from Michael Moretti, "Medical-Laser Technology Responds to User Needs", *Laser Focus World*, pp. 89-102 (March 1989). Other articles are cited, as appropriate.

58. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 43-1 (January 1992).

59. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 45-2 (January 1992) quoting the Health Industry Manufacturers Association, International Affairs.

60. *OITDA Activity Report, Vol. 5, For Fiscal Year Ended March 31, 1992*, Optoelectronic Industry and Technology Development Association, p. 5 (1992). The datum is for the 1991 fiscal year used in Japan (ending March 31, 1992) rather than for the calendar year 1991. The datum was published in yen and was converted to dollars by dividing by 134.59 yen per dollar. This conversion factor is the average annual exchange rate for calendar 1991, based on a Federal Reserve Statistical Release of January 24, 1992. An average annual exchange rate for Japan's fiscal year was not available.

61. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 63 (January 1992).

62. Dennis L. Werth, "Laser Diodes Growing Up", *Photonics Spectra*, p. 133 (April 1989). Frost & Sullivan, Inc. in an announcement of its market study "Medical Lasers", p. 1 (Spring, 1987). Radical keratotomy corrects nearsightedness by changing the curvature of the cornea through small radial cuts made outside the visually significant central region. Corneal sculpting corrects nearsightedness and farsightedness by ablating a thin layer of corneal material from the visually important region. Thomas F. Deutsch, "Medical Applications of Lasers", *Physics Today*, p. 59 (October 1988). David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 67 (January 1992). Glaucoma, retinal disease, birthmarks and port-wine stain removal have been cited by David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 64-66 (January 1992). Dentistry is discussed by Dr. Ricki Lewis, "Lasers: A Cure for Dentaphobia?", *Photonics Spectra*, pp. 74-75 (April 1992) and by Stephen A. May, "Expanding Laser Applications Foster Technological Evolution", *Photonics Spectra*, p. 99 (November 1991).

63. C. Breck Hitz in a laser course presented at NIST in June 1988.

64. Irving J. Arons, "Will Corneal Sculpting with Excimer Lasers Eliminate Eyeglasses?", *Laser Focus World*, p. 62 (March 1992).

65. David Kales, "1989 Laser Economic Review and Outlook", *Laser Focus World*, p. 104 (January 1989). The author does not indicate if the $6 billion figure refers to laser sales, laser-based surgical equipment sales, surgical services, or some combination of these. He also does not indicate if he is referring to the world market, but that can be safely assumed since his entire article addresses the world market.

66. David Kales, "1989 Laser Economic Review and Outlook", *Laser Focus World*, p. 112 (January 1989).

67. "Resonant Leaky-Wave Antiguides Make High-Power Diode Laser Array", *Laser Focus World*, p. 9 (May 1989). Philip Speser, "Doing Business in Space", *Laser Focus World*, p. 19 (May 1989).

68. "CLEO '88 Takes in New Topics", *Laser Focus*, p. 116 (April 1988). CLEO is the Conference on Lasers and Electro-Optics sponsored by the Optical Society of America and by the Lasers and Electro-Optics Society (LEOS) of IEEE.

69. David Begley, "Lasers for Spaceborne Communications", *Photonics Spectra*, p. 73 (February 1989). David Begley and Bhogi Boscha, "Laser Diodes Conquer The Challenge of Space Communications", *Photonics Spectra*, p. 147 (April 1989).

70. *OITDA Activity Report, Vol. 5, For Fiscal Year Ended March 31, 1992*, Optoelectronic Industry and Technology Development Association, p. 5 (1992). The data are for the fiscal years used in Japan (ending March 31) rather than for the calendar years. The data for Japan's fiscal year 1991 in the table were published in yen and were converted to dollars by dividing by 134.59 yen per dollar. This conversion factor is the average annual exchange rate for calendar 1991, based on a Federal Reserve Statistical Release of January 24, 1992. An average annual exchange rate for Japan's fiscal year was not available.

71. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 37-16 (January 1992).

72. Howard Fleming, Office of Microelectronics and Instrumentation, International Trade Administration, U.S. Department of Commerce (September 30, 1992).

73. *1992 Electronic Market Data Book*, Electronic Industries Association, p. 34 (1992).

74. *OITDA Activity Report, Vol. 5, For Fiscal Year Ended March 31, 1992*, Optoelectronic Industry and Technology Development Association, p. 5 (1992).

75. *1992 Electronic Market Data Book*, Electronic Industries Association, p. 34 (1992).

76. The 7.9 percent market share comes from the *1990 Disk/Trend Report: Optical Disk Drives*, Disk/Trend, Inc., p. SUM-10 (July 1990). The $0.9 billion figure come from Disk/Trend, Inc., as quoted from its *1991 Disk/Trend Report: Optical Disk Drives*, p. SUM-14 (July 1991).

77. After correction of some outyear calculations, these percentages were produced from data by Dataquest, "Research Newsletter: Dataquest's Electronics Industry Forecast", p. 6 (May 1989).

78. *1990 Disk/Trend Report: Optical Disk Drives*, Disk/Trend, Inc., p. DT10-11 (July 1990).

79. InfoCorp Computer Intelligence, as reported in *PC Week Buyers' Guide*, p. 105 (October 26, 1992). World production is reported as 650,000 drives in 1991 and 900,000 in 1992, with the U.S. consuming 447,200 in 1991. World market revenue is reported as $264 million in 1992 and is projected to rise to $575 million in 1996.

80. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 27-22 (January 1992) quoting data from BIS Strategic Decisions, a Massachusetts-based market research and consulting company.

81. Arnold Mayer, Norbert Schroder, "Opto-Electronic Components and Systems" by Prognos AG of Basel, Switzerland (August 1988).

82. Laser radar is often called "lidar" for light detection and ranging. Laser radar is also sometimes called "ladar" for laser detection and ranging.

83. David Kales, "Review and Forecast of Laser Markets: 1992", *Laser Focus World*, p. 67 (January 1992). Robert J. Nordstrom, "The CO_2 Laser in Remote Sensing", *Photonics Spectra*, p. 89 (February 1992).

84. "$2-Million Ozone Lidar System Ships to Japan", *Laser Focus*, p. 28 (June 1988).

85. Yvonne A. Carts, "Laser Radar: Solid-State Lidar Development Promises Longer Lifetimes", *Laser Focus World*, p. 23 (June 1989).

86. Yvonne A. Carts, "Laser Radar: Solid-State Lidar Development Promises Longer Lifetimes", *Laser Focus World*, p. 23 (June 1989).

87. Emmett W. Chappelle, Darrel L. Williams, Ross F. Nelson, and James E. McMurtrey III, "Lasers May Help in Remote Assessment of Vegetation", *Laser Focus World*, p. 123 (June 1989). The authors are from the National Aeronautics and Space Administration and the U.S. Department of Agriculture.

88. Emmett W. Chappelle, Darrel L. Williams, Ross F. Nelson, and James E. McMurtrey III, "Lasers May Help in Remote Assessment of Vegetation", *Laser Focus World*, p. 123 (June 1989). The authors are from the National Aeronautics and Space Administration and the U.S. Department of Agriculture.

89. "Laser Detects Herbicide on Tobacco", *Lasers & Optronics*, p. 26 (September 1988).

90. Bernhard H. Klimt, "Review of Laser Marking and Engraving", *Lasers & Optronics*, p. 61 (September 1988). Klimt indicates that "marking and bar coding represent a $2.5 billion-per-year industry" and that marking is only 1 to 2% of the combined market; thus virtually the entire $2.5 billion is bar coding. Klimt does not indicate whether the market he is characterizing is U.S. or world, but world is probably intended and will be assumed here. Certainly that assumption will not overstate the world market.

91. The Federal Aviation Administration (FAA) of the Department of Transportation has issued a regulation, effective in 1993, requiring that all commercial aircraft be equipment wind shear detection systems. Mark E. Storm, "Coherent 2-µm Sources Burst Into Wind-Shear Detection", *Laser Focus World*, p. 117 (April 1991). Perhaps be 3600 jetliners will be affected at a cost of perhaps $50,000 per aircraft, for a total cost approaching $200 million. FAA is not specifying the technology to be used; the marketplace will decide. Microwave radar, laser radar, and passive infrared systems are the primary candidates for forward looking detection. FAA and NASA are jointly investing $24.8 million over five years (beginning in 1987) to stimulate development of advanced wind-shear detection systems in their Forward Looking Technology Program. The systems are described in *Airborne Wind Shear Detection and Warning Systems: First Combined Manufacturers' and Technologists' Conference*, Federal Aviation Administration, U.S. Department of Transportation, and National Aeronautics and Space Administration (January 1988). Lidar is discussed by Yvonne A. Carts, "Laser Radar: Solid-State Lidar Development Promises Longer Lifetimes", *Laser Focus World*, p. 23 (June 1989). A lidar windshear detection system is being flight tested in 1992 by NASA. "A Laser Devense Against the Dreaded "Windshear" Effect, *Photonics Spectra*, p. 116 (August 1992). "Wind shears are the primary cause of death of passengers on US commercial airliners." Stephen A. May, "Expanding Laser Applications Foster Technological Evolution", *Photonics Spectra*, p. 100 (November 1991). Frank Allario, Richard R. Antcliff, Philip Brockman, and Mitchel E. Thomas, "Laser Remote Sensing Supports Aviation Research and Safety", *Laser Focus World*, p. 59 (April 1992).

92. Ivan Dryer, "Laser Color Display and Entertainment Applications", *Laser Focus World*, p. 69 (September 1991). The new association is the International Laser Display Association (ILDA).

93. James Lafuse and Terry Turpin, "Hybrid Optoelectronic Processors Generate Images in Real Time", *Laser Focus World*, p. 123 (September 1992).

94. Jon Grossman, "Military Laser Systems: A Costly Search for Ultimate Weapons", *Photonics Spectra*, pp. 84-90 (July 1991).

95. W. Conard Holton, "The Writing on the Wall: Military-Based Photonics Firms Get the Message", *Photonics Spectra*, p. 64 (June 1992).

96. *Report by the National Measurement Requirements Committee*, National Conference of Standards Laboratories, NMRC 89-01, p. 1-1 and Chapter 4 (January 1989). The survey was conducted in 1982 and was updated periodically through 1988.

97. The phase front measurement determines the degree to which an array of diode lasers looks like a single laser from a distance, that is, how well the lasers are working together.

98. Spectral purity measurements for the low frequency (near infrared) laser diodes designed for optical fiber communications will be developed with resources being pursued separately. Spectral purity measurements for other lasers, which operate at a much greater diversity of frequencies, will be addressed by the expanded program addressed here.

CHAPTER 9

OPTICAL-FIBER COMMUNICATIONS

April 1993

Chapter 9

OPTICAL-FIBER COMMUNICATIONS

SUMMARY

Optical fibers are hair-thin strands of glass that guide beams of light carrying information. The fibers offer high information capacity, low loss of signal strength over distance, low noise, high security from interception, and good environmental tolerance.

Optical fibers are used in many communications applications, including those traditionally associated with telephone services and those traditionally associated with computer services. Those associated with telephone services include long-haul terrestrial lines and undersea cables. Those associated with computer services include local-area networks and metropolitan-area networks. Optical fibers also serve in cable-television installations. In the future, optical fibers are likely to serve in the subscriber loop that connects individuals and businesses to each other and to long-distance terrestrial and undersea cables. Once these last connections have been made, individual users will gain direct access to services requiring the high-information capacity of optical fibers. More generally, these several applications will progressively merge.

These new communications capabilities are made possible by advances in both component technology and network architectures. Many components are required to provide optical-fiber communications services. Examples include laser diodes that generate light and place information on it, fibers that carry light, switches that redirect light to desired destinations, detectors that recover information from light, and many other components. These components are assembled into optical-fiber communications equipment that is integrated into networks. The networks make and manage the interconnections of senders and receivers and assure the reliable transfer of information between them.

Increasingly, new networks are being designed to support data in digital format. The digital format enables correcting for signal degradation during transmission and storage and supports sophisticated signal processing, such as data compression. The new networks will accommodate all kinds of information traffic, including voice, video, computer data, and text, in a manner that is transparent to the users. Key among the newer network designs is the Narrowband Integrated Services Digital Network (ISDN) and an emerging high-speed version, Broadband ISDN, which is being designed especially for optical fibers. Broadband ISDN will be capable of delivering up to 622 megabits per second of information to an individual user. It will support services in education, health care, business, government, transportation, publishing, entertainment, and other areas of interest or endeavor. Direct access to the great libraries of the world, to the newspapers of the world, to diverse high-speed computers, to manufacturing databases, and to many other services may become possible through the enormous information capacity provided by optical-fiber technology.

The development of network architectures, such as Broadband ISDN, is an international labor. This labor has been undertaken in pursuit of a widely held world vision of an international communications network with high compatibility across diverse types of domestic equipment and many network architectures. The requirements that such networks must meet, and the standards that embody those

requirements, are developed and promulgated under the auspices of such bodies as the International Consultative Committee for Telephone and Telegraph (CCITT) and the International Organization for Standardization (ISO). The United States and many other countries of the world have contributed significantly to the development of these international networking standards.

This international activity is creating a growing market for telecommunications components and equipment using optical-fiber technology. The world market for components alone is estimated at $4 billion for 1992 and is expected to grow to $8 billion by 1998. During this period, consumption of components will increase by factors of two to five, in quantity, for different types of components. The components will be used by many countries to complete terrestrial communications lines and to provide dozens of new undersea lines, some of considerable extent; the longest currently planned covers 15,000 kilometers from Japan to the United Kingdom. The components will also be used to begin the process of extending optical-fiber connections directly to end users through the subscriber loop (fiber-to-the-home). If this process is completed, the cumulative world investment in optical-fiber communications networks will be many hundreds of billions of dollars.

Competition for these huge markets will be keen. The U.S., Japan, and Europe are all engaged in high-quality research and development efforts and have significant production capabilities. For optoelectronic products broadly, Japan is in the lead in both technology and commercialization and is pulling further ahead. For optoelectronic products used in optical-fiber communications, in particular, the competitive situation is more difficult to determine in the absence of better worldwide production data by country. However, the U.S. has been outstanding in the production of optical fiber cable. Optical-fiber cable currently accounts for over half of the dollar value of all components for optical-fiber communications systems (1992). At the same time, Japan and Europe also have excellent production capabilities for fiber cable and have been pursuing the technology and manufacturing processes for the other components with determination and success.

The rate of technological progress in this field has been staggering, and the potential for further advances remains great. Yet, at the present time, newly installed optical-fiber communications systems exploit only about 0.1 percent of the information capacity of optical fibers. Current network architectures seek to exploit this small fraction efficiently. The limitation is in the components. New components will be required to do better. Breakthroughs in component design can, and likely will, dramatically widen the options available to networks designers.

U.S. industry is pursuing several goals in its efforts to assure its competitiveness in future world markets for components. U.S. industry is attempting to reduce costs and increase quality for components that accommodate current capacity levels. U.S. industry is also developing new components with dramatically improved performance levels to exploit the untapped information capacity of optical fibers. For both classes of components, U.S. industry is pursuing the development of optical integrated circuits. Such circuits enable improvements in performance, quality, and cost. Optical integrated circuits will be particularly important for components used in the high-volume subscriber-loop markets for which low unit costs are a must. Also important is U.S. industry's pursuit of the development of microwave integrated circuits. These electronic circuits are needed to serve several functions within, or supportive of, optical-fiber communications systems: (1) provide signal compression, time-division multiplexing and demultiplexing, high-speed electronic switching, and wavelength-division demultiplexing [for high-performance coherent systems]; (2) provide low-cost radio interconnections between mobile users and cable networks; (3) provide low-cost satellite communications systems to back up the cable systems in time of failure; and (4) provide low-cost satellite communications systems to connect regions of the world which cable cannot directly access. These regions may be difficult to access for economic or physical reasons deriving from factors such as low population density, remoteness, or rough terrain.

All of these goals for U.S. industry are highly measurement intensive, and important parts of the measurement capability needed by U.S. industry are lacking. This shortfall adversely affects research and development, manufacturing, marketplace exchange, and after-sales support of products for optical-fiber communications systems. The basic types of capability needed are (1) new measurement methods, (2) new national measurement reference standards to support the accuracy of the measurement methods, and (3) new measured materials reference data for the properties of important optoelectronic and other materials. For the measurement methods, the improvements needed take several forms: better accuracy, increased dynamic range, broader frequency coverage, increased simplicity, improved documentation, and improved standardized definitions of important measured quantities. Representative measurement needs for components, materials, and networks follow.

For components, new measurement capability is needed to support eleven groups of components, such as optical fibers, laser diodes, connectors, modulators, and detectors. Representative needs include measurement reference standards to support wavelength measurements on laser diodes, measurement methods and measurement reference standards for return loss (reflected power) from nearly all components, improved measurement methods for linearity in modulators, and improved measurement methods for polarization mode dispersion in isolators.

For materials, new measurement capability is needed to support the use of compound semiconductors, glasses, electro-optical materials (such as lithium niobate and lithium tantalate), dielectric coatings, and ferrimagnetic garnets. For example, for compound semiconductors, new measurement reference standards are needed for alloy content, doping concentrations, and layer thicknesses. Also needed are new materials reference data for excited state emissions and stimulated emission cross-sections for dopants in glasses to support development of optical amplifiers (especially for 1310 nanometers). Materials measurements are especially important to the realization of optical integrated circuits.

For networks, difficult measurement challenges must be successfully addressed to provide sophisticated measurement methods and testing strategies. These methods and strategies must support network commissioning, performance monitoring, and fault location. They must support networks employing diverse data packaging, multiplexing, and switching technologies. Examples of these technologies include high-speed digital data packets and frames, time-division and wavelength-division multiplexing, and virtual-circuit switching, wavelength-switching, and broadcasting.

CAPABILITIES OF OPTICAL FIBERS

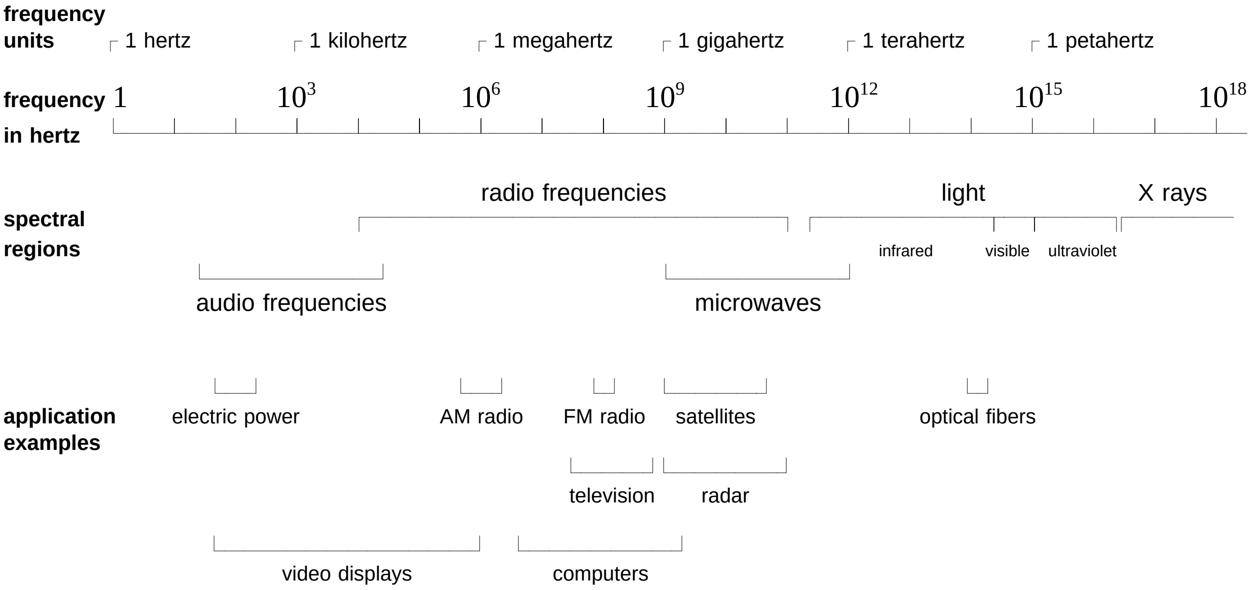
Optical fibers are a capable medium for use in terrestrial communications systems. Optical fibers are hair-thin strands of glass that guide beams of light carrying information. The fibers offer many desirable properties, including high information capacity, low loss of signal strength over distance, low noise, high security from interception, and good environmental tolerance.

High Information Capacity

The high information capacity of optical fibers is attributable to the high frequency of the light that travels through the fibers. That frequency falls in the upper end of the infrared region, as shown in Figure 6 below. A typical frequency for this light is 200 terahertz. [A terahertz is 10^{12} hertz, or 1000 gigahertz.] The frequency of light determines its color. The infrared region in which optical fibers for communications applications operate is below the deepest red that the human eye can see, so the light used is invisible.

Light is often described by its wavelength, rather than its frequency. The wavelength is just the distance that the light travels in one cycle of its frequency. For the frequencies used in optical fibers, this distance is about 1 micrometer or 1000 nanometers. The frequency and the wavelength of a lightwave are inversely related, that is, the higher the frequency, the lower the wavelength.[1]

Figure 6
Overview of the Frequency Spectrum



All electromagnetic waves, including light, have an information capacity that is proportional to their frequency. This means that the light in an optical fiber can, in principle, carry 20,000 times more information than a typical microwave signal with a frequency of 10 gigahertz. In fact, lightwaves in an optical fiber should be able to carry at least 2000 gigabits per second of digital data. This capacity is enough for 2000 channels of high-definition digital television or for 16 million digital telephone calls, without signal compression. With signal compression, at least 10 times more channels or calls could be carried. Thus, a single optical fiber, if its full information capacity could be exploited, could easily carry all of the telephone traffic of the United States.

Of course, a single fiber cannot go everywhere. Also, the rate at which information can be sent through a single fiber is presently limited to much lower levels by the technology available for the non-fiber components. However, even the lower levels are noteworthy. Speeds of 1.7 gigabits per second are being employed in the newer installations. This capacity is high enough to send all of the text in an encyclopedia in one or two seconds. Soon, new systems will carry data at 2.5 gigabits per second; and several experimental systems with capacities of 10 gigabits per second[2] or higher have been tested.

Low Attenuation Over Distance

Lightwaves passing through optical fibers gradually lose strength, that is, they are gradually attenuated, as they travel down the fibers. Losses in modern fibers are quite low, typically 0.2 decibels per kilometer of fiber for newer fibers. This is equivalent to about 5 percent per kilometer. Additional losses are incurred at interconnections between fiber segments. These interconnections are necessary because the maximum length for a continuous coil of optical fiber is limited by practical considerations, such as the size and weight of the resulting coil. Lengths of 5 to 7 kilometers are common. When all sources of loss are considered, modern long-distance optical-fiber lines achieve overall losses of about 0.5 decibels per kilometer, or about 11 percent per kilometer.

Virtually Noiseless Transmission

Optical fibers are virtually free of unwanted thermal noise. That is, they are free of unwanted additions to the light that they carry created by thermal vibrations. This freedom from noise results from the fact that the frequencies of the light in optical fibers are so high that they are above the frequencies generated by thermal motion at ambient temperatures. A simple indication of this fact is that an electric burner on a stove must be heated to a very high temperature before it starts emitting light (glowing red). Thus, unlike wires and coaxial cable, optical fibers provide a virtually noise free path for signals. Their performance with regard to noise is limited principally by the noise in the other components used in optical-fiber communications systems.

High Security from Interception

The light in an optical fiber remains so well confined that "tapping" some of that signal becomes very difficult and is generally easily detected. These characteristics give optical fibers very desirable characteristics for maintaining data security, that is, resistance to undetected interception. Unfortunately, optical fibers, like any other cable technologies, are highly vulnerable to interruption through accident or sabotage since they are, by nature, exposed over many kilometers. For this reason, optical-fiber systems require back up by alternate fiber routes and by wireless radio-frequency systems. Microwave satellite systems are commonly employed.

Good Environmental Tolerance

Optical fibers are high in environmental tolerance. Since most optical fibers are made of glass, they are insulators. For this reason, they are virtually immune to electromagnetic interference.[3] Electromagnetic interference complicates the use of virtually all wireless radio-frequency technologies, as described in Chapter 12, Electromagnetic Compatibility, beginning on page 381.

The insulating character of optical fibers provides perfect electrical isolation between the electrical systems that fibers interconnect. The fibers are relatively unaffected by the presence of either high voltages or strong magnetic fields. The fibers are also relatively inert chemically, so they are largely immune to hazardous chemical environments.

COMPONENTS IN OPTICAL-FIBER SYSTEMS

Optical-fiber communications systems are composed of a number of types of components that perform three basic functions: (1) generate lightwaves and place information them; (2) transmit and direct lightwaves; and (3) recover information from lightwaves.

Generating Lightwaves and Placing Information on Them

Sources/Transmitters

The lightwaves used in optical-fiber communications systems are generated by lasers and light-emitting diodes. The lasers are generally semiconductor diode lasers, although lasers based on optical fibers have recently been developed[4] and may soon come into use. [For a more detailed discussion of lasers, see Chapter 8, Lasers, beginning on page 183.]

Laser diodes come in two principal forms: Fabry-Perot and distributed feedback (DFB). A comparison of their advantages is shown in Table 82. Light-emitting diodes are less expensive and may be used with either multimode or single-mode fibers when lower light levels and a lower information capacity (lower modulation bandwidth) are acceptable. Laser diodes achieve higher performance at the expense of higher cost and complexity. The properties of these sources will be discussed further later in this chapter.

Table 82
Comparison of Light Sources

Desirable Characteristics	Light-Emitting Diodes	Fabry-Perot Laser Diodes	Distributed-Feedback Laser Diodes
high launched optical power	low	high	high
simple electronics	simple	less simple	less simple
low cost	low	intermediate	high
narrow linewidth	wide	narrow	very narrow
high modulation bandwidth	low	high	high

The light sources are assembled into packages called transmitters. The transmitters contain supporting electronic circuits that may provide a variety of services: control the optical power output of the source, stabilize that output, and provide cooling (through a thermoelectric effect). The transmitters may also contain isolators. These are components that prevent the return to the source of any light that may be reflected by other components down the line from the source. Returned light can interfere with the operation of laser diodes.

The transmitters usually contains a fiber "pigtail". The pigtail is an optical-fiber lead that is connected directly to the internal parts of the transmitter. That connection must be made with care to assure that the light produced by the source is not reflected or absorbed but rather is carried out of the transmitter. Pigtails are connected to the inputs and the outputs of virtually all components that lack natural fiber leads; the pigtails facilitate interconnection with other components.

Direct Modulation

In most optical-fiber systems, information is placed on a lightwave by a process called *direct modulation*. Direct modulation is accomplished by varying the power supplied to the light source. This type of modulation is particularly suited to digital systems, that is, to systems that communicate based on switching among discrete levels, such as those represented by strings of 1s and 0s. The optical power output from the source may be varied from full on, to correspond to one digital state, to full off, to correspond to a second digital state. The result is a string of pulses of light that are very square. That is, the optical-power level rises quickly on the leading edge of a pulse, reaches a steady maximum level, and then falls very quickly at the trailing edge of a pulse.

The rate at which the optical output power of sources can be raised and lowered by *direct modulation* is limited. However, for semiconductor laser diodes, modulation rates of 25 gigahertz have been attained; and the theoretical limit appears to be very high, perhaps 40 gigahertz.[5] This rate exceeds the speeds of the best supporting electronic systems (20 gigahertz) that must assemble the information that is to be transmitted.

External Modulation

An alternative to direct modulation is external modulation. External modulation employs a separate component called a *modulator*. A modulator is an optical window that can pass or block light in any proportion. Its optical properties may be varied from transparent to opaque by applying an electrical voltage. This variation can be quite linear; that is, the optical power level of the lightwave coming out of the modulator can be made to vary in direct proportion to the level of the electrical signal that controls the modulator.

For analog optical-fiber systems, such as those used for optical-fiber-based cable-television systems, modulators are especially important. Modulators vary the optical power of a lightwave in a way that enables the lightwave to carry, simultaneously, all of the radio-frequency signals of many television channels. At locations near to customers, the radio-frequency signals are recovered from the light and sent to individual users through coaxial cable, in a manner compatible with ordinary television sets. This cable-television application requires high linearity from the modulator since any departure from linearity will degrade the resulting signal. This analog modulation process is much more demanding than the simple on/off direct modulation process used in digital telecommunications systems.

For digital optical-fiber systems, modulators may be able to provide higher rates of modulation for digital systems than direct modulation can deliver. Such high rates may become important in the future if the speeds of digital electronic signal-processing circuits can be increased. However, even at lower speeds, the high linearity of modulators makes their use important in digital systems. The well-shaped digital light pulses made possible by the high linearity of the modulators are helpful in overcoming an important type of degradation that the pulses experience during transmission, called chromatic dispersion. Chromatic dispersion is discussed further beginning on page 227. The well-shaped pulses are also important to the realization of coherent systems. Coherent systems offer high levels of performance in transmitting information over optical fibers. They are discussed beginning on page 233.

Time-Division Multiplexing

To make full use of the high information capacity of optical-fiber communication systems, many signal sources (such as telephone calls) must be merged together. This is accomplished by a process called *multiplexing*. The most common form of multiplexing is time-division multiplexing. A multiplexer combines many low-speed signals together in a stream of data by sending one part of one signal, then one part of another, and so on, switching among them at high speed. Because the speed of the multiplexer is so fast, it can keep up with the flow of incoming data from many signals.

At the receiving end, a demultiplexer breaks the incoming stream of data down, using codes embedded in the data stream, and reassembles the individual low-speed signals. When special network circuits are used to perform these functions, all of this happens with such small time delays that it appears to be occurring in real time. That is why no time delays are detected by callers speaking over long-distance optical-fiber-communications lines.

At present, time-division multiplexing is done in electronic form only. That is, all incoming signals are in electronic form and are multiplexed together in electronic circuitry. The final composite high-speed electronic signal is used to vary the optical output power of a light source by direct modulation. The speeds at which multiplexers must operate have pushed into the microwave region (data rates above 100 megabits per second, and frequencies above 1 gigahertz). Thus, microwave technology is required to realize these products, and microwave integrated circuits will be necessary to implement multiplexers with minimum cost. [For more information on microwave integrated circuits, see Chapter 7, Microwaves, beginning on page 147.]

In the future, it may be possible to multiplex incoming optical signals directly, using optical devices. Major advances in optical-component technology will be required to accomplish this goal.

Wavelength-Division Multiplexing

A second approach to achieving higher data rates on a single fiber is to send several lightwaves down the fiber simultaneously, each with a different frequency (or equivalently, a different wavelength). *Wavelength-division multiplexers* are devices that merge the different lightwaves together in a way that they do not interact with each other. Since each frequency is associated with a different wavelength of light, this process is referred to as *wavelength-division multiplexing*. At the receiving end of the system, the wavelengths must be separated for individual processing.

Each wavelength in a wavelength-division multiplexed system can carry data that has been assembled with time-division multiplexing, so both types of multiplexing can be used simultaneously.

Transmitting Lightwaves

Fibers and Cables

Optical fibers are thin strands of glass or plastic with a round cross-section. Glass is by far the dominant material; and silica, or silicon dioxide, is by far the dominant glass. Modern glass optical fibers have diameters of typically 125 micrometers. [A micrometer is 1×10^{-6} meter.] For comparison, a human hair has a diameter of roughly 70 micrometers. Plastic optical fibers are somewhat larger in diameter, typically 1 millimeter. Plastic optical fibers are highly resistant to impact and vibration and can employ lower cost sources and connectors, leading to lower installation cost. However, they provide lower information capacity and are suitable for use over short distances only.[6] The focus from here on will be on glass fibers, unless noted otherwise.

An optical fiber is composed of two dissimilar regions: a small round *core* that runs down the center of the fiber and a *cladding* that surrounds the core. Both are made of glass, but the core has been treated, usually with addition of germanium oxide. This treatment gives the core a higher index of refraction than the cladding.[7] This difference causes the cladding to reflect light back to the core and thus to confine light to the core. The fiber thus guides light down the core.

Before optical fibers are installed in systems, they are coated with a protective material called a "buffer" which is similar to the insulation on ordinary wire. Its diameter is typically 250 to 500 micrometers. Several fibers are often bundled together and covered by a protective sheaths to form *cables*. The cables generally contain strength members to assure that the fibers themselves are under no physical stress. Such stress can disturb their optical properties.

Optical fibers can be divided into two basic classes: multimode and single mode. They differ most obviously in the diameters of their cores. Modern multimode fibers have cores of typically 50 to 60 micrometers diameter. Single-mode fibers have cores of typically 7 to 10 micrometers diameter. The two fibers differ in other ways, some of which are traceable to the differences in core diameter. These other differences relate directly to desirable characteristics for the fibers. The differences are summarized in Table 83 and are discussed below. Single-mode fibers have the edge for the first two characteristics in the table, while multimode fibers win on the last three.

Table 83
Comparison of Multimode and Single-Mode Optical Fibers

Desirable Characteristics	Multimode Fibers	Single-Mode Fibers
high information capacity over distance	low	high
low cost	intermediate	low
simple sources	simple	complex
relaxed geometric specifications	relaxed	demanding
simple interconnections	simple	demanding

Multimode Fibers

Multimode fibers have some advantages relative to single-mode fibers. It is easier to launch a high level of optical power into the core of a multimode fiber because of the core’s larger diameter. That is, the larger core provides more area for accepting light; also, the core will accept light arriving over a wider range of angles relative to the axis of the core. This characteristic enables relatively large and low-cost light-emitting diodes to be used as light sources for multimode fibers for many applications. The larger core also relaxes the accuracy required in the geometric specifications of the fibers during manufacturing. Further, when joining two multimode fibers end-to-end at interconnections, the larger core relaxes the alignment that must be achieved between the fibers to assure the passage of a lightwave from one fiber to the other with minimum attenuation.

Single-Mode Fibers

Single-mode fibers have some advantages relative to multimode fibers. The small diameter of the core of a single-mode fiber permits a lightwave to travel only along a single path, or equivalently, in a single mode. This property gives single-mode fibers their most important advantage relative to multi-mode fibers: Single-mode fibers can carry digital light pulses farther without distortion. In contrast, the core of a multimode fiber is large enough that different parts of a lightwave bounce off the inside of the cladding at different angles. Thus, different parts of a lightwave follow different paths with different lengths and arrive at the receiving location at slightly different times. These different times of arrival broaden individual light pulses. This broadening increases with the length of a fiber and is one of several types of *dispersion* that afflict optical fibers. Dispersion caused by differences in path length is unique to multimode fibers. Dispersion may also be caused by differences in the speed of travel for different parts of a lightwave; this effect will be discussed beginning on page 227. However, all types of dispersion arise from different times of arrival for different parts of a lightwave, and all types of dispersion cause pulses to broaden increasingly with distance down a fiber. Single-mode fibers have a second advantage relative to multimode fibers. The single-mode fibers are less expensive to manufacture.

Single-mode fibers have some disadvantages relative to multi-mode fibers. Because the cores of single-mode fibers are so small in diameter, the cores must be aligned with remarkable accuracy at interconnections; otherwise light will be lost. To achieve such good alignment, all of the following must be true: the cladding must be made very round and very accurate in diameter; the core must be made very round and very accurate in diameter; and the core must be located precisely in the center of the fiber.

A second disadvantage of single-mode fibers is that they accept light only from sources of small cross-sectional area and only over a small angle relative to the axis of the core; that is, the light must not be too divergent. To produce light of this type, a laser must be used; and the laser's diameter must be comparable to that of the core of the fiber. Lasers are more expensive and less reliable than the light-emitting diodes that are commonly used with multimode fibers. [Lasers are discussed in detail on page 185 in Chapter 8, Lasers.]

Attenuation in Fibers

As a lightwave travels down a fiber, the strength of the lightwave is attenuated by scattering and absorption in the fiber. The scattering is caused by unavoidable irregularities in the fiber. The attenuation caused by scattering falls off smoothly with increasing wavelength (and thus with decreasing frequency).[8] The absorption behaves very differently; it peaks at certain wavelengths. The absorption is caused by impurities in the fibers, principally a water-related molecule. The impurities are impossible to remove totally. However, they have been reduced by modern manufacturing to very low levels. Between the regions of absorption there are three principal "windows" of low absorption. The wavelengths at which these windows occur are shown in Table 84[9] along with the attenuation levels and relative chromatic-dispersion levels achieved in good fibers operating near these wavelengths. Attenuation is lowest, about 0.2 decibels per kilometer, for fibers operating at 1550 nanometers because that wavelength is longest and thus experiences the least scattering. Operation at wavelengths longer than about 1.6 nanometers is prevented by a sharply arising absorption caused by the excitation of thermal vibrations in glass.

Other types of fibers are in the offing that will provide even lower levels of attenuation. A special *Z fiber* has been made. It uses pure silica glass in the core, plus silica glass containing fluoride glass in the cladding. By removing the germanium oxide from the core, scattering is lowered in the core, yet the fiber still confines light to the core because the fluoride glass lowers the refractive index of the cladding below that of the pure core. Also, other new types of fibers based entirely on the halides (particularly the fluorides) are in the research stage and offer the prospect of even lower attenuation. They push the wavelength caused by thermal excitation to longer values than the 1.6 nanometers of silica glass. This permits operation at longer wavelengths were scattering is reduced.

Table 84
Attenuation and Dispersion in Optical Fibers

Wavelength (nanometers)	Attenuation per Kilometer (decibels)	(percent)	Chromatic Dispersion (relative)	
850	1.81	52	fairly large	(natural form)
1310	0.34	8	zero	(natural form)
1550	0.19	4	fairly large	(natural form)
			zero	(dispersion-shifted)

Plastic optical fibers, for comparison, have losses of 65 to 124 decibels per kilometer. This high level of attenuation limits their use to short distances. They operate at visible wavelengths, which is convenient for providing light sources. A typical performance level for commercial plastic fiber is 10 megabits per second over distances of 70 meters.[10]

Dispersion in Fibers

The dispersion caused by optical paths of different lengths within multimode fibers was discussed above. Dispersion degrades a digital optical signal in a distinctive way. It rounds off the sharply rising leading edges of digital light pulses. It also rounds off the quickly falling trailing edges of light pulses. In the process, dispersion tends to broaden optical pulses. When dispersion is particularly great, the broadening of adjacent pulses may cause them to overlap each other. When this overlap becomes too great, the end of one pulse becomes impossible to distinguish from the start of the next pulse, and data are lost.

While single-mode fibers do not experience the high level of dispersion associated with the multiple path lengths in multimode fibers, they do experience two other types of dispersion: chromatic dispersion and polarization mode dispersion. These effects broaden light pulses in a similar manner.

Chromatic Dispersion

Lightwaves of different frequencies travel at different speeds down an optical fiber.[11] Thus, if a desired signal is impressed on a lightwave containing a wide enough range of frequencies, the components of the lightwave representing different frequencies will arrive at the receiving end of the fiber at different times, again broadening individual pulses. This effect will occur even though the components of the lightwave associated with different frequencies travel along the same path within the fiber. This type of dispersion is called chromatic dispersion, that is, dispersion caused by differences in color, or chroma.

The use of laser light helps to reduce chromatic dispersion because lasers produce coherent light that is nearly monochromatic (equivalently, single frequency, single wavelength, single color). In contrast, multimode fibers are often illuminated with light-emitting diodes; they produce incoherent light that may contain a broad range of frequencies. The breadth of the frequencies emitted by a light source is referred to as the *linewidth* of the source. As shown in Table 82, light-emitting diodes provide *wide* linewidths (about 60 nanometers or 10 terahertz). Fabry-Perot laser diodes provide *narrow* linewidths (about 2 nanometers or 300 gigahertz). Distributed-feedback laser diodes are more complicated and expensive but provide *very narrow* linewidths (about 0.00007 nanometers or 10 megahertz). Thus, distributed-feedback lasers have the greatest potential for reducing dispersion in fibers and thus for transmitting high data rates over long distances. The numerical values for the linewidths shown above are those obtained in an unmodulated (steady "on") condition.

Fortunately, optical fibers made of silica glass have a natural zero of chromatic dispersion at 1310 nanometers, as shown in Table 84. Because of this very desirable feature and because attenuation is lower at this wavelength than at 850 nanometers, most long-distance optical-fiber lines operate at this wavelength. However, specially designed optical fibers have recently been made with a zero of chromatic dispersion at 1550 nanometers.[12] This wavelength is another "window" of attenuation, with an even lower level of attenuation than 1310 nanometers. These new fibers are called *dispersion-shifted fibers*; that is, the zero of dispersion has been shifted from 1310 nanometers to 1550 nanometers. Many new optical-fiber systems will be designed to operate at this longer wavelength.

Although chromatic dispersion has been reduced to zero at specific wavelengths in optical fibers, lightwaves in optical-fiber systems will always experience some chromatic dispersion.[13] This is true because linewidths of light sources are increased, relative to their steady "on" values by either or both of two effects. The first effect is *chirping*; it occurs in laser diodes but not in light-emitting diodes. Chirping is a variation in the wavelength of a laser diode when it is directly modulated.[14] Chirping can be eliminated by the use of an external modulator, with an increase in system cost and a decrease in optical power output compared to direct modulation. The second effect is a natural broadening of the linewidth caused by the modulation process itself; this effect occurs regardless of the nature of the light source. Broadening caused by modulation cannot be eliminated. It increases with the modulation rate, that is, with the digital data rate. Therefore, a lightwave carrying information will always have a linewidth that is greater than the linewidth of a lightwave from an unmodulated (steady "on") source, no matter how narrow the linewidth of the source when unmodulated. In spite of this broadening during modulation, narrower sources are still better for reducing chromatic dispersion.

Polarization Mode Dispersion

Dispersion can be caused by polarization effects, too. Polarization is the orientation of the electric field of a lightwave with respect to the axis of the fiber (that is, with respect to the direction of travel of the lightwave). Unfortunately, in optical fibers, lightwaves of one polarization travel at a speed slightly different from the speed of lightwaves of another polarization. This difference arises from the fact that an optical fiber cannot be made perfectly axially symmetric. Even if a fiber could be made perfectly axially symmetric, the fiber could not be maintained in such an ideal condition during use, if only because of the distortions produced by unavoidable small strains along its length. Further, the polarization of the lightwave in an optical fiber tends to wander as it travels down a fiber. These effects lead to so-called polarization mode dispersion. This type of dispersion is small, however, and becomes a practical concern only for very long lines with lengths greater than 1000 kilometers.

Interconnections

Fibers are joined end on end in several ways. One way is to use *connectors*. They are the method of choice if the junctions must later be reopened. As shown in Table 85 below, good connectors introduce losses of 0.3 decibels or about 11 percent, in the strength of the transmitted light through the connectors. The losses may increase over time as the junctions degrade. Fibers may also be *spliced*, that is, joined with epoxy or with welds (fused) to form junctions that are permanent. Splices typically introduce losses of 0.2 decibels or less, or about 5 percent or less. Splices do not degrade over time.

Table 85
Loss Levels In Good Optical-Fiber Connections

Type of Connection	Loss in Decibels	Loss in Percent
connector	0.3	7
epoxy splice	0.2	5
fused splice	0.1 or less	2 or less

Splitting and Combining Light Paths

Incoming light paths may be split into two or more outgoing light paths by *splitters*. Similarly, two or more incoming light paths may be combined into one outgoing light path by *combiners*. Splitters and combiners are referred to collectively as *couplers*. Both types of couplers introduce losses in the optical power of the lightwaves passing through them.

Amplifying and Regenerating Signals

For long runs of optical fiber, attenuation from the fiber, from interconnections, and from components takes its toll. To restore the optical power of a lightwave, amplifiers are used periodically along the length of the run. Similarly, dispersion takes its toll. To restore pulses distorted by dispersion to their ideal "square" shape, regenerators are used.

The processes of amplification and regeneration have traditionally been accomplished together in devices called repeaters. To date, repeaters have been electronic devices, rather than optical devices, because a means of directly amplifying and regenerating light pulses has been lacking. When the electronic approach is used, an incoming lightwave is sent into a detector [as discussed further below, beginning on page 231]. The detector converts the lightwave pulses into an electronic replica of the optical power level in the pulses. The resulting electronic digital pulse train is processed to restore the pulses to ideal shape (very square) and thus accomplishing regeneration. The regenerated electronic pulses are used to modulate a light source. They impress their ideal shape once again on a lightwave of high optical power, thus accomplishing amplification. The new lightwave is sent on its way.

Direct optical amplification has been an industry goal for some time and has been accomplished and commercialized at the 1550-nanometer wavelength using amplifiers made from optical fibers doped with the element erbium. These amplifiers will appear in the next generation of submarine cable systems, beginning with a cable in the Caribbean in 1993 and followed by transatlantic and transpacific cables in 1996.[15] However, no commercial optical amplifier has yet been developed for 1310 nanometers. This wavelength is the one used in most installed long-distance systems; but progress is being made at 1550 nanometers, particularly with praseodymium-doped fibers[16] and recently with neodymium-doped amplifiers based on phosphate glasses.[17]

Optical amplifiers are much simpler devices than electronic amplifiers. They also are capable of amplifying several different wavelengths simultaneously and thus can support wavelength-division multiplexing much easier than electronic approaches. Optical amplifiers can serve four functions in optical-fiber communications systems as outlined in Table 86. They serve as (1) post-amplifiers to raise the output of laser diodes from milliwatts to tens of milliwatts; (2) in-line amplifiers along the length of long-distance lines; (3) in-line amplifiers in networks to compensate for losses caused by other components such as signal splitters; and (4) preamplifiers for light arriving at detectors, to help overcome the inherent noise in the detectors.[18] Optical amplifiers themselves have remarkably low noise. Optical-fiber amplifiers can provide gains of 20 to 40 decibels (factors of 100 to 10,000.)[19]

Optical amplifiers do not regenerate signals. To reduce the need for regeneration, fibers with a minimum of dispersion are very desirable. With dispersion-shifted fibers operating at 1550 nanometers, dispersion is so low that attenuation is the greater problem. That is, amplification is needed every 100 kilometers or so to overcome attenuation, but regeneration is needed only every 1000 kilometers or so. Thus, optical amplifiers can be used at 100-kilometer intervals, up to 1000 kilometers of total travel before regeneration is required.[20]

Table 86
Functions of Optical-Fiber Amplifiers

laser source post-amplifiers
long-haul in-line amplifiers
network in-line amplifiers
detector preamplifiers

Switching

Information from different light paths is dynamically routed from one path to another by *switches*. At the present time almost all switches are electronic. That is, at every switch the information on a lightwave must be recovered and converted to electronic form, and then switched electronically to another path. After switching, the information is converted back to a lightwave by modulating another light source.

In the near future, switching will be done optically. The first stage of optical switching will be electronically controlled optical switches; the first commercial implementations are just now emerging. These switches are optical devices that divert lightwaves to other paths without converting them to an electronic form first, but the control circuits are electronic. Later, the control circuits may also be optical in nature.

Recovering Information from Lightwaves

Wavelength-Division Demultiplexing

If several lightwaves with different wavelengths have been sent down a fiber, they must be separated before they can be converted to electronic form and processed. This separation is accomplished by a wavelength-division demultiplexer.

Detecting

Detectors convert the information from each lightwave of a given wavelength into electronic form so that it can be processed by electronic means. Detectors are assembled into packages with their associated electronics to form *receivers*.

There are two principal types of detectors used in optical-fiber communications systems. Both are semiconductor devices: positive-intrinsic-negative photodiodes (or PIN photodiodes) and avalanche photodiodes (or APDs). They are compared in Table 87 against a list of desirable characteristics for detectors. Generally speaking, the PIN photodiodes have lower cost and lower noise. However, they offer no gain. The avalanche photodiodes have high gain but also higher noise. A PIN photodiode, when used with an optical-fiber amplifier as a preamplifier to overcome the internal noise of the PIN photodiode, provides sensitivity and signal-to-noise ratio approaching the theoretical limits.

Table 87
Comparison of Detectors

Desirable Characteristics	PIN Photodiode	Avalanche Photodiode
low cost	low	high
simple electronics	simple	complex
low temperature sensitivity	low	high
high light sensitivity	low	high
low internal noise	low	high

Time-Division Demultiplexing

If an incoming lightwave contains data that have been interleaved from several sources by time-division multiplexing, then once the lightwave has been converted into an electronic signal by a detector, *time-division demultiplexers* must separate out the segments of the data associated with each of the original slower speed signals. Time-division multiplexing and demultiplexing are done only by electronic circuitry at this time, but optical approaches may be possible in the future. Optical time-division multiplexers and demultiplexers may look much like optical switches that are capable of very high-speed operation. An optical time-division multiplexer would interleave light pulses from multiple incoming paths, and an optical time-division demultiplexer would separate arriving light pulses into multiple paths.

HIGH-PERFORMANCE TECHNOLOGIES

There are two high-performance technologies in the research stage which may ultimately enable major improvements in optical-fiber communications systems. Each technology attacks one or more of the limitations in achieving *high information capacity over long distances*. The limitations arise principally from three sources: attenuation, dispersion, and lack of wavelength selectivity.

These high-performance technologies are *systems* technologies in this sense: The entire systems employing them must be designed to accommodate their special needs.

Soliton Systems

Solitons are light pulses with a special smooth shape. They arise naturally in optical fibers when very short pulses of light with high peak power are launched into optical fibers. Pulse widths below 100 picoseconds are used to stimulate the formation of solitons. [A picosecond is 10^{-12} seconds.] Solitons contrast with the "square" pulses used in present communications systems. However, the two types of pulses may contain comparable energy levels. The solitons just compress that energy into a shorter time interval and thus achieve higher peak optical power levels. A comparison of the advantages of square pulses versus solitons in provided in Table 88 and is discussed below.

Solitons have a remarkable property. They propagate without dispersion even in a medium that causes dispersion for square pulses.[21] That is, solitons maintain their shape with great accuracy over long distances. Their ability to do this is based on the mutual cancellation of two different effects: one effect makes the high-frequency components of pulses outrace the low-frequency components, and a second effect does the opposite.[22] Thus, the use of solitons raises the possibility of eliminating regeneration and of relying solely on the use of optical amplifiers.

The lack of dispersion enables solitons to provide higher information capacity over longer distances in optical fibers. At the present stage of research in the development of soliton systems, it appears that single-wavelength data rates of at least 20 gigabits per second will be possible at transoceanic distances without regeneration. Wavelength-division multiplexing will enable multiplying this rate further.[23] Note that solitons improve ease of *access* to the information capacity of a fiber, but they do not increase the information capacity of a fiber.

Solitons have some disadvantages, too. The light sources required to produce the very short pulses that give rise to solitons are somewhat complicated and expensive compared to those used to produce square pulses. Also, for solitons to retain their special shape, they must be maintained at a high power level. To maintain this level, optical amplifiers in systems using solitons must be spaced at distances that are about one-third of the spacing used for square pulses (30 kilometers versus 100 kilometers). Further, the gain of these amplifiers must be carefully equalized over all wavelength channels in a wavelength-division-multiplexed system.[24]

The use of solitons with long-wavelength (halide) fibers, which offer greatly reduced attenuation, may someday eliminate the need for amplifiers as well. The prospect of eliminating both amplifiers and regenerators from long-distance lines is sufficiently attractive to motivate further research. This prospect is especially attractive for undersea lines where maintenance of amplifiers and regenerators is especially costly, and where eliminating components requiring electrical power is very important.

Table 88
Square Pulses Versus Solitons

Advantage	Solitons	Square Pulses
lack of regeneration	✓	
access to higher information capacity over distance	✓	
simpler sources		✓
simpler amplifiers		✓
greater amplifier spacing		✓

Coherent Systems

Coherent systems employ a process called coherent detection to provide high sensitivity and high wavelength selectivity. High sensitivity is necessary to overcome attenuation in optical fibers so that noise in amplifiers and noise in detectors will not degrade the signal. High selectivity is necessary to separate closely spaced wavelengths so that more wavelengths may be passed through a single fiber in order to exploit more of the fiber's information capacity through wavelength-division multiplexing.

Coherent detection is more complicated than conventional detection. In conventional detection, incoming lightwaves fall on a detector that directly converts optical pulses into a matching train of electronic pulses. In coherent detection, an incoming lightwave from an optical fiber is *mixed* with a lightwave from a local laser light source, and the combination is applied to a detector. This process increases sensitivity and selectivity through the mechanisms described in the next two subsections. This process can also reduce signal degradation caused by dispersion. Coherent systems require light sources with very narrow linewidths and thus benefit from the use of transmitters employing external modulators to eliminate chirping.

This process of coherent detection is already used at radio frequencies in virtually all modern radio and television receivers. When a listener tunes a radio or changes the channel on a television set, the user is changing the frequency of an internal radio-frequency generator. This generator is the analog of the local laser light source mentioned above. The power output from the radio-frequency generator is mixed with the incoming off-the-air signal, and the mixture is sent to a detector. This process, when used with either optical or radio frequencies, is called heterodyne detection.[25]

Sensitivity

The high sensitivity achieved with coherent detection results from a special effect: The level of the detected signal is increased in proportion to the power from the local laser light source. Thus, a strong local source helps to overcome the inherent electronic noise within the detector, thus increasing the sensitivity of the detection system.

While the sensitivities achieved with coherent detection are remarkable and important, they are not as strong a motivation for coherent detection as they once were. The reason for this change is found in the emergence of optical amplifiers. When optical amplifiers are used as preamplifiers in front of detectors, the combination produces levels of sensitivity similar to those obtained with coherent detection. This high level of performance is possible because of the very low noise of the optical amplifiers and the detectors.

Selectivity

Coherent detection provides unparalleled wavelength selectivity. That is, coherent detection enables separation of very closely spaced wavelengths in wavelength-division multiplexing schemes.

Selectivity is a difficult goal to achieve. Generally speaking, optical wavelengths, like radio-frequency wavelengths, that are spaced no closer than 0.1 percent of their nominal values can be separated with relatively simple techniques, generally employing wavelength-selective filters alone. For example, in lightwave systems operating at 1550 nanometers, wavelengths spaced about 1.5 nanometers apart can potentially be separated by filters alone. However, the filters or other components needed for even these simple techniques are still in the research stage.

In comparison, coherent detection techniques can separate much more closely spaced wavelengths. To accomplish this separation, the local laser light source used in the mixing process described above is set for a frequency just slightly different from the frequency of the incoming lightwave. When the combination is applied to a detector, the detector produces an electrical signal whose frequency is equal to the *difference* between the frequencies of the two light sources. This difference frequency is usually chosen to fall in the lower part of the microwave region (1-20 gigahertz) to enable the difference frequency to be processed by microwave electronics. For the microwave difference frequency so created, even a 0.1-percent selectivity capability is adequate to separate very closely spaced wavelengths (and greater selectivity is available from sophisticated microwave electronics). That is, 0.1 percent of the microwave frequency of 20 gigahertz is very much smaller than 0.1 percent of the frequency of the incoming lightwave of 200 terahertz or 200,000 gigahertz.

An example is provided in Table 89. Instead of 66 channels spaced 1.5 nanometers apart, it is possible to have 6,600 channels spaced 2 gigahertz apart. This increased number of channels permits more convenient access to the information-handling capacity of a fiber. Very low rates of data can be sent efficiently on each channel. For example, a 2-gigahertz channel can easily carry data at a rate of 200 megabits per second. Thus, coherent detection provides a way of splitting down the information capacity of the optical fiber without the need for high-speed time-division multiplexing and demultiplexing electronics. Note that coherent systems, like soliton systems, improve ease of *access* to the information capacity of a fiber, but they do not increase the information capacity of a fiber.

Table 89
Channel Capacity of Conventional Versus Coherent Detection
(example at 1550 nanometers or 200 terahertz)

Characteristic	Conventional Detection	Coherent Detection
channel width (wavelength)	1.5 nanometers	0.015 nanometers
channel width (frequency)	200 gigahertz	2 gigahertz
channels per 100 nanometers	66 channels	6,600 channels

Dispersion

Coherent detection also helps in coping with dispersion. Because coherent systems can support many wavelengths, each carrying less information, the broadening of the linewidths of laser diodes caused by modulation can be reduced. This reduction in dispersion can be used to increase regenerator spacing in long-distance lines.

Conventional Versus Coherent Detection

The relative advantages of conventional detection and coherent detection are summarized in Table 90. Overall, coherent detection provides significant performance advantages by giving up simplicity in the optical components in return for simplicity in the electronic components. Major advances in optical components will be necessary before coherent detection will realize its full technical potential and become economically attractive. In particular, realization of optical integrated circuits will be needed to enable cost reductions for the somewhat more complicated optical devices required. Also necessary will be the realization of low-cost microwave integrated circuits. They are discussed in Chapter 7, Microwaves, beginning on page 147.

Table 90
Conventional Versus Coherent Detection

Desirable Characteristics	Coherent Detection	Conventional Detection
access to higher information capacity over distance	✓	
increased regenerator spacing	✓	
simpler supporting electronics	✓	
simpler sources		✓
simpler supporting optical components		✓

NETWORK BASICS

Optical-fiber communications systems realize their full potential for service when employed in networks. Networks interconnect users wishing to exchange information and can provide a variety of services that support that process, including, for example, error correction and data compression.

Applications

Networks are often broken down into categories by application, as shown in Table 91. The first three applications are generally associated with traditional telephone services. The latter two are generally associated with computer services. However, the differences are blurring in the face of intense worldwide efforts to achieve *interoperability* (compatible operation) across diverse network *architectures* (network designs) and diverse hardware implementations of networking equipment.

For applications associated with traditional telephone services, the subscriber loop interconnects the individual users of a network with the central-office switches of the telephone companies in two steps:

Table 91
Network Applications

Application	Alternate Terminology	Physical Extent
long-haul terrestrial	trunk lines	any distance on land
subscriber loop	local loop	several kilometers
undersea	submarine	any distance undersea
local-area network	-------	local, typically 2 kilometers
metropolitan-area network	-------	regional, typically 50 kilometers

feeder and distribution segments. The feeder segments run from the telephone company's central-office switches to remote terminals. The distribution lines run from the remote terminals to the customers' homes and businesses.[26] The central-office switches connect with other central-office switches serving nearby areas and with long-haul and the undersea lines for access to more distant locations.

For applications associated with computer services, networks interconnect primarily businesses, educational institutions, and government agencies. The local-area networks (LANs) generally run within or between buildings, or they run between larger facilities that are located near each other. They generally operate within an area about 2 kilometers in diameter,[27] although they can cover much larger areas. The metropolitan-area networks (MANs) interconnect local-area networks within a given metropolitan area and may interface directly with high-speed computers.[28] The metropolitan-area networks generally operate within an area about 50 kilometers in diameter.[29] There are more than 25 cities in the U.S. currently using metropolitan-area networks.[30] Some definitions of the metropolitan-area networks include cable television systems.[31]

Transmission Technologies

The nation's communications networks employ a variety of transmission technologies and will likely continue to do so for some time. These transmission technologies are outlined in Table 92. Radio-frequency waves include microwaves. These technologies are listed in order of increasing information capacity.

Table 92
Network Transmission Technologies

Signal Carrier	Transmission Media
conducted electronic signals	twisted-pair wires
radio-frequency waves	coax free-space
lightwaves	optical fibers

Most individual homes and businesses are served by twisted-pair wire telephone circuits. A smaller number of businesses and many educational and government organizations are served by higher speed transmission media. Long-distance signals are carried by microwave satellites and by optical-fibers;

both have very high information capacity. However, because the link to homes and businesses is usually low speed, most end users of telecommunications services do not have direct access to information services requiring high information capacity. Bringing such capacity directly to end users is now the subject of considerable national and international attention.

Advantages of Digital Techniques

Digital techniques are being employed in preference to analog techniques in almost all new network systems for a variety of reasons, summarized in Table 93. While digital techniques have been expensive to implement in the past, because of the complex circuitry that they require, advances in the development of large-scale integrated circuits and very-large-scale integrated circuits have dramatically reduced that cost.

Table 93
Advantages of Digital Techniques in Networks

accommodate diverse signal sources
enable correcting for limited degradation of signal
enable line sharing through time-division multiplexing
facilitate network control
support signal processing

Digital techniques enable reducing information from diverse signal sources to a common digital format. This common format facilitates the development of network architectures that can accommodate data from all types of sources, provided that their special differences can also be accommodated. These differences are discussed in the next section.

Digital techniques enable correcting for degradation of signal quality occurring during processes such as transmission and storage of information. Since digital signals assume only a small number of states that are readily distinguishable, they are relatively robust in the presence of noise. Through the process of regeneration, described above, digital signals can be restored to a pristine quality, provided that they have suffered only a limited amount of degradation. Also, digital signals facilitate the use of error-correction procedures. These procedures detect errors in transmission and achieve correction by retransmission or other methods.

Digital technology enables line sharing through time-division multiplexing. Data from multiple signal sources are readily interleaved once they are in digital form. Digital technology also facilitates monitoring the status of networks and controlling their functions.

Digital technology enables sophisticated signal processing such as signal compression to reduce the information capacity required to send a signal. At the transmission end of a system, signal compression reduces the total data that must be sent. At the receiving end, signal decompression restores the signal to its original form. The compression techniques may be lossless or lossy. Lossless compression techniques enable recovering the original signal without any change at all. Lossless compression is essential for computer data. Lossy compression techniques throw away some information that is not essential; in return, they achieve a higher level of compression. For example, lossy techniques for compressing audio signals throw away information that cannot be heard by the human ear, such as soft sounds masked entirely by loud sounds. Progress in compression techniques has been so significant of late that the information capacities required to transmit diverse signals have been reduced on a regular basis. Computer data and advanced video systems place the heaviest demands on system capacity and thus place a premium on data compression. Compression techniques for advanced video systems, such as high-definition television, are described beginning on page 349 in Chapter 11, Video.

Digital Signal Differences

While all digital signals may look alike, the requirements that they place on networks may vary considerably. The differences arise in information capacity, timing, and accuracy.

Information-Capacity Differences

An overview of the different information capacities required for various types of familiar services is provided in Table 94, according to data from Bell Communications Research.[32] The table shows that very different levels of information capacity are required for different services. However, for a given type of service, capacity requirements are being reduced by progress being made in the compression of data for transmission. Thus, the range of capacities listed for services such as conventional and high-definition television is rather wide to reflect both compressed and uncompressed approaches.

Table 94
Information Capacity Required by Communications Services

Service	Information Capacity
Residential	
telephone call (voice)	64 kilobits per second
conventional television (NTSC)	45-120 megabits per second
high-definition television (HDTV)	150-800 megabits per second
Business	
local-area network (LAN)	1-100 megabits per second
metropolitan-area network (MAN)	10-150 megabits per second

Timing Differences

The differences attributable to timing are outlined in Table 95 for services related to voice, video, and computer data. The *video* referred to in the table, and hereafter, is full-motion video, as opposed to still frames or images, which for the purposes described here can be considered as computer data.

Three types of timing concerns are addressed in the table. The first timing concern is *data rate*, and the question is whether a constant data rate is required. The second timing concern is the *source/destination timing relationship*, and the question is whether the source and the destination must maintain a fixed timing relationship with each other, even though that relationship is looser than constant data rate. For example, the source and destination may have to maintain a fixed-interval timing relationship; that is, the destination may need from the source a certain amount of information within each of a series of uniform time intervals, even though a constant data rate is not required. The third timing concern is time delay. Time delay, as used here, refers to a uniform delay in the delivery of all data, and the question is whether a uniform delay of a given size is acceptable. The time-delay answer may differ depending on whether the service provided is interactive (bidirectional, like a conversation) or non-interactive (unidirectional, like a radio broadcast).

Table 95
Digital Signal Timing Differences

Service	Data Rate	Source/Destination Timing Relationship	Time Delay Interactive /	Non-Interactive	Stream/Bursty
voice/audio					
uncompressed	constant	important	minimum	unimportant	stream
compressed	variable	important	minimum	unimportant	bursty
video					
uncompressed	constant	important	minimum	unimportant	stream
compressed	variable	important	minimum	unimportant	bursty
computer data					
file transfer	variable	unimportant		unimportant	stream
computer terminal	variable	unimportant	moderate		bursty

Uncompressed voice transmission for telephone calls requires delivery of a constant data rate to make natural sounding speech. Uncompressed voice transmission also requires a minimum of delay for natural interactions. Voice (or more generally, audio signals) that are being sent on a non-interactive basis still require delivery at a constant rate but are no longer sensitive to time delay. That is, a uniform one-second delay in the delivery of non-interactive information may not matter, but a one-second delay in the delivery of voice for phone calls (interactive) would be highly noticeable and objectionable to the communicating parties. Compressed voice for telephone calls (interactive) does not require a constant rate of delivery for some methods of compression. For example, there may be no data delivered at all when no one is speaking. However, when such data are transmitted, the source/destination timing relationship is important to permit decompression at the receiving end in a manner that produces a constant rate for the listener. Also, the time delay must still be minimal for compressed phone calls. The situation is similar for video. In video, for example, one method of compression sends only the changes in the video image from one moment to the next. These changes may be nothing, small, or complete changes of all picture elements (pixels). Whatever the changes, they must be delivered in time to reproduce each new frame of the picture within a fixed interval in time, so a source/destination timing relationship must still be maintained. Thus, while the compressed video data may not require a constant data rate during transmission on the network, a source/destination timing relationship still remains important. Also, compression and decompression processes take time and add their own delay, thus their use can increase the requirements on networks to deliver low delay.

For computer data, different needs arise for different applications. For networks that are connecting interactive computer terminals to host computers, the data rate may be variable and its precise value may be relatively unimportant at long as it is fast enough to serve the user's needs. However, the time delay tolerable may be only moderate (or even minimal) since the user may be looking for constant feedback from the host computer, much as the parties to a telephone conversation do. For background operations, such as file transfer, the data rate may be variable and its precise value may be relatively unimportant again, provided only that it is fast enough to serve the user's needs. Time delay is generally unimportant since file transfer is not an interactive process from the standpoint of the user.

These various differences in timing requirements are sometimes described in terms of *stream* data versus *bursty* data. Stream data are lengthy and relatively continuous in nature. Bursty data contain relatively short, sporadic bursts of data.[33] The terms stream and bursty are friendly and useful terms, but they are not very precise. As Table 95 shows, stream data may be constant rate (uncompressed voice) or variable rate (file transfer). There are several additional factors that complicate the characterization of data from a given source as either stream or bursty.

> *Network speed*: For very fast networks, such as those employing time-division multiplexing to interleave data from many sources, the data from all senders may look bursty because no single source can feed data into the network fast enough to require the network's continuous attention. For the same reason, data moving in a slow-speed part of the network may be stream data, but the same data may be bursty in nature when it enters a high-speed part of the network.
>
> *Network design*: The method used by a given network to handle a given type of data can affect its classification. For example, voice data are handled as stream data by some networks and as bursty data by others. This is true because some networks send a continuous stream of data whether a caller is speaking or not, uniformly filling up the quiet time with null data. Other networks fill the quiet time with data from other sources, providing a bursty nature.
>
> *Source non-idealities*: Data, even as it comes from the source, may not be consistently characterizable as stream or bursty. For example, file transfer data, while nominally continuous and thus generally characterized as stream data,[34] may be interrupted frequently. These interruptions may occur when data are read directly from a hard disk on which individual files are fragmented (disorganized) and unbuffered (unsmoothed by intervening high-speed electronics), a common occurrence.

For these several reasons, and perhaps others as well, the characterizations of stream versus bursty in Table 95 may be indicative, but they are not universal.

Data-Accuracy Differences

Different types of services may differ also in the accuracy required of the digital data delivered by the network. For example, computer data must generally be delivered with the highest accuracy, in fact, perfect accuracy. However, uncompressed voice and video data can often tolerate some errors. Compressed voice and video are generally more sensitive to errors since each unit of compressed data generally contains more information; therefore, any error occurring may have a greater effect on the uncompressed result.

Network Standards

Networks are among the most complicated of systems. To assure their proper operation, they are built to detailed specifications that are embodied in standards. These standards govern nearly every aspect of network operation. They may be thought of as falling into three principal groups: (1) hardware standards, (2) architectural standards, and (3) procedural standards, or protocols. The first group specifies the performance of the underlying hardware of the network (such as the wavelengths to be used) and the nature of the physical interfaces of the hardware to the network (such as plug design). The second group specifies the structure of the network and generally breaks that structure down into separable, but mutually supportive, functional units that are embodied in both hardware and software. The third group specifies the procedures used when comparable functional units of the network architecture interact. This latter group addresses such processes as switching and routing strategies and error-correction procedures. None of these three types of standards specify the internal design of either hardware or software of the network. Rather, the standards specify the behavior necessary for compatible operation on the network.

To support the design of network systems on an international basis, the International Organization for Standardization (ISO) developed the Open Systems Interconnection (OSI) reference model shown in Table 96.[35] This model provides an architectural standard for networks. The model calls each functional unit in the architecture a *layer*, and there are seven layers in the OSI model.

The lowest layer represents the hardware that actually transmits the data. Its requirements relate most closely to the hardware standards and thus to the measurement needs developed in this chapter for individual optoelectronic components in optical-fiber communications systems. The higher layers provide for packaging data, for controlling the network, and finally for services that users would recognize in support of their applications. Each layer relies on the lower layers for support.

For two stations on the network to communicate with each other, they must have the same set of layered functions. Also, corresponding (or *peer*) layers in the two systems must follow the same *protocols*, as noted above. Each layer in each station on the network communicates with the corresponding layer in another station by using both stations' corresponding lower layers. Thus, the physical layers of two stations come closest to having direct communication with each other.

The layers are structured so that equipment and software associated with each layer can be designed as independently as possible. The OSI model does not specify how the functions performed by a given layer are to be implemented.

The OSI model is instructive because it details the factors that must be considered in designing networks. Specific network designs may or may not employ the OSI model unaltered. But whether or not they do, the basic functions of the OSI model must generally be addressed.

NETWORK SWITCHING TECHNIQUES

The primary business of a network is to route data from a sender to a receiver. Over the years many switching techniques have been developed and implemented to accomplish this task. Current and contemplated approaches that seem to be of most importance now are outlined in Table 97.[36] The focus here is on techniques applicable to digital systems employing guided transmission media rather

Table 96
Open Systems Interconnection (OSI) Model for Networks by the International Organization for Standardization (ISO)

Layer No.	Layer Name	Layer Function
7	APPLICATION	Provides services directly to users. transaction server, file transfer protocol, and network management
6	PRESENTATION	Transforms data to support diverse services. encryption, text compression, and reformatting
5	SESSION	Controls connections among applications. establishing, managing, and terminating connections among applications
4	TRANSPORT	Assures reliable data transfer end-to-end in the system. end-to-end error recovery and flow control
3	NETWORK	Governs transmission and switching functions. establishing, maintaining, and terminating hardware connections
2	DATA LINK	Assures reliable data transfer in the hardware. synchronization, error control, and flow control for blocks of data
1	PHYSICAL	Governs hardware performance at the bit level. signal voltage swing; bit duration; mechanical, electrical, and procedural characteristics for establishing, maintaining, and deactivating physical connections

than free-space approaches, although both are supported by some techniques.

The switching techniques in Table 97 are currently used, or may soon be used, with one or more of three types of guide transmission media introduced above: twisted-pair wires, coax, and optical fibers. In the next major section, beginning on page 250, the selected network transport platforms and architectures employing the switching techniques will be discussed. That discussion will be easier to understand once the underlying switching techniques have been introduced.

The switching techniques in Table 97 may be grouped in different ways. One way is reflected in the numbered brackets on the left side of the table. Group 1 represents older techniques whose use is continuing into the present. Group 2 includes newer techniques that are in service now or that appear to have a certain and near-term future. Group 3 represents newer techniques whose future is not as clear but which offer noteworthy capabilities.

Each switching technique is described briefly below. The discussion is organized according to the headings used in the left column of Table 97. The techniques are described in a generic manner, although examples of real implementations are provided. There are two principal purposes for this generic approach: (1) to surface the fundamental differences between the techniques in a simplified,

Table 97
Alternative Switching Techniques for Networks

	Switching Architecture	Data Packaging	Distinguishing Path Characteristics	Data Favored
1	Dedicated (unswitched)	none	permanent, dedicated	stream
	Circuit Switching			
	Space Division	none	temporary, dedicated	stream
	Time Division	repetitive time-slot frames	temporary, shared	stream
2	Packet Switching			
	Datagram	packets	dynamic, shared	bursty, stream close second
	Virtual Circuit			
	Packet Switching	packets	temporary, shared	stream, bursty close second
	Frame Relay	frames	temporary, shared	stream, bursty close second
	Asynchronous Transfer Mode	cells	temporary, shared	stream, bursty close second
3	Wavelength Switching	none	temporary, dedicated	stream
	Broadcasting (unswitched)	none	temporary, dedicated	stream

and thus accessible, form; and (2) to support describing other network switching technologies as combinations of the generic ones. This approach is necessary because of the immense variety of real implementations. Because of that variety, not every basis for comparison can be addressed here; rather, certain defining differences have been selected for discussion.

A real network may employ different switching techniques in different parts of the network. For example, local interconnections may employ one switching technology, and long-distance interconnections may employ another.

Dedicated (Unswitched) Circuits

Dedicated circuits have been available for some time. They contain no switches to redirect data flow. Dedicated circuits are permanently available to the customers purchasing them and are unshared (from the perspective of the customer). That is, there is no competition for the use of the circuits at their full capacity at any time. However, dedicated circuits are costly and provide access to only one receiving location. Dedicated circuits are most efficient for stream data. In that sense they favor stream data, as noted in the right column of Table 97. They can carry bursty data, too, but they are underutilized when so doing.

Circuit Switching

Circuit-switched techniques reserve a fixed amount of information capacity for the user, but on a temporary basis; thus there is no competition for the use of circuit-switched interconnections once they are made. This characteristic is the hallmark of circuit-switched techniques. They employ a call set-up process that sets the switches in the network in order to determine the path to be taken by information. The lack of competition makes circuit-switched techniques very desirable for carrying data with the special timing requirements described in Table 95. In fact, virtually all voice data carried by the U.S. telephone companies are carried on circuit-switched lines. There are two principal variations of circuit switching: space-division and time-division.

Space Division

In space-division circuit switching, the sender is connected to a switch by a dedicated circuit. The switch is capable of connecting the sender to a variety of receivers, perhaps through a number of other switches and the lines interconnecting those switches.

Space-division circuit switching is most efficient for stream data, since any unused information capacity associated with gaps in bursty data would be wasted. Space-division circuit switching is particularly important for services with special timing requirements, such as voice.

The preeminent example of space-division circuit switching, using a so-called crossbar switch, has served in the subscriber loop of the telephone system for decades. Each telephone subscriber has access to a dedicated set of wires that connect directly to a cross-bar switch at the telephone company. This switch directly interconnects the subscriber's line with similar lines from other subscribers. This circuit-switched approach has been used for transmission media, like wires, that have limited information capacity.

Time Division

Time-division circuit switching is generally used for higher speed lines that can carry more than one call at a time. This variation of circuit switching uses time-division multiplexing. The data from each sender are grouped into units of fixed length called frames, and the units are interleaved on the high-speed system in a special way. Each sender acquires *temporary* ownership of a certain, repeating time slot in the multiplexed data stream, whether or not any data are sent in that slot. Because of the time-division multiplexing, the path is *shared* in the sense intended in Table 97. However, the time slot is not. Thus, the sender has unrestricted access to the full information capacity of a specified time slot, once the call set-up process has been completed.

The call set-up process tells the network switches how to route the frames falling in a given time slot. Thus, the frames are switched through the system based on their position in time alone. This approach has very low overhead, that is, it requires a minimum of *control data* mixed in with the information to be sent, the *payload data*, and thus requires a minimum of time to process that control data. The control data that are present are principally for establishing synchronization, that is, for assuring the position of the frame in time. Because of the low overhead, this switching technique works well for stream data and is especially important for data with the special timing requirements described in Table 95. This switching technique can be used with bursty data as well, but it is less efficient for two reasons: (1) if the connection is not broken between bursts, then any unused space within a given time slot cannot be allocated to other users; (2) if the connection is broken between bursts, then time-consuming call set-up signals are required to reestablish the connection for subsequent bursts.

The most important example of time-division circuit switching is TDM bus switching. It is used widely by the telephone companies and has virtually replaced the space-division switch.[37] The "TDM" refers to time-division multiplexing, but the use of this terminology in the name of this particular switching technique is not intended to suggest that other switching techniques do not use time-division multiplexing; many do.

In one respect time-division circuit switching is a precursor to the packet-switched techniques that are described in the next section: that is, the data in time-division circuit switching are "packaged" into frames to facilitate transmission over a network segment with much higher information capacity than any single user requires.

Packet Switching

The newer switching techniques receiving the most developmental effort at this time are all based on packaging adjacent data bits in one form or another. These packages are variously designated as packets, frames, or cells. They are referred to collectively as *packets*; and the techniques employing them are referred to collectively as *packet-switched techniques*. The OSI model, introduced above, is intended to support the development of protocols for packet-switched networks.

The switching systems employing packets treat the transmission media as transport *pipes* with high information capacity to be shared by all users sending information in the direction of the pipes. Packets from different users are intermixed by time-division multiplexers to achieve the high data rate necessary to take advantage of the information capacity of the pipes.[38] Thus, a given pipe in a packet-switched network is shared by many users.

When packets arrive at switches, they are stored momentarily, are checked for various types of information embedded in them, and are redirected along paths that ultimately take them to their destinations. At their destinations, the received packets are separated again through time-division demultiplexing and reassembled to provide a replica of the original data stream prior to packaging.

Packet-switched techniques have the limitation that individual users are not necessarily provided unrestricted access to a fixed amount of information capacity in the pipe and the switches, even on a temporary basis. Rather, special arrangements must be made to guarantee capacity, usually at a premium price. This is a major difference between packet-switched techniques and circuit-switched techniques. Therefore, when traffic is heavy on packet-switched networks, the network may experience *congestion*. That is, the network may slow down, from the perspective of a individual user, as it time-shares its limited information capacity among many competing users.

Packet-switched techniques have the virtue that they can accommodate users with different requirements for data rates and can even mediate between users employing different data rates on the same call through a process usually called speed conversion.[39]

Packet switching focuses on the efficient utilization of a pipe whose information capacity is limited but generally higher than that needed by an individual user. There are a variety of approaches to packet switching that are of current importance. They can be divided into two principal groups: *Datagram* and *virtual-circuit.*

Datagram

In datagram packet switching, each packet contains not only payload data but also enough control data for a switch to route the packet properly to its destination. For this reason, datagram packets are said to be fully addressed. There is no call set-up process for setting switches. Rather each packet is treated as a separate entity. Each packet is checked by each switch to determine where it is going, and the switch makes a decision on how to route the individual packet to the next switch on the way to the destination. That decision may differ for successive packets transmitted in sequence from the same sender. Thus, successive packets may follow different paths; and, in this sense, the routing is dynamic. Because of this routing approach, the packets may arrive out of order at the receiver. The receiver must temporarily store them and reassemble them in proper order. This reassembly process relies on a sequence code that is part of the control data in each packet.

The packets are checked for errors at every switch but are merely discarded if they are found to contain errors. The end points of the system must observe that a packet is missing, request a retransmission, and store other packets until the missing packet has been retransmitted, received, and inserted in proper sequence.

Datagram packet switching is especially efficient for bursty data. This efficiency derives in part from the lack of a need for the frequent call set-up signals for switch setting that would be required to support bursty data with some other switching technologies. Further, a high level of flexibility is possible in routing each packet to its destination through any available pipe in a complex network with multiple pipes; thus, efficient use is made of the overall information capacity of such a network. Datagram packet switching is also robust with respect to breakdowns in the network. Rerouting around such breakdowns is readily accomplished, and any packets lost in the breakdown process can be recovered through retransmission.

Datagram packet switching can support stream data, too, but with somewhat less efficiency because of the additional overhead built into each packet to assure its proper routing as an independent entity. However, in very fast networks, very little data, as noted above in the section, Timing Differences, is really in stream form.

Datagram packet switching is not well suited to data with the special timing requirements described in Table 95. The irregular time delays associated with the several processes of packet transmission mentioned above militate against achieving the three types of timing requirements in that table: constant data rate, a fixed timing relationship between sender and receiver, and minimal time delay. For this reason, datagram packet switching is generally not used for voice or full-motion video.

The Internet Network, employing the TCP/IP transmission protocol, is an example of a datagram packet-switched network.[40] It is a computer network created by the Defense Advanced Research Projects Agency and first implemented in 1980, primarily to interconnect government agencies and universities.[41]

Virtual Circuit

Virtual-circuit switching techniques, of which there are several varieties, overcome some of the limitations of datagram packet switching at the expense of some its virtues. Common to all virtual-circuit approaches is a call set-up process that sets all of the switches in the network. However, the circuit created is virtual rather than real because it does not exist as a physical interconnection but rather as a set of instructions stored in the switches. Those instructions tell the switches to route packets to a specific follow-on switch according to specific control data, called virtual-circuit identifiers, in the packets.

In a virtual-circuit switching system, all packets follow the same path from the sender to the receiver. Therefore, they arrive in sequence and do not give rise to the need for resequencing by the receiver. The data packets traveling a given virtual circuit are time-division multiplexed with data packets following other virtual circuits. Thus, the path is a shared one.

The virtual-circuit techniques generally favor stream data in transmission efficiency, relative to bursty data. The stream data can take advantage of the reduced overhead associated with a single call set-up process that stores the virtual path in the switches; this process compares favorably with the greater overhead required to make switching decisions for every packet individually in the datagram approach. However, in fast networks, as noted above, all data tends to look bursty anyway, so this distinction is not a major one.

Virtual-circuit switching is not as inherently robust against failures occurring along the network as datagram switching, since those failures can break the circuit entirely. However, sophisticated methods for redirecting virtual circuits in the case of failures are coming into use and are minimizing this advantage of datagram packet switching.

Three variations of virtual-circuit switching, and their principal differences, are discussed below: virtual-circuit packet switching, frame relay, and asynchronous transfer mode.

Virtual-Circuit Packet Switching

Virtual-circuit packet switching is the most basic form of virtual-circuit switching. Error recovery is performed at every switch along the path. That is, every packet is checked for errors at every switch. If an error is found, a retransmission is requested to replace the defective packet, before transmission to the next switch is initiated. The sequencing of the packets is strictly preserved by this process.

Perhaps the most prominent example of virtual-circuit packet switching is X.25, a designation of the CCITT. This particularly switching technique is used widely by the telephone companies.

Frame Relay

Frame relay is another variation on virtual-circuit switching. Data to be transmitted on the network are packed into frames rather than packets. Frames are generally larger than packets. These frames are different from the very simple frames used in time-division circuit switching. Among other differences, the frames used in frame relay must contain additional control data to identify the virtual circuit that they are to follow.

Frame relay handles error correction differently from virtual-circuit packet switching. Frame relay is designed to reduce overhead by taking advantage of the high reliability of modern networks in transmitting error-free data. Frame-relay checks each frame for errors at each switch, but performs no error correction at the switches. Rather, if an error is discovered, the entire frame is discarded (as in datagram packet switching). The decision to attempt a recovery from errors is left to the discretion of the sending and receiving systems at the end points of the connection. These systems may elect to implement error recovery when data accuracy is critical, as for most computer applications. Or they may elect to ignore error recovery when data accuracy is not as critical, as is often the case for voice or video transmission. Frame relay also implements techniques for improving the allocation of the information capacity of the network. The net effect of the differences, relative to virtual-circuit packet switching, is to reduce delay and thus to achieve higher throughput of information.[42]

Asynchronous Transfer Mode

Asynchronous transfer mode (ATM) is the third variation of virtual-circuit switching. ATM is still in the developmental stage, so its capabilities have not yet been fully realized.

ATM packages data into cells. The cells are smaller than either the packets or the frames used in the above switching technologies. They are also fixed in length (53 bytes), whereas packets and the frames used in frame relay may assume varying lengths. The fixed length and the small size of the cells used in ATM facilitate very rapid processing. The cells are designed for use in very high-speed systems. These high-speed systems offer the best possibility yet for accommodating data with special timing requirements, such as voice and full-motion video, in packet-switched environments.

Like frame-relay, ATM relies on the high reliability of the modern networks to reduce the need for time-consuming error recovery processes. Only ATM goes one step beyond frame relay in pursuing a reduction. ATM shortens the error checking process further by checking for errors only in the control data (header) of the cell, not in the payload of the cell. Cells with errors are discarded.[43] Error recovery may be accomplished end to end for a system using ATM cells throughout.

Wavelength-Division Switching

The time-division approaches used in circuit-switched and packet-switched techniques are focused on making highly efficient use of the information capacity associated with a single *carrier* in the network. In an optical-fiber network, this carrier is a single lightwave (with a specified wavelength) travelling through the fiber. The highest information capacity carried by a single lightwave in currently installed systems is just under 2 gigabits per second. This capacity is 0.1 percent of the capacity of the fiber itself, which is easily 2000 gigabits per second.[44] The practical limitation is set by the capabilities of available components.

There are alternate approaches when using optical fibers. They employ more than one wavelength, using wavelength-division multiplexing, in an effort to open up more of the information capacity of the fiber. They make efficient use of just one wavelength less important. Two of these approaches are described here: wavelength switching, and broadcasting. Because these techniques employ wavelength-division multiplexing, they have a *parallel* character; that is, they employ multiple lightwaves, each independent of the others, travelling together down the fiber. In contrast, the time-division-multiplexed circuit-switched and packet-switched approaches have a *serial* character; that is, they employ a single lightwave containing an interleaved sequence of data from multiple users.

Time-division multiplexing can be employed on each wavelength in a wavelength-division multiplexed system, so both serial and parallel techniques can be implemented simultaneously.

Wavelength Switching

Wavelength switching techniques assign individual users to different wavelengths. When a call is initiated, call set-up signals are sent to switches that create a virtual circuit that routes a specific wavelength to a specific receiver. This approach enables a given wavelength to be temporarily dedicated to a given user in a given part of the network. The same wavelength may be used simultaneously for another call in another part of the network (wavelength sharing).

One appeal of this technique is that it accommodates stream data with high efficiency and at very high data rates since there is no need for any address or virtual-circuit identifier within the data stream itself. The wavelength used is the only virtual-circuit identifier required.

This technique provides access to all of the information capacity associated with a given wavelength. For this reason, it is not subject to congestion and can accommodate the special timing requirements described in Table 95 for services such as voice and video.

This technique accommodates bursty data, but not very efficiently, since periods between bursts are not employed. However, there is no overhead required in the bursty data itself. With access to more of the information capacity of the optical fiber through wavelength-division multiplexing, efficiency in the use of a given wavelength recedes somewhat as a concern.

Time-division multiplexing can be used on individual wavelengths in a wavelength-switched system. For example, a certain wavelength may be temporarily switched to connect City A with City B; and all data between those cities may be time-division multiplexed onto that wavelength. Several of the time-division switching technologies introduced above could be used for this purpose. Of course, when time-division techniques are introduced on individual wavelengths, then the limitations of those techniques come into play.

Broadcasting

Broadcasting employs no switches at all. Instead, the optical fibers in a region are interconnected through couplers (combiners/splitters) to form enough branches to reach everyone on the network. Thus, a signal from any one sender is sent throughout the fiber network, much as a television station broadcasts to everyone over the air.

Broadcasting is generally considered a regional technique. Switching may be needed between regions to isolate signals using identical wavelengths in interconnecting regions. In fact, broadcasting may ultimately be used just for the subscriber loop, and other switching technologies may be used to interconnect the loops.[45]

Broadcasting has several advantages. In addition to eliminating switches, it readily accommodates services with the special timing requirements in Table 95, including those with requirements for high information capacity such as high-definition video. Broadcasting also facilitates addition of new services without a requirement for prior planning, and it requires much less capital investment in the network.[46]

Broadcasting has some disadvantages. Broadcasting, even more so than wavelength switching, requires access to an enormous number of different wavelengths since no wavelength sharing is possible within the broadcast network. Thus, very high levels of performance in wavelength separation are needed from supporting components. The payoff from the realization of coherent systems would be particularly high for broadcasting. Also, because broadcasting provides unrestricted access to all information on the network, encryption is necessary to protect information not intended for all users.

Broadcasting has been implemented in a unidirectional (non-interactive) form on a single wavelength by the cable television companies that use optical fiber. Since the networks employing this technology represent one of the fastest growing segments of the optical-fiber communications industry, the continued use of at least this variation of broadcasting seems assured. Another application of broadcasting is the Passive Optical Network (PON). This is a bidirectional (interactive), multi-wavelength implementation that has been the subject of substantial developmental effort, particularly in Great Britain.[47]

NETWORK TRANSPORT PLATFORMS AND ARCHITECTURES

Several designs for network architectures and for network transport platforms, which provide the *pipes* mentioned above, seem to be of importance at this time. They are summarized in Table 98, along with the maximum levels of information capacity that they offer and other information of relevance. The Synchronous Optical Network is a transport platform designed to serve in other network architectures. The other three entries in the table are network architectures.

Table 98
Optical-Fiber-Based Networks

Network Name	Abbreviation	Maximum Information Capacity (megabits per second)	Transmission Medium	Status
Narrowband Integrated Services Digital Network	N-ISDN	1.5	wires	in service
Broadband Integrated Services Digital Network	B-ISDN	622	fiber	future
Synchronous Optical Network	SONET	2488	fiber	entering service
Fiber Distributed Data Interface	FDDI	100	fiber	in service

These architectures and platforms are in various stages of evolution. Each is discussed briefly below and is related to the switching techniques employed. There are many other network architectures in place or emerging in this very dynamic field of communications.

Integrated Services Digital Network

The Integrated Services Digital Network (ISDN) is largely the child of the U.S. telephone companies and has matured into an international standard. Two versions of ISDN are under development. One version is for low information rates and is called Narrowband ISDN. The other version is for high information rates and is called Broadband ISDN. The major characteristics of either form of ISDN are properly reflected in its name. ISDN provides an integration of services, including audio, video, text, and computer data, through a digital network. That is, these several forms of data can be sent from, and to, any user in an ISDN system, for simultaneous presentation at the receiving end.

ISDN provides for a common interface between users and networks. This common interface enables manufacturers to make equipment compatible with ISDN. ISDN is intended ultimately to provide for universal access. Universal access means that a user who can connect to a network supporting ISDN and located anywhere in the world can communicate with anyone else in the world who can also connect to a network supporting ISDN. This is a very ambitious goal; even approaching it will require many more years of effort. ISDN must interface with diverse existing networking architectures used throughout the world to accomplish this goal.

Narrowband Integrated Services Digital Network

Narrowband ISDN is sometimes called Baseband ISDN, or simply, ISDN. This low-speed network architecture has just recently come into service over the telephone lines of many countries, including the U.S., but it is still evolving. An international protocol standard for Narrowband ISDN was first published by the CCITT in 1984 and was revised and expanded in 1988[48] and 1992.

Narrowband ISDN provides data at a maximum rate of 1.544 megabits per second to a single user (primary rate service). However, more typical levels of service to single users are 144 kilobits per second (basic rate service) or 64 kilobits per second.[49] For these relatively low data rates, wires may be used as the transmission medium.

Narrowband ISDN specifies the interface between the user and the network to which the user connects. The Narrowband-ISDN interface to networks is made through lines that have the characteristics of time-division circuit-switched lines although no switching is necessary to make this connection. Data traveling through this interface are carried in Narrowband-ISDN frames that provide both timing and time-division multiplexing. These frames are not themselves used in a switching process, and thus are to be distinguished from the frames used in frame relay. These frames are of fixed length, but different fixed lengths are used for different capacities of purchased service. The highest capacity service, 1.544 megabits per second, uses a frame containing 193 bytes of information.[50]

The networks to which Narrowband-ISDN users connect handle all of the switching that is needed to serve ISDN customers. They may employ a variety of switching technologies. Voice data from Narrowband-ISDN users is virtually always switched by time-division circuit switching in the networks. Computer data are usually switched by virtual-circuit packet switching in the networks.[51] However, virtual-circuit frame-relay switching for computer data is expected to come into wider use. The CCITT has recommended that this switching technique be used over the entire speed range covered by Narrowband ISDN, principally because of its higher throughput. Therefore, it will likely replace virtual-circuit packet switching as the dominant packet-switched technology supporting Narrowband ISDN.[52]

Broadband Integrated Services Digital Network

The Broadband Integrated Services Digital Network is also being developed as an international standard. The CCITT published the first standards for Broadband ISDN in 1990,[53] but the development of this new architecture is still in its infancy. The very latest results from research in network architecture are still driving its evolution.

Broadband ISDN will support the same data mix as Narrowband ISDN. However, Broadband ISDN will operate at much higher speeds. The current definition provides for a maximum data rate of 622 megabits per second to a single user.[54] Higher data rates may be supported in the near future. A more common level of service to a single user will be 155 megabits per second. Even this slower speed is 100 times the maximum data rate offered by Narrowband ISDN. While Narrowband ISDN can operate over wires, Broadband ISDN will require optical fibers to support its higher level of information capacity.

Broadband ISDN will be capable of supporting the most demanding services, including high-speed computer data and high-resolution, full-motion, full-color video, like that contemplated for high-definition television. Broadband ISDN will very likely employ Asynchronous Transfer Mode as its switching technology.[55] [The emergence of high-definition television is discussed in detail in Chapter 11, Video.]

Synchronous Optical Network

The Synchronous Optical Network, or SONET, is being developed by the Bell operating companies through their research arm, Bell Communications Research.[56] SONET, as noted above, is a network transport platform, specifically designed for optical fibers. It is intended to carry data that support various network architectures.

Phase 1, of several contemplated phases in the development of SONET, was implemented as a U.S. national standard by the American National Standards Institute in 1988. SONET is also being implemented by the CCITT as an international standard called Synchronous Digital Hierarchy (SDH). SONET systems, as defined in Phase 1, are currently being installed by U.S. telephone companies. At present there is every indication that SONET will be broadly adopted by U.S. telephone companies.

SONET carries data in very large frames of fixed size (810 bytes). Like the frames in the Narrowband-ISDN interface described above, these SONET frames are not themselves used in switching.[57] Rather, SONET frames will carry data packages that are switched by other switching technologies. Because of the promise that ATM cells show for high-speed processing, SONET will employ ATM cells. The ATM cells will be carried in the large SONET frames. The ATM cells, in turn, may carry packets or frame-relay frames, each split down to fit into sets of the ATM cells. At switches in the network, the ATM cells will be extracted from SONET frames. They may then be switched as ATM cells, or they may be reformatted into packets or frame-relay frames for switching in those formats. When such reformatting occurs, error correction will be conducted by requesting retransmission of entire packets of frame-relay frames found to contain defective data. These packets or frame-relay frames may then be fed into networks that are designed to carry them. Thus, SONET, in combination with ATM, provides not only transport services but an interface to networks employing other switching techniques.

It is this combination of SONET and ATM cells that the U.S. telephone companies will most likely use as the transport platform to implement the services of Broadband ISDN. In essence SONET and ATM will fulfill the requirements of the lower layers of the architectural model for Broadband ISDN, that is, those layers closest to the physical hardware.

SONET is defined for a variety of transmission speeds, ranging from 52 megabits per second to 2488 megabits per second.[58] SONET is designed for operation over single-mode optical fibers operating at wavelengths of 1310 nanometers or 1550 nanometers. As examples of specifications that relate to the hardware, SONET standards specify: the spectral characteristics of the lightwaves to be used at both 1310 and 1550 nanometers, such as linewidths and pulse shapes; the optical logic levels; and the interface power levels.[59]

Fiber Distributed Data Interface

The Fiber Distributed Data Interface (FDDI) is a network architecture embodied in a set of standards of American National Standards Institute. The first of several standards were published in 1987, and the last are now nearing publication. These several ANSI standards have been adopted internationally by the International Organization for Standardization, typically within a year or so after each ANSI issuance.

FDDI was designed to serve the needs of computer users principally and thus operates at a rather high speed: 100 megabits per second. The primary function of FDDI is to serve as a *backbone* network to interconnect local-area networks operating at lower data rates. However, high-performance workstations are sometimes attached directly to FDDI, particularly when they are serving graphics applications, which are highly data intensive. FDDI technology is being used in some metropolitan-area networks.[60] A variation of FDDI, called FDDI-II, is under development to accommodate services, such as voice and video, that present special timing requirements.[61]

FDDI employs a *closed ring* of optical fiber along which individual users or local-area networks are positioned at *stations* along the ring. Each station is thus connected to one neighboring upstream station and one neighboring downstream station by optical fiber. The data transmitted around the ring are contained in packets. Each station on the ring may initiate the sending of a packet. Each of the other stations on the ring, in turn, picks up this packet from the upstream neighbor and repeats it (retransmits it) to the downstream neighbor. The station, or stations, to which the packet is addressed copy the information in the packet and then forward the packet along. When the packet reaches the station of origin, that station terminates the packet.

Only one station at a time may originate packets. To determine whose turn it is, a special packet called a *token* is passed around. The station currently holding the token may transmit for a specified period of time. When finished, that station passes the token to the next downstream station, and so on. Because of this topology and mode of operation, FDDI is called a token-ring network.

Station-to-station interconnections, or *links*, on FDDI rings, may be spaced as far apart as 2 kilometers. FDDI supports rings that are 200 kilometers in circumference with several hundred stations. Links are often very short. For the shortest, under 100 meters, twisted pairs of wires are coming into use in the place of optical fiber. Thus, both optical fiber and twisted pairs may be used on the same FDDI ring in such a variation. SONET may be used as a link in the ring to accommodate more distant stations, at the expense of the time delay associated with long SONET lines.[62]

FDDI does not fall neatly into just one of the switching categories introduced above. Rather, it has characteristics of at least two. On an overall basis, it is much like a broadcast approach, since all stations on the ring handle the data transmitted, even though they do not all access it. FDDI also has aspects of a packet-switched system. There is no switching process in the usual sense; however, individual packets are addressed to specific stations and are copied by, and thus transferred into, the designated stations. However, as noted above, a receiving station does not terminate a packet, as in a regular packet-switched network; rather, this function is left to the transmitting station. Thus, FDDI works as a sort of time-shared broadcast network employing packets.

WORLD MARKETS AND U.S. COMPETITIVENESS

The world markets for components, equipment, and installed systems based on optical-fiber technology are vast. However, data are not available for all three of these product categories or for all regions of the world. Data for components are available and are presented in some detail below. No data have been located for equipment. Some data are available for the installed-cable part of installed systems.

Based on the data that are available and that are presented below, the world market for optical-fiber communications components is between $4 billion and $5 billion for 1992. Also, the worldwide cumulative investment in installed optical-fiber communications systems is already in the tens of billions of dollars. The worldwide cumulative investment in installed systems will be many hundreds of billions of dollars if fiber is carried directly to individual subscribers.[63]

World Market for Components by Application

The projected world market for optical-fiber communications components is shown by application in Table 99[64] for the years 1992 through 1998, according to data from Market Intelligence Research Corporation (MIRC). The network types used in Table 99 are the same as those introduced in Table 91. The MIRC definition of metropolitan-area networks includes cable-television systems based on optical fiber.[65] MIRC's estimate for the world market for 1992 of $4.4 billion may be compared with the estimate of Kessler Marketing Intelligence of $4.5 billion.[66]

For the world market as a whole, the MIRC data in Table 99 suggest growth from about $4 billion in 1992 to about $8 billion in 1998. This increase translates into a compound annual growth rate of about 11 percent over this six-year period.

An inspection of Table 99 leads to the rankings shown in Table 100.[67] Table 100 shows the anticipated changes from 1992 to 1998 in the relative market shares of the world market for optical-fiber communications components by application area. The shaded applications in the table are those

Table 99
World Market for Optical-Fiber Communications Components by Application (1992 to 1998)

	1992		1994		1996		1998	
Application	$millions	percent	$millions	percent	$millions	percent	$millions	percent
Long-Haul Terrestrial	2,216	51	2,517	45	2,730	40	2,679	32
Subscriber Loop	808	19	1,280	23	2,072	30	3,300	40
Undersea	667	15	807	14	696	10	381	5
Local-Area Networks	463	11	686	12	979	14	1,357	16
Metropolitan-Area Networks	213	5	301	5	421	6	613	7
TOTAL	4,367	*101	5,591	*99	*6,897	100	8,330	100

* rounding error

whose rank has changed from column to column, that is, from year to year, in the table. A rank of 1 has been assigned to the application consuming the most components, and so on. From 1992 through 1996, long-haul-terrestrial and subscriber-loop applications are expected to dominate in market share. However, in 1996 local-area networks are expected to surpass undersea applications. Then in 1998 subscriber-loop applications are expected to surpass long-haul terrestrial applications, as planned long-haul applications near completion worldwide. The first step in using fiber in subscriber loop is the installation of fiber to the curb (FTTC), which is already moving from the trial stage to deployment.[68] The second step is extending fiber all the way to the home (FTTH), which is expected to begin in 1994.[69] Metropolitan-area networks are expected to surpass undersea applications, as planned undersea installations are completed worldwide in 1988.[70] Thus, over the period 1992 to 1998, short-distance applications are expected to gain market share relative to long-distance applications.

Table 100
Ranking of World Market Shares of Optical-Fiber Communications Components by Application (1992 to 1998)

Rank	1992	1994	1996	1998
1	Long-Haul Terrestrial	Long-Haul Terrestrial	Long-Haul Terrestrial	Subscriber Loop
2	Subscriber Loop	Subscriber Loop	Subscriber Loop	Long-Haul Terrestrial
3	Undersea	Undersea	Local-Area Networks	Local-Area Networks
4	Local-Area Networks	Local-Area Networks	Undersea	Metropolitan-Area Networks
5	Metropolitan-Area Networks	Metropolitan-Area Networks	Metropolitan-Area Networks	Undersea

The changes in the relative market significance of the applications will be reflected in anticipated changes in the world market for components, as shown in Table 101, according to data from MIRC.[71] The table suggests that fiber cable will account for the largest share over the period 1992 to 1998, but that fiber cable will represent a gradually declining market share. Both transmitters/receivers and connectors/couplers will continue to increase in market share as they are called upon to serve in greater numbers in the subscriber loop and in local-area networks.

The MIRC data in Table 101 may be compared with data from Kessler Marketing Intelligence, providing a similar, but not identical, breakdown for the same year, 1992: $4.5 billion total composed of $3.0 billion for cable (67 percent), $1.3 billion for transceivers (29 percent), and $0.2 billion for connectors (4 percent).[72] Presumably, the inclusion of couplers with connectors in the Kessler data would move the percentage breakdowns somewhat closer to those of MIRC but would produce a somewhat higher overall market projection for 1992.

World Market by Component

The world market data for components in Table 101 are broken down further by application for the years 1992 and 1998 in three tables: Table 102[73] for fiber cable, Table 103[74] for transmitters and receivers, and Table 104[75] for connectors and couplers. Each table shows the market in both dollars and quantity. Each table also shows the average unit cost of each component used in each application and the average unit cost across all applications. The right column in each table shows the percentage reduction in cost anticipated for the components used in each application. The figure at the bottom of the right column shows the percentage reduction in cost anticipated across all applications. In each of these tables, the applications are listed in the same order as in Table 99.[76]

Table 101
World Market for Optical-Fiber Communications Components
by Component Group (1992 to 1998)

	1992		1994		1996		1998	
Component	$millions	percent	$millions	percent	$millions	percent	$millions	percent
Fiber Cable	2,555	59	3,136	56	3,574	52	3,868	46
Transmitters/Receivers	1,444	33	1,939	35	2,634	38	3,569	43
Connectors/Couplers	369	8	517	9	689	10	893	11
	*4,367	100	*5,591	100	6,897	100	8,330	100

* rounding error

Fiber Cable

In Table 102, the market for long-haul fiber cable dominates the 1992 market in both dollars and quantity, but the market for subscriber-loop cable is expected to approach the dollar value of the long-haul cable by 1998 and to surpass the kilometer value.[77]

Table 102
World Market for Fiber Cable by Application
(1992 and 1998)

	1992			1998			
Application	$millions	kilometers	$/kilometer	$millions	kilometers	$/kilometer	Percent Change in $/kilometer 1992 to 1998
Long-Haul Terrestrial	1,336	5,553	241	1477	7,376	200	-17
Subscriber Loop	376	2,138	176	1367	9,965	137	-22
Undersea	578	463	1,249	289	283	1,022	-18
Local-Area Networks	155	503	309	472	2,174	217	-30
Metropolitan-Area Networks	109	522	209	263	1322	199	-5
TOTAL	* 2,555	* 9,178	278	3,868	21,120	183	-34

* rounding error

Cable for undersea applications is very expensive, about four to seven times more costly per kilometer than cable for any other application. This cable must sustain high pressure, be water tight, have high strength to support installation and recovery activities, be corrosion resistant,[78] be resistant to physical attack by sharks,[79] and have high reliability.

Cable for local-area networks is somewhat more expensive than cable for applications other than undersea. This is true because multimode cable is generally the cable of choice for local-area networks.[80] While multimode cable was originally less expensive than single-mode cable, cost reductions achieved in manufacturing single-mode cable for the enormous long-haul market have made it the less-expensive product.[81] Nevertheless, multimode cable continues in use in local-area networks because the other components used with it are less expensive than those needed to support single-mode fiber. Therefore, system costs are reduced.

Transmitters and Receivers

For transmitters and receivers in Table 103, the market for long-haul terrestrial applications is the largest one for 1992, although, in terms of units used, the subscriber loop is virtually identical for that year. By 1998, the market for the subscriber loop is expected to be larger in both dollars and units, although the local-area networks are also expected to use a large number of units. The unit cost is much higher for the long-haul lines, in part because they use more expensive laser-diode transmitters to deliver high optical power levels into the fibers. The metropolitan-area networks use laser diodes primarily but also some light-emitting diodes.[82] In contrast, the subscriber loop uses light-emitting diodes primarily and relatively few laser diodes.[83] The local-area networks employ virtually all light-emitting diodes.[84] Also important, receivers of lower sensitivity often suffice for the subscriber loop and the local-area networks, further contributing to the lower unit cost for these components in these applications. Transmitters and receivers have the highest unit cost for undersea applications, in which high reliability is paramount to minimize costly repairs at sea. However, transmitter and receiver cost is small relative to the cost of the electronic repeaters used in undersea lines ($800,000 each).[85] The repeaters perform both the regeneration of digital pulses and the amplification process, as noted above. Hence the strong motivation, mentioned on pages 229 and 232, to eliminate the need for repeaters.

Table 103
World Market for Transmitters/Receivers by Application
(1992 and 1998)

	1992			1998			
Application	$millions	units(K)	$/unit	$millions	units(K)	$/unit	Percent Change in $/unit 1992 to 1998
Long-Haul Terrestrial	713	1,545	461	1,011	4,300	235	-49
Subscriber Loop	327	1,515	216	1,485	9,721	153	-29
Undersea	79	135	589	75	152	495	-16
Local-Area Network	241	1,891	128	701	10,379	68	-47
Metropolitan-Area Network	83	197	424	296	922	321	-24
TOTAL	*1,444	*5,282	273	*3,569	25,474	140	-49

* rounding error

The strong market sales for long-haul applications in the later years of the six-year period of 1992 to 1998 are attributable to maintaining existing lines at current performance levels and to upgrading existing lines with more powerful transmitters and more sensitive receivers to increase system throughput.[86] A further factor may be putting the so-called *dark fibers* into service. These are fibers installed before they are needed, alongside the in-service fibers, to provide for future capacity. An estimated 25 percent of all installed fiber is dark fiber at the time of installation.[87]

Connectors and Couplers

For connectors and couplers in Table 104, long-haul terrestrial applications dominate the 1992 market; but by 1998 subscriber loop applications are expected to surpass the long-haul applications in both dollars and units. Local-area networks have a strong presence in 1992 and will have an even stronger one in 1998. Once again, components destined for undersea use are more expensive than those for the other applications.

Based on data for 1991 for long-haul applications, the average unit cost for couplers is about fourteen times the average unit cost for connectors. Therefore, couplers can weigh heavily in the average unit cost of the combination of connectors and couplers, even when the couplers are present in relatively small numbers.[88] This is important to the market data for future years since couplers, particularly couplers in the form of wavelength-division multiplexers, are expected to be used in increasing numbers during the six-year period to increase information capacity without installing new fiber.[89]

The average prices of connectors alone, for 1992, have been estimated by Fleck International: $15 for single-mode connectors and $4 for multimode connectors. By 1996, these average prices are expected to fall to $8 and $2.65 respectively.[90]

Table 104
World Market for Connectors/Couplers by Application
(1992 and 1998)

	1992			1998			
Application	$millions	units(K)	$/unit	$millions	units(K)	$/unit	Percent Change in $/unit 1992 to 1998
Long-Haul Terrestrial	167	3,798	44	191	5,371	36	-18
Subscriber Loop	105	1,330	79	448	8,299	54	-32
Undersea	10	39	254	16	79	205	-19
Local-Area Network	67	1,489	45	184	5,958	31	-31
Metropolitan-Area Network	20	384	52	54	1,571	34	-35
TOTAL	369	*7,039	52	*893	*21,279	42	-19

* rounding error

Looking at Table 102, Table 103, and Table 104 together, it is evident that components for optical-fiber communications systems are generally expensive. Further, steep cost reductions must be achieved to generate the suggested market levels. These cost reductions are 19 percent to 49 percent, in *current dollars*, which are the type of dollars used in the table. Thus, the costs shown for 1998 reflect the addition of six years of inflation, which may add 30 percent or more.[91] When the effects of inflation are removed, the costs in 1998, expressed in *constant 1982 dollars*, must be reduced by factors of 1.6 to 2.5 over the six-year period, 1992 to 1998.[92] These fairly steep cost reductions present major challenges to manufacturers.

World Market for Components by Consuming Geographic Region

The projected world market for optical-fiber communications components is shown by consuming geographic region in Table 105[93] for 1992 through 1998, according to MIRC.

Over this period, the North-American share decreases somewhat; and the shares of the rest of the world increase somewhat. However, the decrease in the North-American share is limited for two principal reasons: Canada is still installing its long-haul terrestrial lines, even though the U.S. has largely finished that process; and in both countries, increases in the purchases of components for short-haul applications are expected to compensate to some degree for decreases in the purchases of components for long-haul applications. The short-haul substitutions will help keep market shares up in other countries of the world, too, even after those countries have completed their long-haul terrestrial lines.

World Market for Installed Undersea Systems

World market data located for installed systems is limited to undersea systems and is shown in Table 106.[94] The table shows the minimum levels of investment expected for installed undersea systems for several years, based on commitments already made. If less certain investments are included, the worldwide cumulative investment may reach $18 billion by 1998. For the five-year

Table 105
World Market for Optical-Fiber Communications Components by Geographic Region (1992 to 1998)

	1992		1994		1996		1998	
Geographic Region	$millions	percent	$millions	percent	$millions	percent	$millions	percent
North America	2,100	48	2,667	48	3,163	46	3,550	43
Europe	869	20	1,077	19	1,362	20	1,780	21
Pacific Rim	925	21	1,191	21	1,507	22	1,949	23
Rest of World	473	11	656	12	865	13	1,050	13
TOTAL	4,367	100	5,591	100	6,897	*101	*8,330	100

* rounding error

period beginning in 1993, the greatest percentages of investment in undersea systems are expected to be in the Atlantic (25 percent), the Pacific (21 percent), and the Far East (16 percent).[95]

Table 106
World Investment in Undersea Optical-Fiber Communications Systems

Year	Annual Investment ($billions)	Cumulative Investment ($billions)
1993	not available	8.8
1994	2.1	10.9
1995	1.1	12.0
1996	2.2	14.3
1997	0.5	14.8
1998	0.5	15.3

Undersea cable installations have been based principally on optical fiber since 1983 and totally on optical fiber since 1988. The first commercial installation was a trans-Atlantic cable in 1988 (TAT-8). Other major installations followed in 1989 (two) and 1991 (one). However, there were a number of earlier installations, principally experimental in nature: 1984 (four), 1985 (two), 1986 (one), 1987 (four). In addition, the total number of fiber installations recently completed or planned (as of October 1991), whatever their size, is quite large: 1988 (six), 1989 (thirteen), and 1990 (seventeen), 1991 (twenty-two), 1992 (fourteen), 1993 (fifteen), 1994 (four), 1995 (four), 1996 (three), and 1997 (one).[96] The longest undersea cable planned will link the United Kingdom and Japan by way of the Indian Ocean and the Mediterranean Sea (15,000 kilometers). It is planned for completion in 1996.[97]

U.S. Market

For the U.S. market, market data have been located for components and for the installed-cable part of installed systems, but no data have been located for equipment. The data that have been located and that are provided below suggest that the U.S. market for optical-fiber communications components falls in the range of $1 billion to $2 billion for 1992. Also, U.S. cumulative investment in installed optical-fiber cable is at least $11 billion through 1991.

U.S. Market for Components

The U.S. market is not broken out separately from the North-American market in the study by MIRC. One can only conclude from that study that the estimated U.S. market is less than the estimated North-American market of $2.1 billion for 1992. For comparison with this figure, ElectroniCast Corporation estimates the North-American market at $1.76 billion for 1992.[98] However, other sources of market data on optical-fiber communications components have reported U.S. market values. Those values are shown in Table 107.[99] The reflect a range of nearly 2 to 1 in the size of the U.S. market.

Table 107
U.S. Market for All Optical-Fiber Components (1992)

Source	U.S. Market ($billions)	Comments
International Trade Administration[99(a)] U.S. Department of Commerce	2.1	Reported in 1993.
Kessler Marketing Intelligence[99(b)]	1.52	Based on a 1991 estimate of $1.276 billion and a compound annual growth rate of 19 percent for 1991 through 1997. Reported in 1992.
Corporate Strategic Intelligence[99(c)]	1.14	Reported in 1990.

One of the sources referenced in Table 107, Corporate Strategic Intelligence, has provided a breakdown of the U.S. market by component. The results are shown in Table 108 for 1992 and 1994. They reflect a compound annual growth rate of 19 percent for the period 1989 through 1994.[100] The data in Table 108 are interesting because they separate transmitters from receivers, and connectors from couplers, suggesting splits of roughly 1.7:1 for the former and 1.8:1 or higher for the latter for the two years in the table.

U.S. Market for Equipment

U.S. market data located for telecommunications equipment do not distinguish between optical-fiber and electronic telecommunications equipment. However, U.S. shipments data and U.S. trade data suggest that the U.S. market for the combination of these categories is $16.8 billion for 1992. Of this sum, $9 billion was network equipment and $7.8 billion was customer premises equipment. The U.S. balance of trade is favorable for network equipment, unfavorable for customer premises equipment, and unfavorable for the combination at a level of $1.8 billion for 1992.[101]

Table 108
U.S. Market for Optical-Fiber Communications Components

	1992		1994	
	$millions	percent	$millions	percent
cable	650	57	988	56
transmitters	227	20	357	20
receivers	142	12	216	12
connectors	77	7	132	8
couplers, splicers, other	49	4	69	4
	1144	100	1763	100

The U.S. market for network equipment of $9 billion for 1992 may be compared with the U.S. market for SONET network transmission equipment of about $200 million for 1992, with projected growth to $1.8 billion by 1996.[102] Of course, there are other types of optical-fiber-based network equipment for which data have not been located, so it is not possible to determine the U.S. equipment market for optical-fiber-based equipment, based on the data presented here.

U.S. Market for Installed Cable

Data located on U.S. investment in installed systems is limited to installed cable. The data have been assembled by the Federal Communications for three of the five applications categories: long-haul terrestrial, subscriber loop, and metropolitan-area networks. The data reflect a cumulative investment of $11 billion for the years up to and including 1991, as shown in Table 109.[103] The data provide a lower limit for the three applications since data from some companies were not available, particularly for the metropolitan-area networks. The data reflect the total investment in fiber cable, deployment, and repeater sites (outside plant), and exclude electronic or optoelectronic equipment. The figure for the subscriber loop includes the interconnecting lines between main switching stations within the domain of a given regional Bell operating company as well as the lines running directly to subscribers.

Table 109
Cumulative U.S. Investment in Installed Optical-Fiber Cable Through 1991

Application	Cumulative Investment ($billions)
long-haul terrestrial	6.5
subscriber loop	4.5
metropolitan-area networks	0.1*
	11.1

* Lower limit since many companies did not report.

U.S. Competitiveness

The U.S., Japan, and Europe are all strong in research, development, and manufacturing for optoelectronic products broadly. However, Japan leads in both technology and commercialization, and that gap is widening.[104]

For the subset of optoelectronic products used for optical-fiber communications systems, the competitive situation is more difficult to assess. Data on world market shares of production for individual countries are generally not available. For the U.S., the lack of hard data on U.S. production of optical-fiber communications components results from the lack of separable shipments categories for these products in the Standard Industrial Classification System. Only data on fiber cable and bare optical fiber are separable.   However, some observations can be made.

Production of All Components

The U.S. is believed to be the largest producer of optical-fiber communications components as a whole, and Japan is believed to be the second largest producer, according to an examination made by the U.S. General Accounting Office in 1992.[105]

This conclusion may be compared with some observations based on Japan's production according to Japan's Optoelectronic Industry and Technology Development Association. That organization estimates Japan's 1991 production at about $1 billion for a subset of the products included in the MIRC world figure; that subset includes fiber cable (likely including bare optical fiber as well), light-emitting diodes for telecommunications purposes, connectors, couplers and assorted other components (multiplexers, demultiplexers, attenuators, isolators, modulators, and switches).[106] Data for Japan's production of other important components, such as laser diodes and detectors, cannot be separated from aggregated reported categories. Thus, Japan's production of all components of relevance likely exceeded $1 billion in 1991 with the other components included. That is, based on MIRC's estimate of the world market (taken as equal to world production) of $3.8 billion for 1991,[107] Japan's share of world production would exceed 26 percent of the world market for 1991. This conclusion is, at least, not overtly inconsistent with the conclusion of the General Accounting Office.

Although hard data are lacking, one market analysis organization, ElectroniCast, has attempted to estimate world market shares of production by region for 1992. The results are shown in Table 110.[108] This estimate provides a somewhat lower percentage for Japan than the above information. Unfortunately, the U.S. fraction is not broken out from the reported value for North America.

Production of Optical Fiber

Somewhat more data are available specifically for optical-fiber cable. World production in 1991 was about $2.29 billion according to Market Intelligence Research Corporation.[109] This figure includes only fiber cable and does not add in separately the value of bare optical fiber prior to cabling; thus there is no double counting.[110] U.S. production in the same year was $0.71 billion for fiber cable alone and some unknown subset of $0.46 billion for uncabled fiber ("unknown" because communications and other applications are reported inseparably in the category "uncabled fiber"). Both captive and non-captive production are included in these two figures.[111] Japan's production was $0.71 billion for its 1991 fiscal year for fiber cable (likely including bare optical fiber as well) for

Table 110
World Market Shares of Production of
Optical-Fiber Communications Components
by Geographic Region (1992)

Region	Percent
North America	53
Europe	30
Japan	17
Total	100

communications purposes.[112] These numbers suggest that U.S. production is somewhat larger than Japan's production, but it is not possible to determine by how much because of the possible double counting of bare fiber in Japan's figures. However, it would appear that the total production of the U.S. and Japan for fiber cable is $1.4 billion or less. This figure would suggest that two-thirds or less of the world's production of fiber cable is attributable to the U.S. and Japan. No conclusion can be reached about the production of bare optical fiber from the available data.

The U.S. is experiencing a positive balance of trade in fiber cable. While the precise ratio cannot be determined because of overlaps in U.S. trade data categories, the ratio of exports to imports for fiber cable is at least 2.7. Neither the balance of trade for uncabled fiber, nor even limits on that balance, could be determined from U.S. trade data because of the overlap of trade categories.[113]

International Plans

Both developing and developed countries have expansive plans for the use of optical-fiber communications systems. Here are some examples of recent plans and developments of importance.

United States

Within the U.S., long-distance and local telephone companies are pushing ahead aggressively with the installation of SONET.[114] Recent court action has cleared the way for these companies to provide information services, not just the transportation platform.[115] Similarly, a recent FCC decision permits cable-television companies to provide telecommunications services.[116]

Europe

The U.S. International Trade Administration reports the following: "Within the European market, Germany, Italy, and the United Kingdom lead the way in fiber optics installation, accounting for over 70 percent of the total market." "In Germany, the Deutsche Bundespost (DBP) Telecom has awarded large contracts for its ongoing implementation of FTTC [fiber to the curb]." The German goal is to reach all households within 15 years.[117]

The European Community is contributing funding to a three-year project called Mundi to "prove the viability of both components and network architecture" for both interactive and distributed services over passive optical networks employing wavelength-division multiplexing. The most elaborate system to be tested will employ 8 channels for one-way transmission, spaced 2 nanometers apart, each with a capacity of 700 megabits per second. The system will also have 16 channels for two-way transmission, spaced 1 nanometer apart, each operating at 155 megabits per second and each with a potential capacity of 700 megabits per second.[118] New opportunities appear to be arising in the countries of the former Soviet Union[119] and in Eastern Europe, too.[120]

Japan

Japan's Ministry of Finance is providing $7.7 million for industrial investment in Broadband ISDN and three-dimensional imaging. Nippon Telephone and Telegraph is increasing its research and development expenditures by $2.3 billion to meet growing competition from its telecommunications rivals.[121] Nippon Telegraph and Telephone had earlier planned to complete a nation-wide fiber-to-the-home system by the year 2015. However, NTT has since decided to let market forces drive the timing, with the presumption that business applications will be the first to mature in using fiber end to end. The estimated $266 billion cost of fiber-to-the-home was a major factor in NTT's decision.[122] The Japanese Ministry of Posts and Telecommunications forecasts that Broadband-ISDN services will cover 77 percent of Japan by 2015.[123]

Latin America

For Latin America, a proposed inter-American fiber optic network (IFON) would link all major cities.[124] The installation will be 21,000 kilometers (13,000 miles) long with termination points in 15 countries.

GOALS OF U.S. INDUSTRY FOR COMPETITIVENESS

U.S. industry is focusing on several goals to improve the competitiveness of the products that it makes for the expanding markets for optical-fiber communications systems. These goals are also important for realizing the full economic impact of the services that the new systems will provide. The goals are outlined in Table 111.

In the simplest terms, these goals seek a reduction in the cost per bit of information transmitted, while maintaining high reliability. Note the importance in particular of the development of optical integrated circuits. They will be essential in reducing the cost of components, particularly as more components are implemented in all-optical form.

Very important also is the effort to complete the development of voluntary standards for hardware, architectures, and protocols. Such standardization reduces system costs and promotes user acceptance and thus expands the size of the markets for which manufacturers compete. As suggested above, the hardware standards apply primarily to interface design and to performance at the interfaces. Other aspects of hardware design do not have to be standardized and remain the subject of intense competitive efforts.

EVIDENCE OF MEASUREMENT NEEDS

The National Conference of Standards Laboratories (NCSL), with broad participation from industry and U.S. Government agencies, has identified measurement needs critical to U.S. progress in many areas of technology. Among the high-priority areas identified is optical-fiber communications. NCSL

Table 111
Industry Goals for Improved Competitiveness of Optical-Fiber Communications Systems

Goal	Comments
Improved Performance	increased information capacity (higher bit rates per wavelength and more wavelengths)
Reduced Cost	new device designs that are manufacturable automation of manufacturing processes development of optical integrated circuits standardization of system hardware, architectures, and protocols
Improved Quality/Reliability	system hardware system software

notes the following: "Our [U.S.] success in this market will depend on our ability to produce low cost, high quality and high performance fiber optic systems and support equipment. This will require more accurate measurements of fiber optic parameters than are presently available from NIST."[125]

The Telecommunications Industry Association has formed several technical working groups, each composed of industry representatives. These working groups focus on individual components, or groups of components, of importance to optical-fiber communications systems. The groups have provided regular guidance on measurement needs critical to the development of the optical-fiber communications industry, and the groups have helped surface those needs which cannot be met by the industry. The working groups have given especially close attention to those measurement needs that must be met to support voluntary industry standards.

Biennial symposia, sponsored by NIST in cooperation with the Lasers and Electro-Optics Society of the Institute for Electrical and Electronics Engineers and with the Optical Society of America, review measurement developments critical to the industry and have contributed information helpful in identifying the primary areas of measurement shortfall.[126]

MEASUREMENT NEEDS FOR COMPONENTS

Many measured quantities require improved measurement support if the components required to support optical-fiber communications systems are to be realized. These components must serve in multiple network architectures and across many applications; and they must provide high performance, low cost, and high reliability. For a given measured quantity, the additional measurement support needed may take one or more of the following forms: measurement methods, measurement reference standards to assure the accuracy of those measurement methods, and materials reference data describing the properties of materials used in making components. For the measurement methods, the types of improvements needed vary with the quantity being measured. Important types of improvements are shown in Table 112.

Table 112
Types of Improvements Needed in Measurement Methods

- improved accuracy
- increased dynamic range
- higher frequencies
- increased simplicity
- documentation
- standardization of
  - - definition
  - - measurement method

The quantities requiring improved measurement support for optical-fiber communications components and materials are shown in several tables that follow. The coverage of each of these tables is shown in Table 113. This section focuses particularly on *components* and is organized by component. It describes the measurement methods and measurement reference standards that are needed for characterizing the performance or compatibility of specific components, including both those in service and those under development in industry. The next major section, beginning on page 281, focuses on measurement needs for characterizing important *materials* used in the components. The final major section, beginning on page 286, introduces the principal challenges that must be addressed to provide measurement support for *networks*. A full treatment of the specific measurement needs of networks is contemplated for a future edition of this document.

Table 113
Locations of Measurement Needs of Optical-Fiber Communications Components and Materials

	Table
Components	
fibers	Table 114
connectors (and splices)	Table 114
sources/transmitters	Table 115
modulators	Table 115
detectors/receivers	Table 115
optical amplifiers	Table 116
couplers	Table 116
multiplexers/demultiplexers	Table 116
isolators	Table 116
optical switches	Table 116
Materials	
semiconductors	Table 120
glasses	Table 120
other	Table 120

Fibers

Most of the measurement needs for optical fibers have been successfully met by recent efforts. However, a few measurement-related needs still must be addressed for multimode fibers; and several improvements in measurement capability are required to support single-mode fibers.

Multimode Fibers

For multimode fibers, standardization of the methods used for launching light into the fibers is needed to support a variety of measurements. The method used can affect the values obtained for such important quantities as attenuation[127] and bandwidth of fibers.

Also, an improved theoretical method is needed for predicting the modulation frequency bandwidth that will be obtained when two or more lengths of multimode fiber are connected in series, or concatenated. The bandwidth for a string of fibers may actually be greater or less than the bandwidth for any one of the individual fibers with the same length as the entire string. The bandwidth is expressed in terms of a so-called concatenation parameter. It describes an average rate of degradation of bandwidth with distance along the fiber; that is, the concatenation parameter describes the bandwidth as a function of length along a series of concatenated fibers, as if the series were homogeneous.

Predicting the bandwidth of even an individual fiber is difficult because of variability in launching conditions, as noted above. The launching conditions for a fiber in a test system may differ from those for a fiber in service. Predicting the bandwidth of a string of fibers from different manufacturers adds more uncertainty. The concatenation parameter is an important factor in the design of optical-fiber systems since overall system bandwidth must be accurately predicted.

Single-Mode Fibers

Three quantities used to characterize single-mode fibers need improved measurement support. These are shown in Table 114 and are discussed below.

Attenuation

Single-mode fibers have attenuation levels of typically 0.2 decibels (5 percent) per kilometer of length at 1550 nanometers. Present measurement capability is adequate if the measurements are made on fibers sufficient long (typically several kilometers) to accumulate attenuation at this level.

However, increased sensitivity for the measurement of attenuation will be needed to support the continued development of long-wavelength fibers. These fibers are made with halides, particularly those containing fluorine. They operate at wavelengths longer than 2 micrometers, where losses caused by scattering within the fibers can be very low. Theoretically, the long-wavelength fibers offer losses as low as 0.001 decibels (0.02 percent) per kilometer. Present measurement methods are not adequate to support loss measurements at such low levels, even in 10-kilometer lengths of fiber, let alone in the shorter lengths that are more practical for research. Measurement methods with sensitivities at least 100 times greater will be required to support research efforts. Lack of this measurement capability to date has not been a great impediment because present test specimens still have higher losses than conventional silica fibers.

Table 114
Measurement Needs for Optical-Fiber Communications Components, Part 1 of 3: Fibers, Connectors, and Splices

Component/Measured Quantity	Measurement Method	Measurement Reference Standard
fibers (single-mode)		
attenuation (long wavelength)	✓	
geometry		
cladding diameter		✓
dispersion		
polarization mode	✓	
chromatic		✓
connectors		
relative power		
return loss	✓	✓
geometry or ferrule		
diameter		
inside	✓	✓
outside	✓	✓
circularity		
inside	✓	✓
outside	✓	✓
concentricity	✓	✓
taper	✓	✓
axial parallelism	✓	✓
surface roughness	✓	✓
splices		
relative power		
return loss	✓	✓

Geometry

Measurements of many geometrical characteristics of optical fibers are important, but only one quantity requires additional attention: *cladding diameter*. The present measurement method is adequate, and a new measurement reference standard has been developed to support needed accuracy levels. However, the standard must be evaluated and then implemented in service. Accurate control of cladding diameter during manufacturing is essential for making low-loss interconnections between pairs of fibers and between fibers and other components.

Dispersion

Present measurement methods for chromatic dispersion are adequate. However, measurement reference standards, in the form of fibers, may soon be needed to support accurate measurements of chromatic dispersion in long, high-bit-rate systems operating at 1550 nanometers. Measurement methods with higher accuracy are needed for *polarization mode dispersion*. The need for new methods has increased as progress has been made in reducing chromatic dispersion.

Chromatic dispersion arises from two sources, from basic materials properties that are present even in bulk materials and from so-called waveguide effects, which have an opposite effect. These effects cancel each other out at about 1310 nanometers, although the contribution from waveguide effects is small. Thus, optical fibers operated at this wavelength have zero chromatic dispersion. For operation at 1550 nanometers, the dispersion-shifted fibers, mentioned above, have been developed. These fibers are manufactured with a special profile for the refractive index across their cross-sections.[128] This profile greatly increases the contribution made by waveguide effects and enables those effects to cancel the materials contribution at 1550 nanometers, providing a zero of chromatic dispersion at that wavelength.[129]

Concerns about both chromatic and polarization mode dispersion have been heightened by the emergence of optical-fiber amplifiers for 1550 nanometers. These amplifiers have dramatically reduced attenuation as a limitation to long-distance communication, leaving dispersion as the surviving principal limitation. Also, the use of higher data rates has broadened the total wavelength range within optical pulses, thus increasing degradation caused by chromatic dispersion. Further, the dispersion-shifted 1550-nanometer fibers are somewhat more sensitive to the effects giving rise to polarization mode dispersion.

Connectors and Splices

Optical fibers are joined in any of three different ways. Connectors are used to join fibers whenever they must be easily disconnected. Epoxy, surrounded by a protective and strengthening sleeve, is used to join fibers permanently, that is, to *splice* them. Fusing (melting) may be used to splice fibers. Splicing is cheaper and provides lower loss levels, so it is the method of choice when no need is foreseen for easy disconnections. Typical loss levels for good connections made in these several ways are shown in Table 85.

Relative Power

Return loss is a measure of the power reflected by an optical component. That power is lost to the transmitted signal. More importantly, the reflected power may be returned to the source and may adversely affect its performance. [This problem is discussed further in the section, Power, on page 271.] Measurement methods with higher accuracy are needed for determining return loss from connectors and splices as well as from a variety of components. Also needed is a measurement reference standard to support the accuracy of the measurements. This standard would provide a known ratio of reflected to incident optical power.

Geometry

Ferrules are the hollow rods that hold the fibers within connectors. Ferrules bring the fibers to be connected into alignment so that light can pass from one fiber to another fiber with minimum loss. Exacting control of the geometry of the ferrules is essential to their success in achieving the needed alignment. Measurement methods for determining the geometric properties of ferrules lack adequate accuracy to achieve the high performance levels needed. In additional, measurement reference standards are needed to support the accuracy of the measurements. The quantities needing improved measurement methods are several and include the inside and outside *diameter* and *circularity* of the ferrules, the *concentricity* and *parallelism* of the hole relative to the outside surface of the fiber, the amount of *taper* inside the hole, and the *surface roughness* of the ferrule.

Sources/Transmitters

Laser diodes are generally packaged with other components to form so-called *transmitters* that contain power-control electronics, thermoelectric cooling elements to remove heat, lenses to guide light into an optical fiber, and fiber pigtails to facilitate interconnection with other fibers. Measurement needs for sources and transmitters are summarized in Table 115. Both laser diodes and LEDs will be supported by meeting these measurement needs. However, the discussion that follows focuses primarily on laser diodes, since they give rise to the more difficult measurement problems.

[In Table 115, the "Notes" are provided for later reference in material beginning on page 281. They mark measured quantities for which improved measurement support is especially critical for the realization of two high-performance technologies: solitons systems and coherent systems.]

Power

Optical *power* output is the single most important characteristic of a semiconductor laser diode used for communications. Present measurement methods are adequate, but measurement reference standards are needed to assure the accuracy of those methods. The standards must support accuracy levels 4 times better (0.1 percent) and power levels 1000 times smaller (down to 10 microwatts) than possible with measurements made with direct use of present NIST national measurement reference standards.[130] Improved accuracy in power measurements will enable improvements in other important and closely related measurements, including those for (1) optical power output versus power-supply current, (2) slope efficiency, which is a measure of the slope of the curve relating the power output to the current, and (3) isolation, which is a measure of the optical power reflected back into a laser diode. This reflected power can impair the operation of laser diodes (and optical amplifiers, as noted below). The sensitivity is generally greater for the sources with the narrower linewidths.[131] This sensitivity is a major factor driving the need for improved measurement methods for *return loss* for many optical components. Since most components for optical-fiber systems are pigtailed, measurements for return loss are generally made on components in this form.

Threshold Current

Threshold current is the minimum power-supply current required to obtain coherent light output from a laser diode. There are several methods presently used to determine threshold current; there is a need to standardize the method used so that threshold current can be determined consistently. Resolving this issue is particularly important now because of the development of new low-threshold-current laser diodes. The new laser diodes offer reduced internal heating and thus may eliminate the need for costly thermoelectric cooling. At present, such cooling is required to assure the wavelength stability, power stability, and long lifetimes of laser diodes.

Beam Profile

Beam profile is the distribution of optical power across a surface perpendicular to the beam of light emitted by the laser diodes. Beam-profile measurements aid in determining how efficiently the beam

Table 115

Measurement Needs for Optical-Fiber Communications Components, Part 2 of 3: Sources/Transmitters, Modulators, and Detectors/Receivers

Component/Measured Quantity	Measurement Method	Measurement Reference Standard	Notes
sources/transmitters			
power		✓	solitons
threshold current	✓		
beam profile			
near-field	✓		
far-field	✓		
modulation			
linearity	✓		coherent
bandwidth	✓	✓	
impulse response			
pulse width	✓		solitons
rise/fall time	✓		
noise			
intensity	✓	✓	coherent
spectral			
purity (linewidth)	✓		coherent
wavelength	✓	✓	coherent
side-mode suppression (DFBs)	✓		coherent
modulators			
relative power			
return loss	✓	✓	
insertion loss	✓		
modulation			
linearity	✓		coherent
bandwidth	✓		
impulse response			
pulse width	✓		solitons
rise/fall time	✓		
switching voltage	✓		
detectors/receivers			
relative power			
return loss	✓	✓	
modulation			
bandwidth	✓		
impulse response			
pulse width	✓	✓	
rise/fall time	✓	✓	
spectral			
responsivity		✓	

of a laser can be directed into the small core of an optical fiber. Measurements of beam profile may be made up close to the laser diode (*near-field*) or far from the laser diode (*far field*). Present measurement methods have adequate performance, but they are rather complicated and expensive to use. Simpler methods are needed to provide more cost-effective measurements.

Modulation and Impulse Quantities

Most laser diodes presently in service in optical-fiber communications systems are directly modulated to place information on a lightwave; that is, the electrical current that powers the laser diodes is varied to produce changes in light output. Unfortunately, the variation in optical output power is not perfectly linear with (proportional to) the current. The departures from linearity must be measured to determine if a laser diode can serve in applications requiring high linearity. One of these applications is cable television which employs amplitude modulation to impress a radio-frequency signal directly on a lightwave.[132] To determine the departures from linearity, new measurement methods for *modulation linearity* are needed to provide improved accuracy and better sensitivity (increased dynamic range at the low end).

Another important quantity is *modulation bandwidth*. This is a measure of the variation of the depth of modulation (percent change in optical output power during modulation) versus the frequency of the modulating signal. Specifically, it is the frequency of modulation that is sufficiently high to produce a 3 decibel (50 percent) reduction in the depth of modulation. Present measurement methods are adequate for modulation frequencies up to 20 gigahertz. But new methods are required to reach modulation frequencies from 20 to 60 gigahertz. Further, a measurement reference standard is needed to support the accuracy of measurements of modulation bandwidth made at these frequencies.

Another way to determine modulation bandwidth is to measure the impulse response of a laser diode, that is, to measure how quickly the output changes in time. A fast-rising current, or a very short current pulse, is used to power the laser diode; and a measurement is made of the optical power output versus time to determine either the *rise/fall times* or the *pulse width*, respectively. Present measurement approaches cannot generate sufficiently short electrical pulses (widths under 20 picoseconds) and cannot measure the resulting optical output. Such high rates of change are necessary to test modulation bandwidths above 20 gigahertz. New measurement approaches will be required.

The importance of measurements for both modulation bandwidth and impulse response for modulation frequencies above 20 gigabits per second is presently limited. The reason is that modern supporting electronic systems cannot readily assemble (through time-division multiplexing) the data required to exploit modulation frequencies above 20 gigahertz. Further advances in the microwave electronics will be needed. However, measurements to support such fast rates may become important in the future. [For further information on microwave technology, see Chapter 7, Microwaves, beginning on page 147.]

Even without advances in time-division multiplexing, wavelength-division multiplexing offers additional information-handling capacity without having to exceed modulation rates of 20 gigabits per second for individual wavelengths.

Noise

Reduction of optical noise, which is optical power that is not a part of the signal, is a major goal to maximize transmission distance, maximize information capacity, and reduce error rates in transmitted data. This quantity, which is generally called relative intensity noise (RIN), can be measured with adequate accuracy now. However, development of a simpler measurement method would enable wider and more economical use. Also, a new measurement reference standard is needed to assure the accuracy of measurements of this quantity.

Spectral Quantities

A variety of spectral quantities require improved measurement support. For *linewidth* (or spectral purity), there are several measurement methods presently used that reflect somewhat different definitions of linewidth and that provide rather poor agreement. These methods need to be evaluated and compared to see if any one of them provides a superior combination of definition and accuracy, and thus motivates standardization. A new measurement method does not appear necessary now.

For *wavelength*, new measurement methods are needed to provide improved stability and accuracy for advanced applications, such as coherent detection. The new methods must also be simpler to use and faster for both advanced and present applications. Further, measurement reference standards are needed to assure that required accuracy levels are obtained at wavelengths critical to the operation of optical-fiber communications systems.

Distributed-feedback (DFB) laser diodes provide relatively sharp lines; that is, they concentrate their optical output power over a very narrow range of wavelengths. But DFB lasers also produce distinct, if relatively weak, sidebands. Sidebands are unwanted signals at nearby wavelengths. These sidebands may be 30 to 40 decibels weaker (a factor of 1000 to 10,000 times weaker) in optical power. Reduction of these sidebands is necessary to minimize degradation of the quality of an optical signal caused by dispersion. Present measurement methods are adequate for measuring sidebands if proper care is exercised. However, improved documentation would promote correct implementation of these methods with less effort.

Modulators

The measurement needs for modulators are summarized in Table 115. Several quantities require improved measurement support.

Relative Power

Like other components, modulators reflect some of the optical power that is incident on them. Measurement methods with higher accuracy, and a supporting measurement reference standard, are required for *return loss*. Modulators are generally manufactured in pigtailed form.

The loss of power through a modulator is determined by its *insertion loss*. Present measurement methods may be adequate, but the methods need to be compared and evaluated to determine if a single, common method, capable of providing uniform results, can be selected as a standard.

Bandwidth and Impulse Quantities

As with laser diodes, improved measurement capability is required for bandwidth, linearity, and impulse quantities for characterizing modulators. For *bandwidth*, measurement support for modulation frequencies above 20 gigahertz is again needed. For *linearity*, improved accuracy is needed. For *impulse response*, a physical source of very short and fast-rising electrical pulses is needed. In principle, an existing measurement concept seems sound; but it cannot be tested without this new source.

Switching Voltage

The switching voltage is the voltage required to reduce the transmitted output power of a modulator to a specified fraction of its full output level. The primary need is for agreement on the definition of switching voltage. This definition combined with existing measurement capability will then constitute a properly defined measurement method.

Detectors/Receivers

Relative Power

Measurements methods with improved accuracy, and a supporting measurement reference standard, are needed for determining the *return loss* from detectors/receivers.

Modulation and Impulse Response

The *modulation bandwidth* of a detector is the value of the modulation frequency of the incident lightwave that is sufficiently high to cause a 3-decibel (50 percent) reduction in the modulated electrical output of the detector. This quantity is also called the *modulation frequency response*. Present measurement methods are adequate for modulation frequencies up to about 20 gigahertz. However, above 20 gigahertz, a comparative evaluation of existing measurement methods is needed to resolve discrepancies and to promote standardization on a single method that will provide uniform results. In addition, the definition of detector bandwidth needs improvement. After completion of the comparative evaluation against the new definition, a measurement reference standard may be needed to support the accuracy of the selected method.

Measurement methods for determining the *impulse response* of detectors, as a means of inferring modulation bandwidth, are generally adequate down to about 20 picoseconds of rise/fall time and pulse width. However, for rise/fall times shorter than 20 picoseconds, a new measurement method is required. Also, for such short rise/fall times, a measurement reference standard is needed to assure the accuracy of the new measurement method. To enable establishing the performance of the standard, a light source capable of producing very short pulses is needed.

Spectral Responsivity

Spectral responsivity is a measure of the power output of a detector as a function of the wavelength of the incident light. Present measurement methods are adequate, but a measurement reference standard is needed to assure the accuracy of the measurements. The measurement reference standard would likely take the form of a detector with very flat spectral response. Characterization of the standard, and characterization of detectors generally for spectral responsivity, would be advanced by the development of a tunable source of light of high intensity and very narrow linewidth.

Optical Amplifiers

Optical amplifiers can be made from laser structures by reducing the reflections at their ends to levels as close to zero as possible. [See the discussion of the principle of laser operation in Chapter 8, Lasers, beginning on page 186.] This modification prevents lasers from generating laser light and enables them to amplify an incoming lightwave. Optical amplifiers for communications applications can be based on optical fibers or on semiconductors. The optical-fiber amplifiers have the advantage of fiber inputs and outputs and thus do not require a separate pigtailing process. Their performance is insensitive to the polarization of the incoming light. The semiconductor amplifiers have the advantage of ready fabrication as part of optical integrated circuits made from semiconductor materials; further, a wide range of wavelengths is readily accommodated. However, semiconductor amplifiers are somewhat polarization sensitive, so their output power can change as the polarization of the incoming light wanders. Both types of amplifiers are broadband; that is, they can amplify lightwaves over a relatively wide range of wavelengths.

The measurement needs for optical amplifiers and for the remaining components to be discussed are outlined in Table 116 and are discussed below.

Relative Power

Measurement methods with higher accuracy, and a supporting measurement reference standard, are needed for determining *return loss* from optical amplifiers, as for other components.

Gain is the ratio of the optical output power to the optical input power of an optical amplifier. Measurement methods for gain in optical amplifiers are generally adequate, even at small levels of the input signal and over adequately broad ranges of wavelength, although standardization on a single method by industry may be needed. However, a good measurement method for determining the gain at a single wavelength in the presence of other wavelengths is needed. This measurement method is increasingly important because of emerging interest in wavelength-division multiplexing. The measurement of gain is difficult for several reasons. The measurement technique must multiplex incoming wavelengths, and demultiplex outgoing wavelengths, with accurate control of all power levels. Also, amplifiers spontaneously emit some light at wavelengths other than those desired. Finally, unwanted reflections must be minimized because the high gain of optical amplifiers, particularly those based on optical fibers, increases their sensitivity to reflections in the measurement system, just as it does in practical systems. This sensitivity is a principal motivation for improved measurements of *return loss* for other components in optical-fiber systems.

Absolute Power

As the power of the optical input signal to an optical amplifier is increased, the gain of the amplifier remains unchanged at low levels of input and then begins to saturate at high levels. Measurement methods with higher accuracy are required for determining the absolute levels of input and output power so that the input *saturation* power level of the amplifier can be determined.

Table 116
Measurement Needs for Optical-Fiber Communications Components, Part 3 of 3: Optical Amplifiers, Couplers, Multiplexers/Demultiplexers, and Isolators

Component/Measured Quantity	Measurement Method	Measurement Reference Standard	Notes
optical amplifiers			
relative power			
return loss	✓	✓	
gain	✓		coherent
absolute power			
saturation	✓		solitons
noise figure	✓		coherent
linearity	✓		
couplers			
relative power			
excess loss	✓		
coupling ratios	✓		
return loss	✓	✓	
multiplexers/demultiplexers			
relative power			
excess loss	✓		
isolators			
relative power			
return loss	✓	✓	
forward loss	✓		
reverse loss	✓		
dispersion			
polarization mode	✓		
optical switches			
relative power			
return loss	✓	✓	

Noise Figure

Noise figure, in this case, is a measure of the unwanted optical power (the noise) added by an optical amplifier when amplifying a lightwave. Improved measurement methods are needed that are capable of distinguishing among the sources of the noise. In particular, noise added by the measurement system (its sources and detectors) must be distinguished from noise created by the optical-fiber amplifier. The very low level of noise in optical amplifiers complicates the measurement process.

Linearity

As for modulators, measurement methods with higher accuracy are needed for determining linearity in optical-fiber amplifiers. The distortion that results from non-linearity is particularly important in analog systems like those used for cable television where even several parts per million of distortion are not acceptable.

Couplers

Couplers may combine or split lightwaves, or do both simultaneously, in optical-fiber communications systems. They may have an arbitrary number of input and output ports. Couplers may be made from optical fibers. For example, a combination combiner/splitter may be made by placing two fibers alongside each other and by slightly fusing the region where they touch. The lightwaves entering the two input leads will be mixed and the combination will then be split and directed to the two output leads, so that each output has some of the signal from each input lead. Alternatively, couplers may be made as optical integrated circuits with pigtails. There are several quantities for which present measurement capability for couplers is not adequate.

Relative Power

For splitters, in particular, a measurement method with improved sensitivity is needed for *excess loss*. Excess loss is the difference between the optical power level entering the input port and the sum of the optical power levels coming out of the output ports. Also, a measurement method with improved accuracy is needed for determining the variation of the *coupling ratio* with wavelength and polarization. The coupling ratio is a measure of the relative power emerging from the output ports of a coupler or a multiplexer. The coupling ratio is also called the *splitting ratio*. Also, a measurement method with improved accuracy and a supporting measurement reference standard are needed for determining the *return loss*. The key challenge is separating reflected light from incident light. Further, a measurement reference standard is needed to assure the accuracy of measurements of return loss. Standardization of the measurement methods for all three quantities is necessary to assure consistent results.

Wavelength-Division Multiplexing/Demultiplexing

Wavelength-division multiplexing/demultiplexing techniques are just coming into use. Presently, most approaches use only two, or at most three, wavelengths, widely separated. The combining (multiplexing) of those wavelengths at the transmission end of the system and the separation (demultiplexing) of those wavelengths at the receiving end can be accomplished by a variety of different components, as shown in Table 117. Couplers that have been made wavelength dependent can multiplex or demultiplex lightwaves. These couplers act as combiners when multiplexing lightwaves and as splitters when demultiplexing lightwaves. Each of the internal optical paths associated with individual input ports of the combiners accepts a certain wavelength and rejects the others. Similarly, each of the internal optical paths associated with each of the output ports of the splitters does the same. Filters, both discrete and integrated (interferometers), and diffraction gratings provide much improved wavelength resolution and can also serve in both multiplexing and demultiplexing functions. Thus, all of these components can perform both functions; that is, they all operate reciprocally.

Since wavelength-division multiplexers and demultiplexers are still at an early stage of development, the full spectrum of measurement needs required to support them is not yet evident. However, the following measurement needs are already evident.

Table 117
Wavelength-Division Multiplexing/Demultiplexing Components

Components	Multiplexing	Demultiplexing
couplers	✓	✓
filters		
discrete	✓	✓
integrated (interferometers)	✓	✓
diffraction gratings	✓	✓

Relative Power

Measurement methods of higher accuracy, and a supporting measurement reference standard, are needed for *return loss* measurements in multiplexers and demultiplexers, as for other components. The improved measurement methods will support filters, both discrete and integrated (interferometers), diffraction gratings, and couplers made as optical integrated circuits.

Measurement methods for *insertion loss* are presently adequate but have not yet been standardized by the industry. However, improved measurement capability for insertion loss may be required as the spacing of adjacent channels is narrowed. In particular, sources of light with very narrow linewidths may be needed to support insertion-loss measurements at the several wavelengths supported by a multiplexer.

Isolators

Isolators reduce the optical power traveling backward in an optical-fiber system. This capability is particularly important to protect laser diodes from returned optical power. For this reason, laser-diode transmitters often contain built-in isolators. Isolators are also important in many systems that use complex circuits, since such circuits provide multiple opportunities for reflections.

Relative Power

Measurement methods with improved accuracy for determination of *return loss* are required to support isolators, as for many other components. Similarly, measurement methods with improved accuracy are required for measuring both *forward loss* and *reverse loss*. Forward loss is measured in the direction that the isolator is designed to pass light. Reverse loss is measured in the direction in which the isolator is designed to block light.

Dispersion

Isolators are constructed of components that are intrinsically polarization sensitive. The property used to achieve isolation is a polarization-sensitive property based on a magnetic effect called the Faraday effect. The Faraday effect causes a rotation of the angle of polarization (direction of the electric field) about the axis of travel of the lightwave in the presence of a magnetic field. The rotation is clockwise for one direction of travel through the material and counterclockwise for the reverse direction of travel, as viewed from the perspective of the arriving lightwave in each case.

Despite the fact that isolators depend on polarization properties for their operation, isolators are generally designed to work with arbitrary incoming polarizations. To accomplish this, the isolators split the incoming lightwave into its component polarizations, process each polarization separately, and recombine the component polarizations at the output. This is an exacting process; small path differences, that is, small differences in the length of the paths as seen by the lightwaves, can introduce *polarization mode dispersion*. Measurement methods with improved accuracy for polarization mode dispersion are required to support isolators, as for optical fibers as noted earlier.

Optical Switches

Optical switches are still in the earliest stages of development, although a number of promising technologies have been demonstrated. Their switching mechanisms may employ either guided waves within optical materials or free-space optics.[133]

The switches will support multiple input and output ports. They will likely employ interference effects to redirect lightwaves. For example, in one approach (the Mach-Zehnder interferometer), the switches employ parallel optical paths whose index of refraction can be altered by applying heat or electric fields (Pockels effect). The change in the index of refraction in a path changes the speed of a lightwave in that path. By permitting lightwaves emerging from different parallel paths to mix, the phase differences introduced by the different speeds of the lightwaves lead to interference effects that can cancel the lightwave coming out of one output port but not out of another. The port favored is determined by the applied heat or electric fields.

The optical switches of the future may use optical control signals rather than electrical control signals. They may rely on certain non-linear properties of optical materials for their operating mechanisms. Because optical switches are still under development, with the first commercial products just now emerging, the technologies that will be used are not yet clear. Thus, the measurement capability required to support them is only partially evident at present.

Relative Power

As for so many other components, a measurement method with higher accuracy is needed for *return loss*.

Speed and Synchronization

If optical switches can be developed for very high switching speeds, perhaps at gigahertz levels, then new measurement methods will likely be needed to achieve adequate accuracy in determining *switching speed* and *jitter*. Jitter is the short-term departure of digital signals from their ideal position in time. If in fact gigahertz speeds are realized in optical switches, then optical switches may be used for time-division multiplexing and time-division demultiplexing as well. As noted above these functions are presently performed entirely electronically.

Special Measurement Needs for High-Performance Technologies

Soliton Systems

Systems that use solitons require components with increased performance levels. Particularly important are laser-diode sources with narrow linewidths (good spectral purity), high power output, and very short pulse duration. Also needed are optical amplifiers with the high power output necessary to preserve the power levels required to assure propagation in the non-linear regime of the optical fibers. In presently contemplated designs, special filters are required, too, in order to combat degradation that gives rise to jitter.[134]

These requirements place special demands on supporting measurement capability. The measured quantities for which improved measurement capability is especially important to support solitons systems are annotated in the "Notes" column in Table 115 and Table 116.

The performance levels achievable with solitons have been demonstrated. Quite recently, Nippon Telegraph and Telephone reported that it had "successfully transmitted soliton signals over a distance of 1020 kilometers at a rate of 20 gigabits per second using erbium-doped fiber amplifiers." NTT used pulses with 10-picosecond duration. NTT believes that rates of 100 gigabits per second will be possible with pulses of 1-picosecond to 3-picosecond duration.[135] AT&T Bell Laboratories has demonstrated error-free transmission at distances exceeding 10,000 kilometers using solitons. AT&T has also demonstrated two-wavelength soliton transmissions over distances of 9000 kilometers.[136]

Coherent Systems

Coherent systems also require components with increased performance levels. Several of these demands are especially important. Laser-diode sources must have very narrow linewidths to support the several modulation methods that are especially attractive for coherent systems.[137] Linewidths of 100 megahertz are the largest tolerable, and linewidths of 1 megahertz or narrower are desirable for some modulation methods. Laser-diode sources must also exhibit high stability, ultra-low noise, and high linearity during direct modulation. The laser diodes must be tunable over wide frequency ranges, typically 10 percent of their nominal frequency, to obtain full benefits.[138] Any external modulators employed must have very high linearity. Polarization behavior in all components must be carefully controlled because the polarization of the local light source must match that of the incoming light. All active components, such as optical amplifiers, must have very low noise levels.

Once again, these several requirements place special demands on the supporting measurement capability. The measured quantities for which improved measurement support is especially critical for coherent systems are marked under the "Notes" column in Table 115 and Table 116.

MEASUREMENT NEEDS FOR MATERIALS

A very wide variety of materials are used in components for optical-fiber communications systems. Those materials can be divided into three major groups: (1) semiconductors; (2) glasses and other amorphous materials; and (3) other materials. The principal materials that are important to specific components are shown in Table 118 and Table 119.[139] These tables focus principally on materials used in commercial components, but selected materials used in developmental components are also shown. A few are labelled "experimental" to indicate the very early nature of the work employing them. The text discusses materials in both commercial and developmental components in detail. For optical amplifiers and lasers, the optically active dopant elements, which are responsible for generating the light, are shown in parentheses after the host material, for example: silica (erbium).

Table 118
Principal Materials in Optical-Fiber Communications Components, Part 1 of 2

Component	Semiconductors	Amorphous Materials	Glasses and Other Other Materials
fibers			
multimode		silica, plastic	
single-mode			
common		silica	
long-wavelength		non-oxide glasses (experimental) (e.g., fluorides)	
connectors			
bodies		plastics	
ferrules		plastics	ceramics
other components		index-matching fluids epoxies	
sources/transmitters			
low power			
laser and light-emitting diodes			
850 nanometers	gallium aluminum arsenide		
1310 nanometers	indium gallium arsenide phosphide		
1550 nanometers	indium gallium arsenide phosphide		
high-power			
solid-state lasers			yttrium aluminum garnet (neodymium)
1320 nanometers			
laser-diode pump			
800 nanometers	gallium aluminum arsenide		
modulators			
discrete			lithium niobate, lithium tantalate
integrated-optic	gallium arsenide		lithium niobate
detectors/receivers			
850 nanometers	silicon		
1310, 1550 nanometers	indium gallium arsenide		
850-1700 nanometers	germanium		

The most important of the *semiconductor materials* are the compound semiconductors, composed of various combinations of gallium, indium, aluminum, arsenic, and phosphorous. These are crystalline materials that are sometimes referred to as alloys because of the broad range of atomic-level mixtures that can be made. A variety of ratios of the constituent elements are necessary to support different wavelengths of operation.

The most important of the *glasses and other amorphous materials* is silica. Silica is used principally in fibers, couplers, and optical-fiber amplifiers. However, the fluorides and phosphates are proving increasingly important to developmental components.

Table 119
Principal Materials in Optical-Fiber Communications Components, Part 2 of 2

Component	Semiconductors	Glasses and Other Amorphous Materials	Other Materials
optical amplifiers			
fiber			
1550 nanometers		silica (erbium)	
laser-diode pump[139]			
980 nanometers	indium gallium arsenide		
1480 nanometers	indium gallium arsenide phosphide		
1310 nanometers		fluoride (praseodymium) [also fluoride (neodymium) and phosphate (neodymium)]	
laser-diode pump[139]			
1017 nanometers	indium gallium arsenide		
integrated-optic			
semiconductors	indium gallium arsenide phosphide		
glasses			
1550 nanometers		silica (erbium)	
1310 nanometers		phosphate (neodymium)	
couplers			
fiber		silica	
integrated-optic		silica	
multiplexers/demultiplexers			
couplers		silica	
filters			
discrete		silica	dielectric coating
integrated-optic			
interferometers		silica	lithium niobate, lithium tantalate
gratings		silica plastics (experimental)	dielectric coating
isolators			ferrimagnetic garnets
optical switches			
integrated-optic			
thermal		silica	
electro-optic	gallium arsenide		lithium niobate

In the category of *other materials*, two ferroelectric materials are especially important: lithium niobate and lithium tantalate. They provide the basis for both individual (discrete) and integrated-optic modulators and for integrated-optic switches. They may soon be used in many other components, too. One class of ferrimagnetic materials, ferrimagnetic garnets, is used for just one component: isolators.

The measurement needs associated with these materials are varied. They are outlined in Table 120.[140] In some cases adequate measurement methods already exist, and the principal needs are for measurement reference standards and materials reference data. The next several sections

Table 120
Measurement Needs for Materials Used in Optical-Fiber Communications Components

Component/Measured Quantity	Measurement Method	Measurement Reference Standard	Materials Reference Data
compound semiconductors			
alloy content		✓	
doping concentration		✓	
layer thickness		✓	
dielectric constant			✓
glasses			
electro-optic non-linear coefficients			
Kerr coefficient			✓
spectroscopic properties			
excited state absorption	✓		✓
stimulated emission cross-section			✓
composition			✓
lithium niobate/tantalate			
impurity concentration		✓	
dopant concentration		✓	
diffusion rates		✓	
solubilities		✓	
composition	✓		
crystalline structure	✓		
dielectric coatings			
refractive index			✓
layer thickness		✓	
ferrimagnetic garnets			
saturation magnetization			✓
saturation rotation			✓

describe the measurement needs that are especially important for the materials used in components for optical-fiber communications systems.

Compound Semiconductors

Measurement methods for *alloy content* of compound semiconductors are presently adequate, but measurement reference standards are needed to support the accuracy of those measurement methods. Similarly measurement reference standards are required to support measurements of *dopant concentrations*, which determine the concentrations of the charge carriers (electrons and holes) in these materials. Measurement reference standards are also needed to support measurement of *layer thickness*. For *dielectric constant*, materials reference data are needed for both the dissipative part of the dielectric constant (absorption) and the non-dissipative part (refractive index) for very thin materials. For example, data for the refractive indices in very thin materials are needed to support development of laser diodes and modulators based on new quantum-well technology. The properties of thin materials differ from those of bulk materials, so currently available data on bulk materials are not sufficient. Particularly important are data showing the variation of the dielectric constant with wavelength. Additional information on the measurement needs for semiconductor materials is contained in Chapter 4, Semiconductors, beginning on page 53.

Glasses

Measurement reference data are needed for the *non-linear optical Kerr coefficient*, which is a measure of the change in the refractive index of a material as a function of the square of the applied optical electric field. These reference data are needed to support the development of switches, modulators, multiplexers, and optical amplifiers.

Improved measurement methods and materials reference data are needed for the rare-earth dopants (such as erbium) that are embedded in glasses to give them special optical properties. In particular, improved measurement methods are needed for the *excited-state absorptions* of the dopant elements. Current measurement methods are difficult to apply and are not giving consistent results. Knowledge of excited-state absorptions is particularly important for the development of optical-fiber amplifiers and optical-fiber lasers. The goal is usually to minimize the absorptions.

Materials reference data are needed for both excited-state absorptions and cross-sections for stimulated emission. The *cross-sections for stimulated emission* describe the degree of interaction of the material with light of various wavelengths, including the pump wavelength and the wavelength of the incoming lightwave to be amplified. These reference data are needed to support the design of new optical-fiber amplifiers for 1310 nanometers, particularly those based on praseodymium and neodymium dopants embedded in fluoride glasses.

Finally, materials reference data are needed that relate the *compositions* of silica and phosphate glasses to the gain produced in optical amplifiers for operation at 1310 and 1550 nanometers. Glasses used in optical amplifiers may contain five or so major constituents (such as oxides of sodium, aluminum, and lanthanum) aside from the optically active dopant ions (such as erbium and neodymium). These other major constituents are present to influence both optical properties, such as operating wavelengths, and physical properties related to the "workability" of the materials during fabrication processes.

Lithium Niobate/Tantalate

Measurement reference standards are needed for both dopants and impurities in the ferroelectric materials, specifically lithium niobate and lithium tantalate, in order to support development of components such as modulators and optical switches. Impurities such as iron, even at trace levels, can degrade optical performance by introducing absorption.

Measurement reference standards are also needed for *diffusion rates* and *solubilities* of rare-earth ions, such as erbium and neodymium, in lithium niobate and lithium tantalate to support development of optically pumped integrated-optic solid-state lasers and optical amplifiers for both 1310 and 1550 nanometers.

Improved measurement methods are needed to probe the *composition* and *crystalline structure* of individual ferroelectric domains, measuring typically 5 micrometers on a side, in lithium niobate and lithium tantalate. The methods must support characterization of thin films; these films can be grown slowly to obtain better control of composition and crystalline structure than would be possible with bulk materials. This measurement capability is needed to support development of optically pumped non-linear parametric amplifiers. These amplifiers do not employ rare-earth ions, but rather use non-linear effects to enable the pump wavelength to power the amplification process occurring at the wavelength of the incoming lightwave. These amplifiers offer potentially higher gain and much higher efficiencies than rare-earth-doped silica-based amplifiers.

Improved measurement methods for composition and crystalline structure are also needed to support continued improvements in acousto-optically tunable filters for wavelength-division multiplexing. These filters are already offered commercially.

Dielectric Coatings

Measurement reference standards are needed to improve the accuracy of measurements of the thickness and refractive index of dielectric coatings. These coatings are critical to the design of filters and gratings employed in multiplexers and demultiplexers.

Ferrimagnetic Materials

Measurement methods for characterizing ferrimagnetic materials are basically adequate, but materials reference data are needed. Of particular importance are data on the saturation magnetization versus temperature, and the saturation rotation versus both temperature and wavelength.[141] These data are needed to support the selection of ferrimagnetic material compositions during the design of isolators. The ferrimagnetic materials in isolators are normally operated with their rotation in saturation.

CHALLENGES FOR NETWORK MEASUREMENTS

Measurements made on networks serve several functions, as outlined in Table 121. They may be used for commissioning a network, that is, for qualifying the performance of the network just after installation. They may be used for performance monitoring while the network is in service. Or they may be used for fault finding when specific difficulties are detected in network performance.

Table 121
Functions of Network Measurements

commissioning
performance monitoring
fault location

Many of the measurements needed for characterizing the performance of individual components can also be used to characterize networks. However, networks present challenges that differ considerably from those of individual components. These challenges are outlined in Table 122. Some of the challenges arise from the fact that networks, when measured for performance monitoring and fault location, are in service at the time; that is, they are carrying communications data. Further, portions of the networks are often located in uncontrolled environments (underground, overhead, or undersea). In contrast, components, when measured, are out of service and are generally located in controlled environments (research laboratories, testing laboratories, or manufacturing facilities). Further challenges arise from a combination of the complexity, diversity, and extent of networks.

Table 122
Challenges for Network Measurements

Challenge	Explanation
transparency	non-interference of measurement systems with communications traffic
environment	accommodation of hostile environments of networks
access	limitations to physical access for measurement systems
longevity	long-term stability (aging) of components and measurement systems
spatial extent	tens to thousands of kilometers
temporal extent	nothing "instantaneous" at optical-fiber data rates
complexity	multiplicity of components with diversity of interactions
architecture	architecture-specific requirements on measurement systems

This diversity of challenges makes the area of network measurements a subject of its own. For this reason, a separate and detailed treatment of network measurement needs is contemplated for a future edition of this document to supplement the more general introduction that follows.

Each of the measurement challenges in Table 122 is discussed below, and selected examples are provided that are important to networks based on the switching technologies that are currently being used, or that are contemplated for use, in optical-fiber communications systems. These switching technologies include the packet-switched networks generally (and especially the virtual-circuit approaches) plus the wavelength-switched and broadcast networks.

Transparency

Measurements on networks must generally be made in a manner that does not interfere with communications data and is thus said to be *transparent* to communications data. The reason for this requirement is that normal data flow cannot be interrupted without causing great inconvenience and cost. Such measurements can be accomplished with different techniques. One approach is to mix digital test data with communications data. Another approach is to generate test wavelengths that run in parallel through the network without interfering with communications data carried on normal network wavelengths.[142] This latter approach is especially suited to wavelength-switched and broadcast networks, but it can be used over individual fiber segments of virtual-circuit frame-relay networks, too.

Environment

The environment of networks, and thus of the measurement equipment used to characterize them, can be very hostile with thermal cycling over temperature ranges of -40 °C to +65 °C, wide humidity variations,[143] vibration,[144] and even ionizing radiation. These conditions are difficult not only for network equipment but also for measurement equipment, and they can complicate the measurement process by introducing changes in lightwave properties, including optical output power, phase, and polarization.

Access

Access to optical-fiber networks is limited by the overt physical constraints of lines that are buried in the ground, carried overhead, or located undersea, inside walls, or inside plenums. Access is also limited by the fact that fibers are difficult to tap. Thus, the same feature that makes fibers secure as data channels also complicates the measurement process. To get around this access problem, fibers can be injected with test light using temporary so-called clip-on splices. However, these splices are difficult to make without introducing either permanent or transitory disturbances into current data.[145] Development of improved methods for temporary access is a major challenge for making measurements on networks.

Longevity

Networks are in service for long periods of time; thus measurement systems installed for permanent monitoring must maintain their performance over long periods of time, as well, and must sense long-term changes. Suitable measurement systems must therefore have high levels of stability, and often high levels of accuracy and sensitivity, as well.

Physical and Temporal Extent

Networks can have enormous physical extent. Long-haul terrestrial and undersea lines are primary examples. This great physical extent complicates end-to-end network measurements. Networks can also have considerable temporal extent, whatever their length; that is, at the high data rates used in optical-fiber communications systems, nothing happens at a rate that approaches instantaneous, even though lightwaves travel through the networks at a significant fraction of the speed of light in air. For example, for networks operating at a data rate of 1 gigabit per second, individual data bits are spaced about 20 centimeters apart as they move down an optical fiber. Thus, a noticeable time evolution is reflected down the length of the fiber.

Timing is everything in virtually all high-speed network architectures and is somewhat more complicated in network architectures employing time-division multiplexing or buffering at electronic switches. There are at least two timing parameters that are of great importance to networks: *delay* and *jitter*. Delay is caused by normal transit times and by the buffering. Delay extends the time required for any bidirectional interactions and is especially critical for interactive services such as voice communications. *Jitter* is the irregularity in the arrival times of incoming digital pulses. Jitter complicates synchronization, which is the process of accommodating the timing of the incoming pulses. This accommodation is essential to enable distinguishing the boundary between one digital bit and another, particularly in a stream of bits of the same value.

Timing, or synchronization, must be established at one of several levels of accuracy, depending on the network architecture. Jitter increases errors in the interpretation of digital data bits, contributing to the *bit error rate*. The bit error rate is the single most important quantity for characterizing the overall performance of a digital network. The bit error rate is simply the fraction of bits that are in error. Thus, bit error rate is a measure of the success of the network in accomplishing its primary mission: delivering data unaltered. Bit error rates as high as 10^{-9}, that is, 1 bit in error out of 1 billion, are sometimes permitted in networks for raw, uncorrected performance of network hardware. However, lower bit error rates are increasingly specified for higher-speed networks. Bit error rates of 10^{-12} are starting to be specified for FDDI systems. Further reduction of errors, to virtually zero, is achieved through sophisticated encoding practices and associated software and hardware. Together, they make detecting errors very reliable and thus support the decisions that the network must make on retransmission of packets, frames, or cells.

Complexity

Because networks are complex, with a multiplicity of components interacting in diverse ways, measurement methods used in networks to determine performance levels and failures, like measurement methods used in many other complex systems, must accommodate a diversity of possible situations. In such environments, measurements must be used in conjunction with *test strategies* that have been designed ahead of time to support efficient and meaningful probing of network performance.

An example of the nature of the measurement problem in complex systems, end-to-end measurements of network performance can show an out-of-specification condition even when all individual components within the system measure within specification.[146]

Architecture

Finally, the measured quantities of importance for networks, and the methods used to measure those quantities, are often dependent on the architecture of the networks being measured. For example, in broadcast networks, conventional optical time-domain reflectometry (a sort of optical-radar approach) is not effective in locating defective components. This is true because these broadcast networks subdivide lightwaves into multiple paths, and combine lightwaves from multiple paths, using couplers; and these paths are indistinguishable when reflections are returned.[147] In contrast, optical time-domain reflectometry is suitable for measurements on point-to-point connections of the type that interconnect adjacent switches in virtual-circuit frame-relay networks.

Similarly, architecture can complicate bit-error-rate measurements and associated fault location. For example, in virtual-circuit frame-relay networks, error recovery for frames is not conducted at individual switches intermediate to the path, but rather is left for the path end points. This approach relies heavily on low error rates in intermediate components. It also means that measurement strategies for these networks must cope with a greater number of components, simultaneously, when seeking to locate error sources.

ENDNOTES

1. Mathematically, the wavelength is defined as the speed of light divided by the frequency. The speed of light in a vacuum (free space) is 3 x 10^8 meters per second. In an optical fiber, the speed is somewhat lower. The frequency of light does not change as it passes from one medium to another, but its wavelength does. The speed of light in a medium is found by dividing the speed of light in a vacuum by the index of refraction of the medium.

2. S. L. McCall and N. K. Dutta, "Diode Lasers Invigorate Communication Technologies", *Laser Focus World*, p. 75 (October 1992).

3. The term electromagnetic interference is associated with so-called *non-ionizing* radiation. In contrast, *ionizing* radiation, which is characterized by higher frequencies and higher energy levels, can darken glass somewhat.

4. Jeff Hecht, "Laser Action in Fibers Promises a Revolution in Communications", *Laser Focus World*, p. 81 (February 1993).

5. P. A. Morton, R. A. Logan, T. Tanbun-Ek, P. F. Sciortino, Jr., A. M. Sergent, R. K. Montgomery, and B. T. Lee, "25 GHz Bandwidth 1.55 µm GaInAsP p-Doped Strained Multiquantum-Well Lasers", *Electronics Letters*, Vol. 28, No. 23, p. 2156 (November 5, 1992).

6. Ken Eben, "Plastic Optical Fiber", *Lightwave*, pp. 36-40 (November 1992).

7. As noted in endnote 1, the index of refraction in a medium, like glass, is the number by which the speed of light in a vacuum must be divided to get the speed of light in the medium. Typical values of the index of refraction in optical materials are between 1.5 and 4.

8. The attenuation is caused by Rayleigh scattering. This scattering is the result of local variations in the refractive index attributable to the irregular arrangement of atoms in a glass and to differences in the sizes of those atoms. Rayleigh scattering decreases with the fourth power of the wavelength, λ, that is, as $1/\lambda^4$.

9. The loss levels in decibels per kilometer were provided by Corning Glass.

10. Ken Eben, "Plastic Optical Fiber", *Lightwave*, pp. 36-40 (November 1992).

11. This is equivalent to saying that the refractive index varies with frequency.

12. This result was achieved by varying the index of refraction along the radius of the fiber’s cross-section in a special manner.

13. C. F. Cottingham, "Dispersion-Shifted Fiber", *Lightwave*, p. 25 (November 1992).

14. Jeff Hecht, "Long-Wavelength Diode Lasers Are Tailored for Fiberoptics", *Laser Focus World*, p. 88 (August 1992).

15. Jeff Hecht, "Laser Action in Fibers Promises a Revolution in Communications", *Laser Focus World*, pp. 75, 81 (February 1993).

16. Jeff Hecht, "Laser Action in Fibers Promises a Revolution in Communications", *Laser Focus World*, pp. 80-81 (February 1993).

17. Current research at NIST.

18. Jeff Hecht, "Laser Action in Fibers Promises a Revolution in Communications", *Laser Focus World*, pp. 75, 81 (February 1993).

19. Jeff Hecht, "Laser Action in Fibers Promises a Revolution in Communications", *Laser Focus World*, p. 79 (February 1993).

20. At 1310 nanometers, where losses in the fibers are greater than at 1550 nanometers, amplification is needed at more frequent intervals. In a currently operating system (AT&T's FT Series G), typical values are these: 40-kilometers spacing between repeaters (which also provide amplification); 20 decibels of loss (factor of 100) between repeaters; 1.7 gigabit-per-second data rate; less than 10^{-12} bit error rate. S. L. McCall and N. K. Dutta, "Diode Lasers Invigorate Communication Technologies", *Laser Focus World*, p. 75 (October 1992).

21. Linn F. Mollenauer, James P. Gordon, and Stephen G. Evangelides, "Multigigabit Soliton Transmissions Traverse Ultralong Distances", *Laser Focus World*, p. 160 (November 1991).

22. The first effect is chromatic dispersion. At wavelengths longer than 1310 nanometers, this effect causes the high frequencies in a light pulse to travel faster than the low frequencies. The second effect is the increase in the refractive index of the fiber at high light intensities. This non-linear effect gives rise to a so-called self-phase modulation that causes the low frequencies to travel faster than the high frequencies. Operation of fibers at 1550 nanometers, which is a wavelength longer than 1310 nanometers, enables these two effects to cancel each other out so that soliton pulses maintain their shape over distance. Hermann A. Haus, "Molding Light Into Solitons", *IEEE Spectrum*, p. 50 (March 1993).

23. Hermann A. Haus, "Molding Light Into Solitons", *IEEE Spectrum*, p. 53 (March 1993).

24. Linn F. Mollenauer, James P. Gordon, and Stephen G. Evangelides, "Multigigabit Soliton Transmissions Traverse Ultralong Distances", *Laser Focus World*, pp. 159-170 (November 1991).

25. There are actually two variations of this process, applicable in both radio-frequency and lightwave systems. In one variation, homodyne detection, the local signal source is set for the same frequency as the incoming lightwave. The resulting electrical signal from the detector then represents the extracted information from the incoming lightwave directly. In the second variation, heterodyne detection, the local signal source is set for a frequency slightly different from the incoming signal. The resulting signal from the detector must be further processed to remove the information.

26. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 2-8 (1992).

27. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 8-1 (1992).

28. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 7-8 (1992).

29. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 7-1 (1992).

30. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 7-2 (1992).

31. Private communication from Market Intelligence Research Corporation, March 5, 1993.

32. *SONET Overview*, Bell Communications Research, "Broadband ISDN Evolution", p. 5 (1991). The information capacity levels shown are for applications in Broadband ISDN.

33. William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. 81 (New York, New York: Macmillan Publishing Company; 1992).

34. William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. 81 (New York, New York: Macmillan Publishing Company; 1992). Ronald J. Vetter and David H. C. Du, "Distributed Computing with High-Speed Optical Networks", *Computer*, p. 13 (February 1993).

35. William Stallings, *ISDN and Broadband ISDN, Second Edition*, pp. 581-595 (New York, New York: Macmillan Publishing Company; 1992). The information in Table 96 is a condensation of Table 8.1 from Stallings, page 585.

36. Key parts of the information in this section were developed with the help of Howard Davidson of NYNEX, the Bell operating company for the New York and New England regions.

37. William Stallings, *ISDN and Broadband ISDN, Second Edition*, pp. 34-39 (New York, New York: Macmillan Publishing Company; 1992).

38. Time-division multiplexing is sometimes used in a more restrictive sense to refer only to that subset of time-division multiplexing processes in which the position of data in time along a line has significance, as in TDM bus switching.

39. William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. 82 (New York, New York: Macmillan Publishing Company; 1992).

40. TCP/IP stands for Transmission Control Protocol/Internet Protocol.

41. Douglas Comer, *Internetworking with TCP/IP: Principles, Protocols, and Architectures*, p. 5 (Englewood Cliffs, New Jersey: Prentice-Hall, Incorporated; 1988).

42. William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. 381 (New York, New York: Macmillan Publishing Company; 1992).

43. Asynchronous transfer mode contains provision for encoding certain cells for retention "unless no other alternative is available." William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. 536 (New York, New York: Macmillan Publishing Company; 1992).

44. The total information capacity (in bits per second) achievable is actually dependent on several factors. At the most fundamental level, information capacity increases (1) linearly with the frequency bandwidth to be occupied by the signal, and (2) logarithmically with the signal-to-noise ratio in the transmission system (Hartley-Shannon Theorem). The information capacity actually achieved in a practical system depends also on the modulation technique employed. The percentage specified in the text reflects a total information capacity of 2000 gigabits per second. This level should eventually be reachable, with suitable advances in the technology of the components used in optical-fiber communications systems.

45. Donald E. Clark and Tetsuya Kanada, "Broadband: The Last Mile", *IEEE Communications Magazine*, p. 94 (March 1993).

46. William E. Burr, *Architectures for Future Multigigabit Lightwave Networks*, Report No. NISTIR 90-4240, National Institute of Standards and Technology, U.S. Department of Commerce, p. 1 (February 1990).

47. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 6-9 (1992). The work has been conducted by British Telecom in pursuit of cost-competitiveness with metallic systems.

48. William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. ix (New York, New York: Macmillan Publishing Company; 1992).

49. The information capacity of 144 kilobits per second is composed of two data channels of 64 kilobits per second capacity each and one control channel of 16 kilobits per second.

50. William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. 340 (New York, New York: Macmillan Publishing Company; 1992).

51. William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. 99 (New York, New York: Macmillan Publishing Company; 1992).

52. William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. 381 (New York, New York: Macmillan Publishing Company; 1992).

53. William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. ix (New York, New York: Macmillan Publishing Company; 1992).

54. Data channels of 140 megabits per second are commonly mentioned. Control channels of 64 kilobits per second are commonly mentioned.

55. William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. 520 (New York, New York: Macmillan Publishing Company; 1992).

56. William Stallings, *ISDN and Broadband ISDN, Second Edition*, p. 546 (New York, New York: Macmillan Publishing Company; 1992).

57. There is provision within the definition of the SONET frames to enable them to be switched directly, but at present there is more interest in using other switching technologies.

58. *SONET Overview*, Bell Communications Research, in the "Introduction", p. 12 (1991).

59. These specifications appear in the voluntary industry standard ANSI T1.106-1988 of the American National Standards Institute, as discussed in *SONET Overview*, Bellcore (Bell Communications Research), p. 15 (1991).

60. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 7-13 (1992).

61. William E. Burr, *FDDI Planning Guide*, National Institute of Standards and Technology, p. 27 (draft document of November 6, 1992).

62. William E. Burr, *FDDI Planning Guide*, National Institute of Standards and Technology, pp. 48-49 (draft document of November 6, 1992).

63. "Communications Markets Conclusions", presented at the OIDA Forum on Market Opportunities, September 15-17, 1993. The OIDA is the Optoelectronics Industry Development Association, a U.S. organization composed of manufacturers of optoelectronic products.

64. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, pp. 4-3, 5-4, 6-3, 7-3, and 8-3 (1992).

65. Private communication from Market Intelligence Research Corporation, March 5, 1993.

66. Michael Fahey, "Worldwide Sales of Lightwave Equipment to Remain Strong", *Lightwave*, pp. 1 and 21 (December 1992). The data were read from a graph, with a possible error of $0.1 billion or so, either up or down.

67. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, pp. 3-4 to 3-5 (1992).

68. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 6-8 (1992).

69. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 6-9 (1992).

70. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 3-29 (1992).

71. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 3-7 for percentages and pp. 4-10, 4-16, 4-23, 5-10, 5-17, 5-23, 6-11, 6-18, 6-25, 7-10, 7-16, 7-22, 8-9, 8-15, and 8-21 for dollars (1992).

72. Michael Fahey, "Worldwide Sales of Lightwave Equipment to Remain Strong", *Lightwave*, pp. 1 and 21 (December 1992). The data were read from a graph, with a possible error of $0.1 billion or so, either up or down.

73. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, pp. 4-10, 5-10, 6-11, 7-10, and 8-9 (1992).

74. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, pp. 4-16, 5-17, 6-18, 7-16, and 8-15 (1992).

75. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, pp. 4-23, 5-23, 6-25, 7-22, and 8-21 (1992).

76. Michael Fahey, "Worldwide Sales of Lightwave Equipment to Remain Strong", *Lightwave*, pp. 1 and 21 (December 1992). The data were read from a graph, with a possible error of $0.1 billion or so, either up or down.

77. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, pp. 4-10 and 6-11 (1992).

78. *1990 World's Submarine Telephone Cable Systems*, Report No. NTIA-CR-91-42, National Telecommunications and Information Administration, U.S. Department of Commerce, p. 35 (October 1991).

79. *1990 World's Submarine Telephone Cable Systems*, Report No. NTIA-CR-91-42, National Telecommunications and Information Administration, U.S. Department of Commerce, p. 44 (October 1991).

80. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 8-8 (1992).

81. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 8-10 (1992).

82. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 7-14 (1992).

83. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 6-16 to 6-17 (1992).

84. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 8-14 (1992).

85. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 5-18 (1992).

86. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 4-4 (1992).

87. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 4-4 (1992).

88. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 4-23 (1992).

89. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 4-5 (1992).

90. Michael Fahey, "Connector Prices Drop as Shipments Rise", *Lightwave*, p. 8 (November 1992). Data provided by Fleck International, Huntington Beach, California.

91. An inflation rate of 4 percent per year over six years adds 27 percent to the cost of a product with no change in the product. An inflation rate of 5 percent per year over six years adds 34 percent to the cost of a product with no change in the product.

92. If X equals the cost in current dollars in 1992, and if Y equals the cost in 1998, expressed in constant 1992 dollars, after cost reductions, then Y can be found by equating the costs in current dollars in 1998, that is: X(1-C)= Y(1+I). The "C" is the cumulative percent price *decrease*, as a decimal, in the price in current dollars from 1992 to 1998. The "I" is the cumulative percent *increase*, as a decimal, attributable to inflation over the six-year period, 1992 to 1998. Thus the factor Y/X is the factor by which the initial 1992 price must *decrease* in *real* terms over the six-year period.

93. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation (MIRC), pp. 3-4, 3-19, 3-23, 3-27, and 3-30 (1992). North America includes Canada and the U.S. (p. 3-19). The Pacific Rim includes Japan, Australia, New Zealand, South Korea, Taiwan, Hong Kong, Thailand, the People's Republic of China, Malaysia, the Philippines, and Singapore. The Rest of the World includes Latin America, the Middle East, Africa, and other regions or countries not elsewhere included (pp. 3-29 to 3-30).

Europe includes both Western and Eastern Europe. Western Europe includes Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Luxembourg, Norway, Portugal, Ireland, Spain, Sweden, Switzerland, Netherlands, and the United Kingdom. Eastern Europe includes Albania, Bulgaria, Commonwealth of Independent States, Czechoslovakia, Hungary, Malta, Poland, Rumania, and Yugoslavia. Private communication from MIRC, March 5 and 8, 1993.

94. Data from Information Gatekeepers Incorporated, Boston, Massachusetts, as published in "Industry Update", *Lightwave*, p. 3 (March 1993).

95. Data from Information Gatekeepers Incorporated, Boston, Massachusetts, as published in "Industry Update", *Lightwave*, p. 3 (March 1993).

96. *1990 World's Submarine Telephone Cable Systems,* Report No. NTIA-CR-91-42, National Telecommunications and Information Administration, U.S. Department of Commerce, pp. 33 and 55-56 (October 1991).

97. Michael Fahey, "Nynex Heads Plan for Longest Fiber Link", *Lightwave,* p. 11 (July 1992).

98. *ElectroniCast Photonics*, Issue Number 4, ElectroniCast Corporation, p. 2 (August 1992). Also reported by Stan Kolodziej, "Market Researcher Sees Surge in Fiber Components Exports", *Lightwave*, pp. 16-17 (October 1992). In this latter reference, the range of components included is not spelled out but appears to include fiber cable, transmitters, receivers, transceivers, amplifiers, connectors and so-called 'interconnect connectors (patch panels and other inside plant equipment that enable "straight through" network interconnects)' as a minimum. Thus the "interconnect connectors" are included in the ElectroniCast data but not in the MIRC data.

99. The superscripts in the table refer to the notes that follow: (a) *1993 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 29-12 (January 1992). (b) John M. Kessler, *Worldwide Markets for Fiberoptics*, presented at 15 Annual Newport Conference on Fiberoptics Markets by Kessler Marketing Intelligence Corporation, no page number (October 19-21, 1992). (c) Corporate Strategic Intelligence (Middlebush, NJ) as reported in "Fiberoptics Industry Report", *Laser Focus World*, p. 153 (July 1990).

100. Corporate Strategic Intelligence (Middlebush, NJ) as reported in "Fiberoptics Industry Report", *Laser Focus World*, p. 153 (July 1990).

101. *U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, pp. 29-2 to 29-4 (January 1993). U.S. shipments of network equipment are estimated at \$10.2 billion for 1992, with a favorable balance of trade of \$1.2 billion. U.S. shipments of customer premises equipment are estimated at \$4.8 billion for 1992, with an unfavorable balance of trade of \$3 billion. Thus the U.S. market for these products for 1992 is estimated at \$16.8 billion (10.2 - 1.2 + 4.8 + 3).

102. John Celentano and Mark Lutkowitz, "Two Suppliers Merit Watching in U.S. Sonet Market", *Lightwave*, pp. 45 and 47 (November 1992).

103. Jonathan M. Kraushaar, *Fiber Deployment Update: End of Year 1991*, Federal Communications Commission, Table 4, page 11, Table 9, page 24, and Table 15, page 36 (March 20, 1992). Page numbers are approximate because electronic format of document yields pagination dependent on receiving system.

104. Michael Shimazu, "The Ball Game is Bigger Than OEICs", *Photonics Spectra*, p. 75 (October 1992). The information provided comes from the book *The New Optoelectronics Ball Game: The Policy Struggle Between the U.S. and Japan for the Competitive Edge* by Philip Seidenberg of Georgia Tech, published by

IEEE [Institute of Electrical and Electronics Engineers] Press in 1992. Ernest Sternberg, "Should We Have a Policy for Optoelectronics?", *Laser Focus World*, p. 59 (September 1992).

105. *High-Technology Competitiveness: Trends in U.S. and Foreign Performance,* Report No. GAO/NSIAD-92-236, U.S. General Accounting Office, p. 57 (September 1992).

106. *OITDA Activity Report, Vol. 5, for Fiscal Year Ended March 31, 1992*, Optoelectronic Industry and Technology Development Association, Japan, p. 5 (1992). OITDA's category "Optical fibers (including cables)", subcategory "Fibers for telecommunications" appears to count both uncabled fibers and cabled fibers and thus may double count those uncabled fibers that are later put into cable. An exchange rate of 134.5 yen per dollar, the average value for calendar 1991, was used to convert Japanese yen to U.S. dollars.

107. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, p. 3-4 (1992).

108. *ElectroniCast Photonics*, Issue Number 4, ElectroniCast Corporation, p. 3 (August 1992).

109. World production has been taken as equal to the world market, which, according to MIRC, was 59.6 percent of the world market of $3.84 billion for all components in 1991. *World Fiber Optic Communication Application Markets*, Market Intelligence Research Corporation, pp. 3-4 and 3-7 (1992).

110. Private communication from Market Intelligence Research Corporation, March 5, 1993.

111. "Insulated Wire and Cable", *Current Industrial Reports*, Issue MA33L(91)-1, Bureau of the Census, U.S. Department of Commerce, p. 5 (September 1992). U.S. production of cabled fiber for communications applications was $709 million for 1991. U.S. production of uncabled fiber for both communications and other applications was $463 million for 1991; the fraction that is for communications is not reported separately by the Bureau of the Census.

112. *OITDA Activity Report, Vol. 5, for Fiscal Year Ended March 31, 1992*, Optoelectronic Industry and Technology Development Association (OITDA), Japan, p. 5 (1992). OITDA's category "Optical fibers (including cables)", subcategory "Fibers for telecommunications" appears to count both uncabled fibers and cabled fibers and thus may double count those uncabled fibers that are later put into cable. An exchange rate of 134.5 yen per dollar, the average value for calendar 1991, was used to convert Japanese yen to U.S. dollars.

113. "Insulated Wire and Cable", *Current Industrial Reports*, Issue MA33L(91)-1, Bureau of the Census, U.S. Department of Commerce, p. 6 (September 1992).

114. *1993 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 29-12 (January 1992).

115. *1993 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 29-14 (January 1992).

116. *1993 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 29-12 (January 1992).

117. *1993 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 29-13 (January 1992).

118. "Europeans Test Limits of Broadband WDM", *Lightwave*, pp. 1, 21, and 24 (February 1993).

119. George Lawton, "Russians Plan Fiber Network", *Lightwave*, p. 1 (October 1992). Also, "Russia and Korea Build Fiber Links with Japan", *Lightwave*, p. 14 (September 1992).

120. *1993 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 29-13 (January 1992).

121. "Japan Steps Up Investment In Broadband-ISDN", *Lightwave*, p. 6 (June 1992).

122. George Miller, "Japan Optimistic at Interopto", *Lightwave*, pp. 1, 16, and 18 (September 1992).

123. "Japan Steps Up Investment In Broadband-ISDN", *Lightwave*, p. 6 (June 1992).

124. *1993 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 29-13 (January 1992).

125. *National Measurement Requirements Committee Report*, National Conference of Standards Laboratories, p. 10 (June 1991).

126. *Technical Digest -- Symposium on Optical Fiber Measurements, 1992*, National Institute of Standards and Technology, U.S. Department of Commerce (September 1992). Sponsored by NIST in cooperation with the Lasers and Electro-Optics Society of the Institute of Electrical and Electronics Engineers and the Optical Society of America, on September 15-17, 1992.

127. Attenuation is greater for higher-order modes, that is, for those modes associated with paths that are highly divergent from axis of the fiber. These modes experience more internal reflections and they travel a path that is longer than that of lower-order modes.

128. In the dispersion-shifted fiber, lightwaves of longer wavelength penetrate deeper into the cladding of a fiber and thus become more sensitive to the refractive-index across a greater cross-section of the fiber.

129. C. F. Cottingham, "Dispersion-Shifted Fiber", *Lightwave*, p. 25 (November 1992).

130. The present NIST measurement reference standard, in the form of a so-called C-series calorimeter, supports accuracy levels of 0.4 percent down to 10 milliwatts of power.

131. Siegmar Schmidt, "High-Performance Optical Return Loss Measurement", *Hewlett-Packard Journal*, p. 79 (February 1993).

132. The cable television industry is gradually moving to digital technology for several reasons: to permit interconnections with other networks, such as metropolitan-area networks; to enable the industry to provide broadened services; and to enable using digital compression. "Cable-TV Industry Moves Toward Digital Systems", *Lightwave*, pp. 1 and 12 (February 1993). Denver-based Telecommunications, Incorporated, "the largest cable TV system operator in the United States", estimates that digital compression will enable the company to provide 500 channels to its subscribers by 1994. "Frontlines: Over 500 Channels", *Lightwave*, p. 3 (January 1993).

133. James J. Refi, "New Technologies Drive Optical Systems Faster and Closer to Home", *Laser Focus World*, p. 150 (December 1991).

134. Hermann A. Haus, "Molding Light Into Solitons", *IEEE Spectrum*, pp. 52-53 (March 1993).

135. "NTT Completes Soliton Experiment", *Lightwave*, p. 10 (November 1992).

136. C. F. Cottingham, "Dispersion-Shifted Fiber", *Lightwave,* p. 29 (November 1992).

137. Conventional detection systems generally employ amplitude-shift keying; that is, they impress the desired digital signal on the lightwave by modulating the amplitude of the lightwave. This technique can be used with coherent detection, too, but greater detector sensitivity can be obtained with frequency-shift keying and phase-shift keying as the modulation techniques. As their names imply, they modulate the frequency or the phase of the lightwave, respectively. The sensitivities obtainable with these three modulation techniques in coherent systems range from 9 to 36 photons per bit at a bit error rate of 10^{-9}. E. E. Bert Basch, "Coherent Fiberoptic Systems: the Next Generation?", *Laser Focus World,* p. 187 (April 1989).

138. Progress is being made toward tunable lasers. "Researchers at NTT report on a world record tuning range of 101 nm [nanometers] in the 1550 nm window based on a superstructure grating DFB laser." "OFC/IOOC 93: High Capacity Networks and Enabling Technologies", *Optics & Photonics News,* p. 48 (December 1992).

139. Jeff Hecht, "Laser Action in Fibers Promises a Revolution in Communications", *Laser Focus World*, pp. 78 and 80 (February 1993). The laser-diode pump wavelengths in Table 119 came from this reference. Optical-fiber amplifiers employing praseodymium in fluoride glasses require pump wavelengths from 1007 to 1017 nanometers. Optical-fiber amplifiers employing neodymium in either fluoride or phosphate glasses are not quite as desirable because they produce wavelengths somewhat longer than the 1310 nanometers required for minimum chromatic dispersion.

140. Pumping wavelengths for optical-fiber amplifiers are discussed by Jeff Hecht, "Laser Action in Fibers Promises a Revolution in Communications", *Laser Focus World*, pp. 75-81 (February 1993).

141. Ferrimagnetic materials used in isolators have a refractive index that is dependent on the applied magnetic field (Faraday effect). This dependency can cause a rotation of the electric field of an incoming lightwave traveling in the direction of the applied magnetic field. The rotation increases with the magnetization of the material, which in turn increases with the applied magnetic field. The rotation will increase linearly with magnetic field for small fields, and the slope of that increase is measured by the Verdet constant. The rotation will saturate at higher magnetic fields. On a plot of the rotation versus the magnetization, the point at which the rotation saturates marks both the saturation rotation and the saturation magnetization. For isolators, materials with a high saturation rotation, especially at small values of saturation magnetization, are highly desirable. Also, materials with stable values of saturation rotation in the presence of temperature and wavelength changes are also highly desirable.

142. Nick Lewis, Peter Keeble, and David Ferguson, "Testing Strategies for Modern Fibre Network Architectures", from British Telecom Laboratories, *Technical Digest--Symposium on Optical Fiber Measurements, 1992,* published by NIST, p. 3 (September 1992).

143. T. S. Frank Lee, "Reliability Testing of a Fiber Optic System for Subscriber Loop Application", from Bellcore, *Technical Digest--Symposium on Optical Fiber Measurements, 1990,* published by NIST, p. 179 (September 1990).

144. Nick Lewis, Peter Keeble, and David Ferguson, "Testing Strategies for Modern Fibre Network Architectures", from British Telecom Laboratories, *Technical Digest--Symposium on Optical Fiber Measurements, 1992,* published by NIST, p. 1 (September 1992).

145. Nick Lewis, Peter Keeble, and David Ferguson, "Testing Strategies for Modern Fibre Network Architectures", from British Telecom Laboratories, *Technical Digest--Symposium on Optical Fiber Measurements, 1992,* published by NIST, p. 3 (September 1992).

146. Nick Lewis, Peter Keeble, and David Ferguson, "Testing Strategies for Modern Fibre Network Architectures", from British Telecom Laboratories, *Technical Digest--Symposium on Optical Fiber Measurements, 1992*, published by NIST, p. 2 (September 1992).

147. Nick Lewis, Peter Keeble, and David Ferguson, "Testing Strategies for Modern Fibre Network Architectures", from British Telecom Laboratories, *Technical Digest--Symposium on Optical Fiber Measurements, 1992*, published by NIST, p. 4 (September 1992).

CHAPTER 10

OPTICAL-FIBER SENSORS

December 1992

Chapter 10

OPTICAL-FIBER SENSORS

SUMMARY

Sensors are measurement devices. Sensors translate quantities to be measured, such as temperature or pressure, into useful forms, usually electrical or optical quantities. These electrical or optical quantities are readily transmitted and processed to provide the information needed to support a diversity of modern systems. Sensors serve manufacturing, transportation, energy, defense, environmental protection, health, safety and many other areas of national concern. The capabilities of sensors are often the limiting factors in the performance of high-technology systems.

Optical-fiber sensors are a versatile class of sensors. Optical-fiber sensors translate the quantity to be measured into a change in one of four fundamental properties of a lightwave: intensity, polarization, phase, or wavelength composition. The element providing the sensing function may be built into the optical fiber itself or it may be located in a special optical device attached to the end of the fiber. The change made in the lightwave by the sensing element is transmitted through an optical fiber to processing components that interpret the change in terms of the original quantity of interest, providing the measured value.

Optical-fiber sensors bring to measurement systems many important advantages, some of which are shared with optical-fiber communications systems. Optical-fiber sensors are small, light in weight, and geometrically flexible. They are highly tolerant of adverse environments, including those with high levels of electromagnetic fields, high chemical reactivity, and high temperatures. Optical-fiber sensors are electrically non-conducting and thus provide electrical isolation between measured systems and other equipment. Optical-fiber sensors are compatible with a broad range of multiplexing and networking technologies similar to those employed in modern optical-fiber communications systems. These technologies enable interconnecting an enormous number of optical-fiber sensors. Some optical-fiber sensors are even capable of distributed sensing; that is, they can make measurements all along their length for some quantities. These multiplexing, networking, and distributed-sensing capabilities of optical-fiber sensors enable them to provide the equivalent of "nervous systems" for aircraft, ships, factories, and other large systems. Further, optical-fiber sensors are compatible with emerging optical-signal-processing technology and thus are ready for the transition to an all-optical technology, should that prove possible in the future.

Optical-fiber sensors are a small but steadily growing part of the expanding market for all types of sensors. The world market for all sensors was about $7 billion in 1991. The world market for optical-fiber sensors was about $164 million in 1990, or roughly 2 percent of the world market for all types of sensors. This market consumed 305,000 optical-fiber sensors in 1990. Between 1990 and 1996, the growth rate in the world-market for optical-fiber sensors is expected to exceed 30 percent per year, leading to a world market greater than $900 million by 1996.

Optical-fiber sensors are serving in a variety of end-user industries. These industries include the following: industrial machinery, metalworking, processing, in-plant process control, medicine,

aerospace, and electrical power. These industries are listed in descending order of dollar consumption of optical-fiber sensors in the world market in 1990. Sensors classified in the *processing* end-user industry are those built into processing equipment; they serve applications in such large industries as petroleum refining, chemical production, and plastics formulation and molding. Sensors classified in the *in-plant process control* industry are those sold separately but are for use with processing equipment; they serve a greater diversity of industries.

Optical-fiber sensors have already been developed for many measured quantities. They show great potential for an even greater variety of measured quantities. Optical-fiber sensors that are already in service may be grouped by measured quantity as follows: displacement and proximity; pressure; temperature; level; chemical and biochemical; magnetic and electrical; flow; and rotation. This order reflects descending dollar sales in the world market in 1990. During 1990, the displacement and proximity sensors were especially important; they represented 55 percent of the dollar value and 93 percent of the unit value of all optical-fiber sensors sold worldwide. Displacement and proximity sensors are expected to remain dominant in 1996 in unit sales and to place second in dollar sales. Rotation sensors (especially optical-fiber gyroscopes) and magnetic and electrical sensors are expected to show the greatest increase in market share, based on dollar sales by 1996, with rotation sensors placing first in dollar sales. The two types of sensors that were used most uniformly across all of the industries served by optical-fiber sensors, based on dollar sales in 1990, were temperature and pressure.

While optical-fiber sensors have tremendous potential, a number of challenges must be addressed to realize that potential and to assure U.S. competitiveness in providing the sensors to the world market. Optical-fiber sensors must be improved in performance, especially in accuracy and selectivity. Selectivity means responding *only* to the quantity to be measured; present sensors often respond to other quantities also, such as strain and temperature, when they are not the quantities to be measured. Optical-fiber sensors must also be improved in dynamic range, particularly at the low end (sensitivity). Their reliability must be demonstrated if they are to gain acceptance in the marketplace. Their cost must be reduced through development of automated manufacturing processes; most optical-fiber sensors are still made by manual methods. They must be packaged in modular form with standardized interfaces to ease use in diverse systems. Finally, improved components for signal handling must be developed. These components must support networking and multiplexing services in many applications. Eventually, the development of optical signal processing will enable even fuller exploitation of the capabilities of optical-fiber sensors.

Addressing these challenges will require major advances in measurement capability for virtually all of the components involved in an optical-fiber-sensor measurement system, whether those components perform illuminating, sensing, transmitting, or signal-processing functions. In all, twenty-one measured quantities require improved measurement support in one form or another. Nineteen of these quantities require new or improved *measurement methods*. For example, a common need is for measurement methods that can determine changes in important optical quantities, such as attenuation and birefringence, in the presence of external mechanical stresses. These stresses include bending, twisting, and axial and transverse stretching and compression; they are broadly experienced by optical-fiber sensors in service. Such measurement capability supports improved selectivity. Another common need is for measurement methods with higher accuracy and increased dynamic range to support the design of sensors with improvements in the same two factors. Two measured quantities -- refractive index and birefringence phase difference -- require *measurement reference standards* to assure the accuracy of measurement methods. Six measured quantities require development of *materials reference data*. Important examples are data that reflect the responses of optical materials to electric and magnetic fields. Also important are data that show the variation of optical properties, such as refractive index, with temperature. These materials reference data are necessary to support the design of new sensors and the development of automated manufacturing processes that can bring down the cost of optical-fiber sensors.

ELECTRICAL AND OPTICAL SENSORS

Sensors are intimately tied to measurements. The purpose of most modern sensors is to translate some quantity to be measured, such as pressure, into some other quantity, usually an electrical or an optical quantity, in a consistent manner.[1] The purpose of the translation is to enable the use of modern electronic and optical instrumentation to make the measurement for any one of several reasons, such as accuracy, speed, cost, sensitivity, or compatibility with a particular application. Here are three examples:

> Temperature: Temperature can be measured by a mercury thermometer. This approach is suitable for some applications. Or temperature can be measured through an electrical translation by determining the electrical resistance of a piece of wire whose resistance increases with temperature in a known manner. This approach enables the use of the temperature information in electronic systems that control or record temperature, whereas the output of a mercury thermometer is difficult to use in a such a way.
>
> Weight: A balance-beam scale can determine weight by purely mechanical means. The user adjusts the balance weights and interprets their positions to make a measurement. Alternatively, an electronic sensor may be used. This sensor is usually a material whose electrical resistance changes with the strain created by the applied weight. These changes are measured and translated into weight. This approach is often less expensive and easier to use.
>
> Plant vigor: Plant vigor, as reflected in such factors as drought effects, nutrient deficiencies, and fungal infestations, has generally been measured by a variety of tests made either on site or in laboratories. Alternatively, airborne optical sensors may be used. The plants are illuminated by a laser beam from an aircraft, causing the plants to fluoresce (emit light). Analysis of the fluorescent light returned to the aircraft enables identification of the plants and assessment of their vigor. This approach is much faster and less costly. This application and others employing optical sensing with lasers are described in Chapter 8, Lasers, on page 200.

Electrical Sensors

The bulk of modern sensors, like those just described as examples, translate the quantity to be measured into an electrical quantity. Electro-mechanical sensors have been the standard sensor technology for decades. They have taken many forms and have provided reliable, robust, and low-cost solutions to sensing.[2] But now new sensor types are emerging. Preeminent among these are silicon sensors. Silicon-based sensors have the advantage that they are readily fabricated with processes similar to those used to make silicon integrated circuits. Silicon-based sensors are also readily built into integrated circuits which provide signal-processing electronics. This close integration produces so-called *smart sensors* which are close to being complete measurement instruments.[3] Silicon sensors can be fragile in many operating environments, but steady progress has been made in shielding them from damage.

Optical Sensors

A second important class of sensors translate the quantity to be measured into an optical signal. Some of these employ lasers operating in a free-space environment, as described in one of the examples above. But an emerging type of sensor employs light in a guided environment. This emerging type of sensor is the optical-fiber sensor.

In optical-fiber sensors, lasers or light-emitting diodes serve as light sources. Optical fibers, or optical devices connected to the fibers, perform the sensing function by translating the quantity to be measured into a change in a lightwave. Optical fibers provide the transmission media for transferring the changed lightwave to processing components. The processing components may be located remotely, that is, they may be located quite some distance from the site to be sensed. Processing components translate the change in the lightwave into a useful representation of the measured quantity. This representation is commonly in electrical form. That form may be an analog signal level or a digitally encoded number. Thus optical-fiber sensors are really measurement systems composed of a variety of components that perform a series of four functions: illuminating, sensing, transmitting, and processing.

ADVANTAGES AND CHALLENGES FOR OPTICAL-FIBER SENSORS

Optical-fiber sensors present both advantages and challenges. They are itemized in Table 123. Some of the advantages are *inherent*, that is, they reflect the basic properties of optical fibers. Other advantages are *potential* in nature. They have been realized in optical-fiber sensors for some measured quantities but they pose *important* challenges for realization in optical-fiber sensors for a broader range of quantities. Three challenges -- specifically high accuracy, high selectivity, and low initial cost -- have been particularly difficult to address across the broad spectrum of optical-fiber sensors and are very important to their commercial success; they have been checked as *critical* challenges in the table.

The inherent advantages of optical-fiber sensors are discussed below. The potential advantages and the associated challenges are discussed beginning on page 319, where the accompanying measurement needs are identified.

Physical Advantages

Most optical-fiber sensors are made of thin strands of amorphous glass. These fibers are very similar to those used for optical-fiber communications. They were described in Chapter 9, Optical-Fiber Communications, beginning on page 217. The fibers used in sensors have the same ability as the fibers used in communications to trap light and to guide it along their length with very low losses. The fibers have three inherent physical advantages that are very important to their versatility as sensors. They are small in size, low in weight, and geometrically flexible. While glass is the dominant material for the fibers used in optical-fiber sensors, other materials are increasingly being evaluated. These include plastics, silicone rubber,[4] zirconium fluoride, sapphire, and chalcogenides.[5] While these materials differ greatly from each other, they share the same three desirable physical

Table 123
Advantages and Challenges for Optical-Fiber Sensors

	Advantages		Challenges	
Characteristic	inherent	potential	important	critical
physical				
small	✓			
low weight	✓			
geometric flexibility	✓			
environmental				
safe	✓			
electrically isolating	✓			
chemically non-toxic				
non-perturbing of measured environment	✓			
tolerant of adverse environments				
electromagnetic radiation	✓			
chemically reactive	✓			
high temperatures	✓			
performance				
high information capacity	✓			
accuracy				✓
selectivity				✓
dynamic range/sensitivity		✓	✓	
reliability				
in service		✓	✓	
on the shelf		✓	✓	
cost				
initial				✓
maintenance		✓	✓	
modularity and standardization		✓	✓	
signal-handling				
networking and multiplexing		✓	✓	
distributed sensing	✓		✓	
optical signal processing		✓	✓	

characteristics with glass fibers.

Environmental Advantages

There are a number of inherent advantages of optical-fibers sensors that may be called environmental in nature because they relate so closely to the environments in which optical-fiber sensors can be used. Several of these are traceable to the non-conducting, or electrically insulating nature, of optical-fiber sensors. This characteristic promotes safety by electrically isolating measuring equipment and its operators from the system being measured; that is, the optical-fiber sensors do not transfer dangerous voltages or currents. The non-conducting nature also makes the sensors non-perturbing to the measured environment and tolerant of environments containing electromagnetic radiation or high electromagnetic fields. The importance of tolerance for electromagnetic interference is discussed in Chapter 12, Electromagnetic Compatibility, beginning on page 381.

Another inherent advantage is traceable to the low chemical reactivity of the sensors, which makes them relatively tolerant of chemically reactive environments. A still further advantage arises from the tolerance of optical-fiber sensors for high temperatures, which makes them useful in high-temperature environments, whether they are measuring temperature or some other quantity.

In special applications, optical-fiber sensors promote safety. Their non-electrical nature is important to measurement applications in process industries employing or making flammable or explosive materials. Their non-toxic nature is important to medical applications. Further, they can replace highly stressed and potentially explosive transformers used to measure electrical current in electric power systems.

Performance Advantage

Optical-fiber sensors have the inherent advantage of high information capacity, just like the optical fibers used for communications systems. This information capacity is, at a minimum, thousands of times greater than that of electronic sensors and potentially much greater. This capacity is so high that no measurement system is likely to exploit it fully. Such capacity is a major factor in enabling optical-fiber sensors to serve as nervous systems for complex and huge systems such as modern ships and factories.

Signal-Handling Advantage

Optical-fiber sensors have an inherent advantage for distributed sensing, that is, for making measurements all along their length. For example, temperature[6] or strain measurements can be made on a distributed basis. This capability is important for potential applications in "smart skins" or "smart materials". For example, aircraft surfaces of the future may contain embedded fiber sensors that report continuously on the strain throughout the surfaces, if the technical and economic challenges can be successfully addressed.[7] Similarly, structural materials may contain embedded fiber sensors that report on corrosion or vibration on a distributed basis.[8] Thermal processing facilities may contain fiber sensors that report on the temperature throughout the facilities.

However, the ability to use optical-fiber sensors to make distributed measurements is highly dependent on sophisticated processing components which separate the measurement information based on its location along the fiber. Most commonly used is a technique called time-domain reflectometry, which interrogates each portion of the fiber separately by noting the difference in arrival time for signals returned from nearer versus farther locations along the fiber. This technique thus employs a sort of optical echo which returns with information about the quantity to be measured. The development of measurement systems to support the distributed measurement of a broad spectrum of quantities is an important challenge.

More generally, optical-fiber sensors have great potential for use with networking and multiplexing techniques, including time-division and wavelength-division multiplexing, that can take advantage of their high information capacity. These techniques are discussed in detail in Chapter 9, Optical-Fiber Communications, beginning on page 217. They enable interconnecting many more individual sensors than would be possible in systems using electronic sensors.

QUANTITIES MEASURED BY OPTICAL-FIBER SENSORS

The quantities currently measured by commercial optical-fiber sensors are described in Table 124. The eight principal classes of quantities are shown in the left column. The right column shows examples of quantities in the classes or provides other information of interest.

Table 124
Quantities Measured by Optical Fibers

Measured Quantities	Notes
displacement/proximity	strain, vibration, sound
temperature	in thermal environment, or remote
pressure	stress, force
chemical/biochemical	composition, acidity (Ph)
flow	in liquids and gases
magnetic/electrical	electric field, magnetic field (thus voltage, current, energy, power)
rotation	angle, angular velocity (especially gyroscopes), angular acceleration
level	in liquids

The potential for expansion of the number of quantities measured by optical-fiber sensors is remarkable. At present there seem to be few limits to the discovery of new mechanisms for measuring new quantities. There also seem to be few limits to the discovery of improved mechanisms for quantities currently measured by existing optical-fiber sensors.

SENSING ELEMENTS AND MECHANISMS FOR OPTICAL-FIBER SENSORS

Sensing Elements

The sensing element is the physical part of the optical-fiber sensor that performs the sensing function. The sensing element may be *intrinsic* or *extrinsic* in nature.

Intrinsic

When the optical fiber itself performs the sensing function, the sensing element is said to be intrinsic in nature. More specifically, intrinsic sensing elements modify the optical properties of the fiber itself.[9] Intrinsic sensing elements may be realized in several different ways: (1) as a natural part of the core of the fiber; (2) by a dopant embedded within the fiber; (3) as a part of the cladding, or outer shell, which is an integral part of every optical fiber[10]; or (4) by a coating that runs along the length of the sensing region of the fiber and acts directly on the optical properties of the fiber.[11]

Extrinsic

When the sensing function is performed external to the optical fiber, then the sensing element is said to be extrinsic in nature. The extrinsic sensing element can take many forms, such as a prism for sensing concentration in liquids, a microbellows for sensing pressure, or a fluorescent coating on the end of the fiber for sensing temperature. When an extrinsic sensing element is employed, then the optical fiber serves only as the transmission medium, although it may require special optical properties. The extrinsic sensing element may be applied directly to the end of the fiber (like the fluorescent coating), or it may be placed against the end of the fiber or near the end of the fiber. The use of an extrinsic sensing element is generally used when no intrinsic element has been found or when known intrinsic elements perform inadequately; they may, for example, have inadequate sensitivity.

Access to Sensing Elements

Whether an intrinsic or an extrinsic sensing element is used, access to the sensing element may be made in any of several ways:

> (1) Two fibers may be used with an active light source. The light source launches a lightwave into one fiber which carries it to the sensing element. The sensing element modifies the lightwave and returns it through the second fiber to the processing components.[12]
>
> (2) One fiber may be used with an active light source. The light source launches a lightwave into the fiber which carries it to the sensing element. The sensing element modified the lightwave and returns in through the same fiber to the processing components.
>
> (3) One fiber may be used without an active light source. The sensing element itself generates the lightwave that is then carried down the fiber to the processing components.

In any of these cases, the sensing element impresses on the properties of the lightwave some change that can be interpreted by the processing components in terms of the quantity to be measured.

Sensing Mechanisms

The nature of the change impressed on a lightwave by a sensing element is known as the *sensing mechanism*. Optical-fiber sensors exploit different mechanisms to measure different quantities. The mechanisms used are shown in Table 125.

Intensity mechanisms

Intensity mechanisms are useful for measuring a wide variety of quantities, including mechanical, chemical, and thermal quantities. For example, pressure can be measured because transverse pressure applied along a fiber increases optical loss. Similarly, bending can be measured because bending a fiber to a small radius increases its loss. Temperature can be measured by incorporating certain materials, especially rare earth elements, into a fiber to make its loss temperature sensitive.

Table 125
Mechanisms of Sensing in Optical-Fiber Sensors

Mechanism	Description
intensity	change in optical power of a lightwave
wavelength	change in color of the lightwave
phase	shift of lightwave in time or space
polarization	change in direction of transverse electric field

Temperature can also be measured by determining the relative intensity of light emitted by a fluorescent coating on the end of a fiber. Fluorescence is the emission of light from a material in response to illumination by other light. The presence or concentration of liquids or gases can be measured by making the outer parts of a fiber slightly porous to induce a loss dependency. Alternatively, the concentration of liquids can be measured by attaching a prism to the end of a fiber, dipping the prism into a liquid, and observing the intensity changes in the reflections produced by differences in the refractive indices[13] of the prism and the liquid.[14] Both temperature and pressure can be measured through their effects on losses at interconnections between two fibers or between fibers and other components. The presence of particulate matter in a medium can be measured by observing the loss of light intensity resulting from the scattering caused by the medium when it is placed between the transmitting and receiving fibers of an optical-fiber sensor.[15] Intensity sensors are very versatile. They are also among the least complex and least costly of sensor types.

Wavelength mechanisms

Wavelength can be used as a sensing mechanism. For example, wavelength can be used directly for measuring temperature; that is, temperature can be determined by measuring the temperature-dependent wavelength composition of the radiation produced by any warm body. The radiation may be generated by the fiber itself or by another component used in conjunction with the fiber. Temperature may also be determined by measuring the temperature-dependent Raman effect of a material. The Raman effect is the emission of light of a slightly different wavelength in response to illumination by light of an initial wavelength. Further, gaseous chemical species may be determined by measuring changes in the wavelength composition of light after absorption of certain wavelengths by the chemical species.

Phase mechanisms

Lightwaves, like all electromagnetic waves, are composed of electric and magnetic fields that vary periodically in strength as they pass a given point. They also vary periodically with distance along their path of travel at a given moment in time. Shifts of corresponding parts of these waves in either time or space reflect changes in the phase of the lightwaves.

Changes in phase are useful for measuring displacement, strain, and rotation. Phase can be also used to measure other quantities that induce strain, including temperature, pressure, and even electric and magnetic fields. In a phase sensor, the sensing element is part of an interferometer. An interferometer is sensitive to small changes in the phase of light, which may be caused by changes in the length of a fiber or in the position, or optical properties, of an object in the optical path of the fiber. Phase sensors are among the most sensitive of optical-fiber sensors. In fact, a strain as small as 1 part in 10^{14} can be detected by a fiber only 1 meter in length.[16] Phase differences, and the interference effects that they produce, are used in optical-fiber gyroscopes where the relative phases of two lightwaves rotating in opposite directions in rings made of optical fiber are compared to sense rotation. The first optical-fiber sensors for magnetic field used a phase mechanism to sense the strain in a fiber coated with a magnetostrictive material; this is a material that undergoes physical changes in dimensions when exposed to a magnetic field.

Polarization mechanisms

The polarization of a lightwave is indicated by the position of the electric field component of the lightwave. That component is always perpendicular to the direction of travel of the lightwave but may assume various positions about that direction. Polarization is an important sensing mechanism for electric and magnetic quantities. When electric fields or magnetic fields are applied to sensors using polarization mechanisms, the nature of the polarization of a lightwave changes with the applied field.

For sensing *magnetic fields*, the change in polarization is caused by the Faraday effect. The Faraday effect is observed when a magnetic field is applied parallel to the fiber and polarized light is launched into the fiber. The polarization of the light is rotated in proportion to the strength of the applied magnetic field and to the distance that the light travels in the fiber. A constant of proportionality describes the sensitivity of the fiber to this change and is called the Verdet constant. Light with a simple form of polarization (linear) that is launched into one end of the fiber will retain that simple form, but the direction of the polarization will be rotated in proportion to the strength of the magnetic field applied to the fiber.

For sensing *electric fields*, the change in polarization is caused by the Pockels effect. The Pockels effect is observed in certain crystalline materials used in conjunction with fibers to form fiber sensors. The Pockels effect is not observed in fibers themselves; they are not crystalline. The Pockels effect is rather complicated. It can change the nature of polarized light from a simple form (linear) to a more complex form (circular) and back again. Each time the simple form occurs, the polarization of the light appears at a right angle with respect to its previous appearance. The distance over which this change takes place is a function of the strength of the electric field applied to the crystalline material and thus provides a measure of the field. The strength of the Pockels effect in a given material is measured by electro-optic coefficients.

SENSOR MARKET AND U.S. COMPETITIVENESS

Sensors of all types are an enabling measurement technology that support a U.S. market of $50 billion to $75 billion per year of automated controls, according to the National Research Council.[17] Sensor performance is often the limiting factor in the performance of the systems that the sensors support.

The world market for sensors of all types was approximately $7 billion for 1991. The U.S. market for sensors of all types was approximately $3 billion for 1991.[18]

World Optical-Fiber Sensor Market by Consuming Region

The breakdown of the world market for optical-fiber sensors by consuming country or region is shown in Table 126, according to a study by Market Intelligence Research Corporation published in 1991.[19] This study is the source of most of the market data presented in this chapter. The world market for optical-fiber sensors was $164 million for 1990, or about 2 percent of the world market for all types of sensors. The U.S. market for optical-fiber sensors for 1990 was $42 million, representing about 25 percent of the world market for optical-fiber sensors in that year. By 1996, the U.S. market is expected to rise to $345 million, representing about 37 percent of the world market in that year.

Table 126
World Market for Optical-Fiber Sensors by Consuming Country or Region
($millions)

Country or Region	1990	1996
United States	42	345
United Kingdom & Canada	24	123
Western Europe	41	211
Pacific Rim	50	209
Rest of World	7	47
	164	935

The Pacific-Rim market of $50 million in 1990 is dominated by Japan. The Western-European market of $41 million for 1990 nearly matches the U.S. market for that year. A different market study found the European market to be $38 million in 1989, roughly in agreement.[20]

The world market for optical-fiber sensors is projected to increase to $935 million in 1996. From 1986 to 1990, the world market for optical-fiber sensors grew at an compound average rate of 16 percent per year. However, from 1990 to 1996 the world market is expected to grow at a compound average rate of 34 percent per year.[21] Estimates from other sources for the growth rate of the world market are also quite high: 21 percent per year from 1988 to 1998[22] and 40 percent per year from 1988 to 1992.[23]

World Optical-Fiber Sensor Market by Measured Quantity

The world market for optical-fiber sensors for 1990 and the projected world market for 1996 are shown in Table 127, listed by measured quantity and in decreasing order of market share on a dollar basis.[24] Note that displacement/proximity sensors and temperature sensors represent the largest shares of the market for 1990 and are expected to remain large in 1996. However, for 1996, rotation sensors, principally optical-fiber gyroscopes, are expected to become the largest segment, assuming their successful and timely implementation. Also, magnetic/electrical sensors are expected to increase in market share.[25] Sensors for other quantities are expected to decrease in market share on a dollar basis.

Table 127
World Market (Dollars) for Optical-Fiber Sensors by Measured Quantity

Measured Quantity	1990 World Market		Measured Quantity	1996 World Market	
	$millions	percent		$millions	percent
displacement/proximity	91	55	*rotation*	420	45
temperature	21	13	displacement/proximity	228	24
pressure	16	10	temperature	69	7
chemical/biochemical	14	9	*magnetic/electrical*	65	7
flow	14	9	pressure	51	6
magnetic/electrical	4	2	chemical/biochemical	49	5
rotation	2	1	flow	46	5
level	2	1	level	7	1
	164	100		935	100

The world market for optical-fiber sensors on a units basis was 305 thousand in 1990 and is expected to increase to 916 thousand by 1996. The breakdown of the market by measured quantity is shown in Table 128 in descending order of units.[26] Displacement/proximity sensors clearly dominate in both 1990 and 1996. Rotation sensors are a distant second even in 1996 and thus achieve their top position in Table 127 by virtue of their high expected unit cost of about $6000 each in 1996. Rotation and magnetic/electrical sensors are expected to increase in share of the world market on a units basis from 1990 to 1996. Sensors for all other quantities are expected to decrease in world-market share on a units basis.

Table 128
World Market (Units) for Optical-Fiber Sensors by Measured Quantity

Measured Quantity	1990 World Market		Measured Quantity	1996 World Market	
	thousands	percent		thousands	percent
displacement/proximity	282	93	displacement/proximity	754	82
pressure	12	4	*rotation*	72	8
temperature	5	2	pressure	40	4
level	3	1	*magnetic/electrical*	20	2
chemical/biochemical	1	-	temperature	15	2
magnetic/electrical	1	-	level	7	1
flow	1	-	chemical/biochemical	5	1
rotation	-	-	flow	2	-
	305	100		*916	100

Dash (-) means less than 500 units or 0.5 percent.

* Total does not match due to effects of rounding.

World Optical-Fiber Sensor Market by End-User Industry

The world market for optical-fiber sensors for 1990 and the projected world market for 1996 are shown again in Table 129, but this time broken down by end-user industry served.[27] The order is decreasing market share; however, the category representing *other* industries has been separated from the ranking and kept at the bottom of the table. Between 1990 and 1996 the sensors consumed in the aerospace and electrical power industries in Table 129 are expected to increase in market share. The increase in sensors for the aerospace industry reflects the strong growth anticipated for rotation sensors in Table 127. The increase in sensors for the electrical power industry reflects the growth anticipated for magnetic/electrical sensors in Table 127.

Table 129
World Market for Optical-Fiber Sensors by End-User Industry

End-User Industry	1990 World Market		End-User Industry	1996 World Market	
	$millions	percent		$millions	percent
industrial machinery	39	24	*aerospace*	457	49
metalworking	25	15	industrial machinery	99	11
process	25	15	process	64	7
in-plant process control	20	12	metal working	63	7
medical	18	11	*electric power*	63	7
aerospace	16	10	in-plant process control	60	6
electric power	10	6	medical	59	6
other	11	7	other	71	8
	164	100		*935	*100

* Total does not match due to effects of rounding.

In Table 129, sensors sold for *in-plant process control* accommodate diverse applications in which the sensor application is unique to a given processing application but for which the sensor is sold independently of the processing equipment. Applications like these arise in research and development toward new products, in the development of industrial processes, or in the installation or monitoring of process equipment.[28] In contrast, sensors sold for the *process industry* focus principally on sensors which are an integral part of processing equipment. The three principal process industries that are candidates for significant use of these optical-fiber sensors are "petroleum refining, chemical production, and plastics formulation and molding."[29]

Relative Concentration of Sensor Use in End-User Industries

The relative concentration of sensors for a given measured quantity in each of the end-user industries during 1990 is shown in Table 130.[30] Each column in the table should be viewed as a separate entity, totalling to 100 percent (within rounding error). Each block (■) in each column indicates that about 10 percent of all of the sensors for the quantity associated with that column, on a dollar basis, were used in industry shown to the left. Each small circle (○) indicates that less than 5 percent of all sensors for that quantity on a dollar basis were used in the industry to the left.[31]

Table 130 shows that rotational sensors are highly concentrated in the aerospace industry. Magnetic/electrical sensors are highly concentrated in the electric power industry. Chemical/biochemical sensors are highly concentrated in the medical industry. Level sensors are somewhat concentrated in the process industry. Flow sensors are somewhat concentrated in the in-plant process control industry. Displacement/proximity sensors are only modestly concentrated in the industrial machinery industry. Temperature sensors, followed very closely by pressure sensors, are the most uniformly spread among the end-user industries. The table also shows that five of the eight types of sensors were used in every, or nearly every, end-user industry in 1990: displacement/proximity, pressure, temperature, level, and flow.

Table 130
Relative Concentration of Sensors For A Given Measured Quantity in Each End-User Industry Based on Dollar Percentage of Worldwide Consumption For 1990

End-User Industry	Measured Quantity							
	displacement/ proximity	pressure	temperature	level	chemical/ biochemical	magnetic/ electrical	flow	rotation
industrial machinery	■■■	■■	■■	■			■■	
metalworking	■■	■	■	○			■	
process	■■	■■	■	■■■■	■		■	
in-plant process control	■	■	■■	■■	■■		■■■■	
medical	○	■■	■		■■■■■■		■	
aerospace	■	■	■■	■■	■	■	○	■■■■■■■■■■
electric power	■	○	■	■		■■■■■■■	○	
other	■	■	■	■	■	■■	■	○

■ about 10 percent of dollar value of sensors in column ○ less than 5 percent of dollar value of sensors in column

U.S. Competitiveness

Information on U.S. competitiveness in optical-fiber sensors is sparse. For 1990, Japan is believed to be the single largest supplier of optical-fiber sensors on a dollar basis.[32] Japan estimates that it will produce $545 million of "optical sensors" in 1991, but the part representing optical-fiber sensors is not known.[33] More than 1000 U.S. companies manufacture sensors of one type or another, and about 125 have optical-fiber sensors on the market.[34] Comparable information on the number of companies in Japan and Europe could not be located.

One source suggested in 1991 that the emergence of a relatively large market for optical-fiber gyroscopes would favor the U.S. as the largest supplier by 1996.[35] However, as of 1992 Japan was already producing and installing low-cost optical-fiber gyroscopes in high-end automobiles.

In terms of research and development, the U.S., Europe, and Japan all have rapidly advancing efforts in "basic research and commercial development" for optical-fiber sensors, although the largest research-and-development effort has been in the U.S., driven by military requirements.[36] In 1991, the U.S. published 38 percent of all technical papers recorded in the international INSPEC Database, maintained by the Institute of Electrical Engineers in the United Kingdom. Western Europe published 33 percent, and Japan published 8 percent.

For optoelectronic products more broadly, the U.S., Japan, and Europe are all strong in research, development, and manufacturing. However, Japan is leading in both technology and commercialization, and that gap is widening.[37]

APPLICATIONS EXAMPLES BY END-USER INDUSTRY

Both current and emerging applications of optical-fiber sensors are important to the various end-user industries. Current applications are already highly diverse. Emerging applications are even more so. What follows are examples of both current and emerging applications, organized by the end-user industry served. All sensors referred to in this section are optical-fiber sensors unless an explicit comparison with other types of sensors is suggested.

Aerospace/Military

The first rotation sensors, in the form of optical-fiber gyroscopes, have been successfully made. The appeal of optical-fiber gyroscopes is based on their small size, lack of moving parts, minimum maintenance requirements, zero warm-up time, and unlimited shelf life. They will be used in aircraft, missiles, ships, and possibly other vehicles. While their performance has been demonstrated, their price remains very high. The average unit price in 1990 was about $17,000. A unit price of $5000 to $10,000 will be needed to make them competitive with conventional mechanical gyroscopes, and the hope is to reach unit prices of $1000 to $2000 with sufficient production volume. To achieve these goals, improved production processes will have to be developed.[38] DOD funding has been important to the development of the gyroscopes. Among current DOD-funded projects is the development of next-generation navigation systems that integrate the gyroscopes with the Global Positioning System.[39] The Global Positioning System is described in Chapter 7, Microwaves, on page 158.

Electromagnetic sensors and displacement sensors are key elements in emerging applications in "smart skins" for aircraft. They can sense electromagnetic events, such as radar scans. They can also sense mechanical events such as flexing under stress, possibly in time for corrective action to be taken.[40] Magnetic sensors have been developed for detecting buried ordinance and ship signatures.[41]

Displacement sensors in the form of underwater microphones (hydrophones) for submarine detection have already been employed by DOD in Project Ariadne.[42] DOE is using a variety of optical-fiber sensors to support nuclear weapons testing and nuclear effects simulation.

Level sensors are contemplated for fuel-detection systems in a variety of vehicles. Level sensors also look promising for controlling damage in ships by detecting the rising water levels caused by collisions or torpedoes.[43]

For ships more generally, optical-fiber sensors are contemplated for a wide variety of quantities. They will likely be merged into optical-fiber communications-and-control systems within the ships. The immunity that they offer to electromagnetic interference is especially important in the presence of the powerful radar and communications systems used on the ships. Large numbers of sensors will be required. A recent Navy ship contains 2,500 sensors and next-generation ships may contain 7,500.

Optical-fiber sensors employed in military applications will have to be rugged. They must work over a wide temperature range (-55 to +125 °C) and under extreme mechanical conditions: high shock levels (several thousand g), high vibration levels (above 15 g), and high acceleration levels (above 100 g). The letter *g* represents the acceleration due to gravity at the surface of the earth, about 9.8 meters per second per second.[44]

Industrial Machinery

Industrial machinery includes "all heavy and lightweight machinery used in manufacturing products worldwide." It specifically includes equipment used for food processing, packaging, pulp and paper manufacturing, air-conditioning, heating, and ventilating.[45]

Both temperature and pressure sensors are promising areas for industrial machinery. One prospective application is engine-oil pressure monitoring.[46] In robotics, both temperature and pressure sensors are used to provide the elements of "touch". Temperature, pressure, and refractive-index sensors have already been developed for the food industry.[47]

Displacement sensors can be used for counting and for determining "thickness, eccentricity, concentricity, diameter, vibration, axial motion, proximity, strain, acceleration, and alignment."[48] Displacement sensors offer high accuracy and small size and weight for automated and robotic production and process equipment. These characteristics are particularly helpful "in designing faster, smaller, and more accurate machines."[49] For example, displacement sensors can monitor the floating height[50] and vibrational modes[51] of the tiny recording heads used in magnetic disk drives. At the other size extreme, displacement sensors also monitor high-speed conveyer systems in the food-processing and pharmaceutical industries.[52]

Metalworking

The largest metalworking industries support the automotive, aviation, construction, and farm-equipment industries. Strong potential for optical-fiber sensors has been shown in the machine-tool and robotics manufacturing sectors.[53]

Displacement sensors are particularly important in metalworking. They are used "to detect the presence of materials in die processes and to verify that drill bits are intact before automatic drilling processes begin."[54]

Flow sensors are important in a wide variety of applications including aerodynamic design of the sheet metal for automobiles. The optical-fiber flow sensors offer "more detailed information than is available from conventional flow monitoring techniques."[55]

In-Process Control, and Process Industries

The process industries include principally the chemical, petroleum, and plastics industries, as noted above. Optical-fiber sensors have the advantage that they can often be placed directly in the process stream.[56] Thus, they can reach previously inaccessible parts of processes, even at high temperatures or under other adverse conditions.[57] They monitor temperature, pressure, flow rates, liquid levels, and other quantities.

Chemical sensors have been used to monitor the curing process in polymer reactions.[58] The sensors can potentially monitor octane levels in gasoline at refineries and pumps.[59] Chemical sensors have been particularly successful in the petroleum and chemical industries and in manufacturing applications requiring chemicals and solvents (as well as in aerospace applications). The sensors are already monitoring processing areas, storage tanks, other containers, and pipelines, all the way to the end users.[60] Here the safety they offer in the presence of flammable liquids is paramount. Chemical sensors are monitoring ground-water contamination around processing facilities such as refineries. They detect xylene and toluene in gasoline spills.[61] More generally, chemical sensors are expected to provide increasing services for environmental applications, whether in the process industries or others; included will be detecting airborne, liquid, and solid contaminants.[62]

Temperature sensors are important to the steel, plastics, and glass production industries where high temperatures are required. Temperature sensors are important for induction-heating processes where immunity to interference is essential.[63] Temperature sensors are key elements in fully automating production processes to achieve better product uniformity and reliability. The food processing, agricultural, pharmaceutical, and plastics industries are showing an interest in temperature sensors for use at lower temperatures as well, typically at 177 °C (350 °F).

Flow sensors are used to monitor the flow of chemicals into processes and the flow of waste materials out of processes, such as the processes used in manufacturing semiconductors.[64]

Electrical Power

Electrical and magnetic sensors may offer improved means for controlling and optimizing the efficiency of the national power grid. An important characteristic of optical-fiber sensors for this application is their immunity to electromagnetic interference in high-power environments. Also of importance is their ability to make measurements all along a power line.

Although the power industry consumed 73 percent of all optical-fiber sensors for magnetic and electrical measurements in 1990, acceptance in this industry has been slow for several reasons, including the long times required for careful evaluation of the new sensing technology and the low rate of turn over in power equipment. The growth of optical-fiber communications technology may spur future use in this industry,[65] which may wish to integrate communications, monitoring, and control functions into a single computer-controlled optical-fiber system.

Medical

In the biomedical area, the small size and non-toxic nature of optical-fiber sensors enables them to be placed in the body to provide continuous monitoring. A chemical sensor has been developed that uses fluorescence techniques to monitor simultaneously the acidity (pH) of blood and the presence of carbon dioxide and oxygen in blood.[66] Chemical sensors are showing promise for determining the concentration of oxygen in hemoglobin by sensing its color.[67] Chemical sensors can be used for tissue analysis, to distinguish normal cells from cancerous cells.[68] Both temperature and pressure sensors also show promise for measurements within the human body. They are simpler and less dangerous to the body than present sensors used for these quantities.[69] Displacement sensors are monitoring ear-drum movement to support research into hearing loss.[70]

GOALS OF U.S. INDUSTRY FOR COMPETITIVENESS

U.S. manufacturers of optical-fiber sensors are pursuing a number of goals in their effort to make optical-fiber sensors competitive with other types of sensors and to assure U.S. competitiveness in the world market for optical-fiber sensors. These goals can be stated as a number of challenges that must be overcome. Some of these challenges are primarily non-technical in nature, and others have significant technical content.

Challenges Primarily Non-Technical in Nature

There are several challenges that are primarily non-technical in nature. A lack of familiarity with optical-fiber sensors among potential users is a major barrier to the adoption of the sensors. That barrier can only be slowly overcome, as shown by the case of optical-fiber temperature sensors. These sensors have been on the market for 10 years, but they are still slow to be adopted.[71] A second barrier is low turnover in equipment employing older types of built-in sensors. A third barrier is the costly downtime that a factory or system must incur to make a major change in sensor technology, entirely aside from the cost of the new equipment itself.

Challenges with Major Technical Components

There are a number of other challenges that have distinct technical components. They are summarized in Table 131 and are drawn from the *important* and *critical* challenges shown in Table 123 above. These challenges fall into four major groups. The first three groups relate to the fibers themselves and address performance, cost, and modularity and standardization of design. The remaining group relates to signal-handling issues that affect optical systems more broadly; they are networking/multiplexing and optical signal processing. In considering many of these challenges, one must keep the excellent performance and continued evolution of electronic sensors in mind. It is likely that electronic sensors and optical-fiber sensors will continue to co-exist indefinitely.

Performance

Optical-fiber sensors have already demonstrated great potential for high dynamic range and reliability for some measured quantities in specific applications, but important challenges remain. Included among these challenges are achieving higher accuracy and higher selectivity, extending dynamic range down to lower values (higher sensitivity), and proving reliability across a wide variety of sensors.

Among these challenges achieving higher accuracy and higher selectivity are particularly critical. There is an interaction between these two challenges. Optical-fiber sensors are often sensitive to quantities other than those that they are intended to measure. Improving their rejection of changes in the unwanted quantities, that is, their selectivity, is a major challenge for design engineers. Independence of mechanical stresses, temperature, and wavelength are especially important. Promising new techniques are emerging to address the selectivity problem. For example, fluorescent decay has potential for temperature measurements. The fluorescent coating on the end of the fiber, after illumination by a light pulse, fluoresces with a rate of decay that is dependent only on temperature. Because only a relative intensity measurement is required, losses due to attenuation in the optical-fiber measurement system have no affect on the result, provided sufficient light is generated to support the processing components. Achieving high levels of selectivity is the first step to achieving desired levels of accuracy. Once higher selectivity is achieved, achieving higher levels of accuracy, beyond the gains made through improved selectivity, will be necessary.

While optical-fiber sensors have already demonstrated great potential for reliability, proving that reliability convincingly will be necessary before optical-fiber sensors can gain wide acceptance. Industry will put these sensors to demanding tests consistent with the advantages of these sensors for

Table 131
Technical Challenges to the Realization of Optical-Fiber Sensors

Performance	
accuracy	In some applications, such as revenue metering for the electric power industry, optical-fiber sensors must have very high accuracy. The fundamental limits to accuracy are set by materials properties and their variation with environmental conditions.
selectivity	The optical properties of materials are sensitive to environmental influences other than those that are to measured. Sensors must be made insensitive to quantities not to be measured.
dynamic range/ sensitivity	Optical-fiber sensors must often measure quantities over a wide range of values. Materials properties must be exploited in refined component designs to achieve this goal. For some measured quantities, optical-fiber sensors are among the most sensitive available; but in other cases higher sensitivity is needed. High sensitivity requires full exploitation of the properties of existing and new materials and the use of special manufacturing techniques.
reliability	In many of the applications for which optical-fiber sensors provide the greatest benefits, they will be subjected to adverse environments. Sensors must be resistant to broad temperature ranges, vibration, high pressures, and ionizing radiation. Their reliability must be established in service and on the shelf awaiting service.
Cost	Reduction in the costs of optical-fiber sensors will require design improvements and the automation of manufacturing processes.
Modularity and Standardization	Optical-fiber sensors must be reduced to standardized, modular, plug-in forms for electronic and optical systems. This is necessary to enable easy substitution in present electronic control systems and to enable incorporation into emerging all-optical control systems.
Signal Handling	
networking and multiplexing	Optical-fiber sensors must be successfully multiplexed and networked in all-optical systems to serve in emerging applications that require multiple sensors, equivalent to sophisticated "nervous systems".
optical signal processing	The emergence of optical signal processing will enable the optimum utilization of optical-fiber sensor data at the high data rates that the sensors can produce.

operation in adverse environments in which other types of sensors are already known to fail. Important testing environments will be chemical processing facilities and nuclear power plants.

Cost

The cost of optical fibers will have to be reduced before they can gain widespread acceptance in many commercial applications. At the present time the costs are high, and reducing those costs is a critical challenge for the future of optical-fiber sensors. The average unit cost of a sensor for 1990, and the projected average cost for 1996, by measured quantity, are shown in Table 132.[72]

Table 132
Average Unit Cost of Optical-Fiber Sensors by Measured Quantity

Measured Quantity	Average Unit Cost ($) Per Sensor		
	1990	1996	Percent Change
rotation	17,000	5,822	-66
magnetic/electrical	3,636	3,218	-11
pressure	1,417	1,276	-10
chemical/biomedical	10,142	9,404	-7
displacement/proximity	322	303	-6
flow	24,166	23,100	-4
temperature	4,333	4,732	+9
level	645	943	+46

The table is arranged in increasing order of the percentage of change (increase) in average unit cost expected from 1990 to 1996. Note that for six of eight categories of measured quantities in Table 132, the anticipated percentage of change in the average unit cost is not really that great for the six-year interval: less than or equal to 11 percent. However, a major reduction in the average unit cost is anticipated for rotational sensors (optical-fiber gyroscopes), as noted above. In contrast, major increase is anticipated in the average unit cost for level sensors. This increase reflects the emergence of more sophisticated, more expensive level sensors and their dominance over the average unit cost. Less sophisticated, less expensive sensors, which had an average unit cost in the range of $200 to $300 in 1989, are expected to fall to about half this level during the period.[73]

The cost changes reflected in Table 132 are based on an estimated world market for sensors that will consume three times as many sensors in 1996 as in 1990, as shown in the unit shipments in Table 128 above. This increase in sensor consumption is likely not great enough to enable realization of the economies of scale that can drive unit costs down significantly. When volume rises to a level sufficient to motivate the development of automated manufacturing processes, then costs can be driven down steeply. Thus, the usual problem is evident in introducing a new technology: low cost requires achieving high volume first, but high volume requires achieving low cost first.

Modularity and Standardization

Sensors are most easily used when they are modularized and standardized, that is, when the sensors are built into "packages" with standardized ways of connecting them to signal lines. Sensors that have been modularized and standardized are easily replaced when in need of repair or upgrading.

Also, sensors that have been modularized and standardized, even if made by different manufacturers, can be used compatibility in a given system. Such sensors give users greater choice of sources and reduce concern about reliance on a single source for continued supply. These benefits increase the likelihood of adoption by users.

Package technology is very important for cost reasons, also; the cost of packages for optoelectronic components broadly represents about 75 to 80 percent of product manufacturing costs presently. Package technology must provide a high level of ruggedness for many applications, including those in the military.

A high level of modularity and standardization has already been achieved for electronic sensors. Optical-fiber sensors will have to reach higher levels before they will be able to compete in the broader sensor market.

Signal Handling

The optimal use of sensors requires sophisticated methods for signal handling, including networking, multiplexing, and signal processing. Again, electronic sensors are well supported in these areas, a distinct advantage. Electronic sensors are typically used as part of control systems which will themselves remain electronic for the foreseeable future. However, optical-fiber sensors can benefit from present electronic methods of signal handling by employing single-processing components that translate their optical outputs into electrical form. As optical-fiber communications systems are used increasingly, many optical-fiber sensors will be able to interconnect with them naturally, will be capable of delivering data rates that take advantage of the information capacity of the fibers, and, as a result, will likely be increasingly used. Further, optical signals in optical-fiber communications systems are relatively easily multiplexed compared to those in electronic systems; that is, a single optical fiber can readily carry the signals from a large number of sensors. In fact, as noted above, certain types of optical-fiber sensors are capable of distributed sensing all along their length. The spatial resolution obtained is limited primarily by the performance of the signal-processing components used with the sensors.

At a future time, when all-optical methods of signal handling become available, optical-fiber sensor will experience an additional impetus for use. For example, a key need is for an inexpensive optical spectrometer, perhaps in the form of an optical integrated circuit. An optical spectrometer is a device capable of determining the wavelength composition of the arriving light. Since many optical-fiber sensors employ a wavelength mechanism, the availability of an inexpensive optical spectrometer would greatly further the adoption of these sensors. Similarly, an inexpensive polarimetric detector, for determining the change in polarization of light, would accelerate the adoption of optical-fiber sensors using polarization mechanisms.

MEASUREMENT NEEDS

Optical-fiber sensors measure a great variety of quantities, so it is not surprising that a wide range of measurement needs arise in developing, manufacturing, and marketing the sensors, and in supporting them after sale (installation, operation, and maintenance). Note that the focus here has shifted from the measurement capability of optical-fiber sensors themselves to the measurement capability that U.S. manufacturers of sensors need to provide these sensors -- a related but somewhat different matter.

Measurement Needs Addressed Here

Not all of the measurement needs associated with optical-fiber sensors are addressed here. Rather, the focus here is on the largest subset of those measurement needs -- those associated with the optical and optoelectronic properties common to a broad spectrum of optical-fiber sensors.

Measurement Needs Not Addressed Here

There are other measurement needs that are not addressed here. They are the measurement needs that are specific to individual classes of sensors and that are generally non-optical in nature. For example, measurement needs associated with specific chemical reactions required to support chemical sensors are not addressed here.

Also, only selected measurement needs of optical integrated circuits are addressed here. They are those related to the special functions served by optical integrated circuits for optical-fiber sensors. However, additional information on measurement needs for semiconductor materials used in optical integrated circuits is contained in Chapter 4, Semiconductors, beginning on page 53. Also, addition information on measurement needs for optical integrated circuits used in optical-fiber communications systems is contained in Chapter 9, Optical-Fiber Communications, beginning on page 217.

Types of Measurement Needs

Three types of measurement needs are examined here: measurement methods; measurement reference standards to assure the accuracy of those measurement methods; and materials reference data to support the design and manufacturing of optical-fiber sensors. Needs related to measurement methods include new measurement methods, improvements in existing measurement methods, and documentation or standardization of new or existing measurement methods or associated definitions of measured quantities to promote proper use.

The types of improvements needed in measurement methods supporting optical-fiber sensors often take the forms shown in Table 133. Especially important are the need for higher accuracy and for increased dynamic range, particularly at the low end, that is, increased sensitivity. Also important is the need for standardized measurement methods that more systematically account for the dependency of measured results on changes in mechanical stresses, temperature, and wavelength, particularly when these quantities are not the intended measured quantity. As an example, an important need is for standardized methods for applying stress as part of a measurement method for attenuation in optical fibers. The standardized method must enable obtaining repeatable results for attenuation in the presence of stress. A similar problem arises when measuring quantities in the presence of temperature and wavelength changes. For special applications, measurements made in the presence of ionizing radiation pose additional similar problems.

Components Requiring Improved Measurement Support

The measurement needs of optical-fiber sensors are usefully referenced to the components within an optical-fiber measurement system that require the improved measurement support. There are a variety of components that are used in optical-fiber sensor systems. They have been described in general terms above. To support discussion of their measurement needs, the components are catalogued in

Table 133
Types of Improvements Needed In Measurement Methods for Optical-Fiber Sensors

- higher accuracy
- increased dynamic range
- standardized methods for accommodating unwanted dependencies on:
  - mechanical stress
  - wavelength
  - temperature

more detail in Table 134, where they have been divided into four groups according to their function in an optical-fiber sensing system. The *illuminating* components provide any illumination needed by the sensors. The *sensing* components change a lightwave (or, in some cases, create a lightwave) in a way that represents the value of the quantity to be measured. As explained above, intrinsic sensing is associated with changes in the optical performance of the fiber itself; extrinsic sensing is provided by components external to the fiber. The *transmitting* components interconnect the sensing components with both the illuminating and the processing components. They also carry, merge (combiners and multiplexers), and separate (splitters and demultiplexers) the lightwaves. The *processing* components interpret the light changes in terms of the measured quantity of interest. Usually the processing components include devices that convert the light into an electrical signal. This conversion requires a detector and often other components (such as polarizers or waveplates) whose selection is determined by the nature of the change in the light to be detected. Increasingly, the processing components will take the form of optical integrated circuits which will carry out more of the processing directly in the form of light.

Table 134
Functional Groups of Components in Optical-Fiber Sensing Systems

Functional Group	Principal Components in Group
illuminating	lasers, light-emitting diodes, super-luminescent diodes, couplers (splitters), lenses, fiber pigtails
sensing	
intrinsic	fibers, claddings, coatings
extrinsic	prisms, microbellows, coatings (fluorescent), microinterferometers, crystals, and many others
transmitting	lenses, fibers, connectors, couplers (splitters, combiners), multiplexers, demultiplexers
processing	detectors, polarizers, couplers (splitters), optical integrated circuits

Some of the components in Table 134 appear in more than one functional group. For example, couplers (splitters, in this case) can be used as illuminating components when monitoring the light level from the source, as transmitting components for dividing the primary signal path, and as processing components.

Quantities Requiring Improved Measurement Support

The quantities for which improved measurement capability is needed are shown in Table 135. The check marks indicate the type of measurement capability needed. The table also provides examples of the types of components that will benefit from the improved measurement capability.

Table 135
Measurement Needs for Optical-Fiber Sensor Components

Measured Quantity	Measurement Method	Measurement Reference Standard	Materials Reference Data	Examples of Components Associated With the Measurement Need
relative power				
inherent attenuation	✓		✓	fibers, integrated circuits
diattenuation	✓			waveplates, fibers
excess loss	✓			couplers
coupling ratios	✓			couplers, multiplexers
return loss	✓			couplers, multiplexers
refractive index		✓	✓	virtually all
phase	✓			fibers, couplers, integrated circuits
polarization				
holding parameter	✓			fibers, integrated circuits
extinction ratio	✓			couplers, multiplexers, demultiplexers, polarizers,
birefringence				
phase difference	✓	✓	✓	fibers, waveplates, polarizers
beat length	✓			fibers
magneto-optic Verdet constant	✓		✓	bulk materials, fibers, integrated circuits
electro-optic linear coefficients			✓	bulk materials, integrated circuits
stress-optic coefficients	✓		✓	bulk materials, fibers, integrated circuits
mode-field diameter	✓			fibers, integrated circuits
geometry				
core concentricity	✓			fibers
core ellipticity	✓			fibers
noise				
intensity	✓			laser diodes
phase	✓			laser diodes
wavelength	✓			laser diodes
linewidth (spectral purity)	✓			laser diodes

For many of the measured quantities, there is a need for reference materials data for critically important materials. One particularly important need is for data describing the variation of the quantity of interest with temperature for a variety of materials. Temperature changes are the single most important source of unwanted variation in the output of a sensor designed to measure some other quantity. The reference materials data are critical to the design of new sensors and to the development of the automated manufacturing processes that can reduce the cost of the sensors.

The measurement needs summarized in Table 135 are discussed in more detail in the sections that follow. The interrelationships among many of the measured quantities are highlighted.

Relative Power

There are a variety of measurements needs that reduce essentially to measurements of the relative levels of optical power in a lightwave. The most common of these is *attenuation*. Attenuation is a measure of the relative loss of optical power of a lightwave as it travels through a material.

Measurement methods for the attenuation inherent in optical fibers are generally adequate. However, extensions are needed to standardize the methods for applying mechanical stresses to the fibers while measuring the effects on attenuation. Stresses from both *microbending* and *macrobending* are important. Microbending is, roughly speaking, the bending of a fiber at intervals of order 1 millimeter with amplitudes (departures from straightness) of order 10 micrometers. Macrobending is the bending of a fiber with a radius of curvature from a few millimeters up to a centimeter or so. Above a centimeter radius, bending an optical fiber has little effect on its optical properties. The specified method for applying mechanical stresses for attenuation measurements is needed both for developing sensors that measure stress and for developing sensors that must be immune to stress when measuring other quantities.

Similarly, standardized measurement methods for determining attenuation in optical fibers in the presence of ionizing radiation are needed to support emerging applications of optical-fiber sensors inside nuclear reactors. Also needed are reference materials data that describe the effects of ionizing radiation on attenuation in optical fibers.

New measurement methods are needed for attenuation inherent in the pathways within optical integrated circuits. Since these pathways are entirely internal, there are usually no accessible external connections to support an external measurement method.

An improved measurement method is also needed for a special form of attenuation -- diattenuation. *Diattenuation* is a measure of the difference in attenuation experienced by two different polarization components of a lightwave travelling through an optical component. Representative components for which diattenuation measurements are important are optical fibers and waveplates. Waveplates are optical devices placed in an optical path to shift the phase of lightwaves. An improved measurement method is needed to enable better distinguishing between the attenuation levels associated with each of the two different polarizations.

An improved measurement method is also needed for determining attenuation in couplers. Couplers combine (combiners) or split (splitters) lightwaves or do both simultaneously. They may have an arbitrary number of input and output ports. For splitters, in particular, a measurement method with improved sensitivity is needed. This need arises because the attenuation quantity of interest is a differential one that involves more than just two measured values. It is the attenuation associated with the difference between the light power level entering the input port and the sum of the light power levels coming out of the output ports. This difference is called the *excess loss*. An improved measurement method is also needed to provide greater accuracy for determining the variation of the *coupling ratio* with wavelength and polarization. The coupling ratio is a measure of the relative power emerging from the output ports of a coupler or a multiplexer. The coupling ratio is also called the splitting ratio. Finally, a measurement method with improved accuracy is needed for determining the relative power reflected from the input ports of couplers, multiplexers, and other optical components as well. This ratio is called *return loss*. The key challenge is separating reflected light from incident light.

Refractive Index

The refractive index is an inverse indicator of the speed of light in a material relative to the speed of light in a vacuum (larger index, lower speed). Light slows down as it moves from a vacuum, or air, into a denser medium, such as a glass fiber. The present level of measurement capability for determining the refractive index is adequate. The principal need is for reference materials data that will provide, in particular, the variation of refractive index with wavelength and temperature for critical optical materials. Also needed are measurement reference standards to assure the accuracy of measurements made for refractive index. Reference materials data on refractive index will support the design and manufacture of virtually all optical components.

Phase

As noted above, some measured quantities are detected through the change that they produce in the phase of a lightwave. Present measurement capability for phase is not adequate to take full advantage of optical-fiber sensors that employ phase mechanisms. The principal need is for greater sensitivity to small phase changes, that is, for greater dynamic range at the low end. An important motivation is the need to determine the phase shifts that occur when fibers are subjected to varying temperatures. Temperature changes cause changes in both the refractive index and the length of a fiber, and these in turn produce phase changes.

Polarization

As a lightwave travels through free space or through an optical material, its electric and magnetic fields are perpendicular to each other and to the direction of travel. As noted above, the polarization of a lightwave is the angular position of the electric field about the axis of travel. Fibers used for telecommunications are generally indifferent to the state of polarization of the lightwave transmitted through them. But fibers used in the sensing mechanism of many optical-fiber sensors and in the transmitting fibers used with those sensors must carefully preserve the state of the polarization. Such preservation is necessary when the polarization is carrying the information about the measured quantity; such preservation is also necessary when a change in polarization might affect processing components that are attempting to detect a change in another characteristic of the light.

For optical fibers, an important measure of polarization is the *holding parameter*. It indicates the degree to which the fiber holds the incoming polarization per unit of distance travelled along the fiber. The *extinction ratio* is a closely related quantity; it is a measure of the extent to which the fiber holds the incoming polarization over the entire length of the fiber. Present measurement methods for these two quantities lack adequate sensitivity for determining (1) low levels of the unwanted polarization, and (2) small changes in the desired polarization. Also needed are standardized methods for the measurement of holding parameter and extinction ratio in the presence of mechanical stress, particularly microbending.

For bulk materials, like those used in polarizers, polarization measurements need additional refinement to achieve adequate sensitivity. Measurement of the extinction ratio with 100 times greater sensitivity than presently possible is needed and cannot be obtained without the refinement.

For couplers, the principal need is for a more comprehensive definition of polarization behavior in couplers. This definition must reflect both the effects of the polarization of the incoming light and the fact that the outgoing light from a given output of the couplers may be elliptically polarized.

Birefringence

Birefringent materials present a different index of refraction to each member of a pair of polarization components of a lightwave. That pair may present a simple polarization (linear) or a more complex polarization (circular or elliptical). In a birefringent material, one of the pair of polarization components travels at a different speed from the other polarization component. This difference in speed causes the polarization of the light to change as it passes through a birefringent material.

The relationship between the pair of polarization components in a birefringent fiber is determined by measuring either (1) the phase difference that develops between the components over distance, or (2) the beat length, which is the distance over which a phase difference of 360 degrees develops.

For the measurement of phase difference in birefringent fibers, the principal need is for a measurement method with improved accuracy. For the measurement of beat length in optical fibers, the principal need is for standardization of the definition of beat length. There are two different definitions in use, each of which is consistent with one of two different existing measurement methods; these two different approaches reflect somewhat different types of phase differences.

When characterizing the birefringence of components made from bulk optical materials, such as waveplates and polarizers, the measurement of birefringence is somewhat easier. The principal need is for reference materials data for important optical materials, showing, in particular, the variation of birefringent properties with temperature and wavelength. These data are needed to support the design of sensing elements. These elements must be compensated for changes caused by temperature and wavelength shifts when these quantities are not the measured quantities of interest. Further, for waveplates, whose function is to produce precise changes of phase, a measurement reference standard is needed to support the accuracy of measurements of phase difference.

Magneto-Optic Verdet Constant

The Verdet constant describes the variation of the refractive index with magnetic field (Faraday effect). Present measurement methods for the Verdet constant are basically adequate, but they need to be augmented by an expanded definition for ferrimagnetic materials. The expanded definition must accommodate both saturation effects at high fields and a shape dependency in measured properties. Most important, however, is the need for materials reference data for important magneto-optic materials exhibiting paramagnetic, diamagnetic, and ferrimagnetic properties. [These properties are explained in Chapter 5, Magnetics, beginning on page 95.] The materials reference data should show the variation of the Verdet constant with temperature. A variety of components are affected including the bulk materials used in extrinsic sensing elements, optical fibers themselves, and optical integrated circuits.

Electro-Optic Linear Coefficients

The linear coefficients describe the variation of the refractive index with the electric field in a lightwave (Pockels effect). Only certain crystalline materials show the linear effects, such as those used in extrinsic sensing elements. Glass optical fibers are amorphous and do not. The primary need is for materials reference data that show the variation of the linear coefficients with temperature, wavelength, and materials composition. Affected components include bulk materials used in extrinsic sensing elements and optical integrated circuits.

Stress-Optic Coefficients

The stress-optic coefficients indicate the effects of physical stresses on the refractive index of an optical material. Those stresses, as noted above, include bending, twisting and axial and transverse stretching and compression. The principal need is for a standardized method for applying stress to an optical fiber so that reproducible results can be obtained for the stress-optic coefficients. Stress can induce differential changes in the two different directions of polarization in a birefringent fiber. Stress can also change the refractive index in a non-birefringent material. Knowledge of the stress-optic coefficients is important for optical fibers, bulk materials used in extrinsic sensing elements, and optical integrated circuits.

Mode-Field Diameter

The mode-field diameter refers to the size of the cross-section of the light emerging from an optical material. For many of the fibers used in optical-fiber sensors, the mode field is not round (as in communications fibers) but rather elliptical. This ellipticity results from the design of certain fibers, such as the birefringent fibers, for polarization preserving properties. Present measurement methods are basically adequate for characterizing the mode-field diameter with one principal exception: A standardized approach is needed to assure that the effects of the elliptical shape on the measurement process are properly accommodated. This need arises also for mode-field diameter measurements made within optical integrated circuits, including measurements made on waveguides with polarization-preserving properties.

Geometry

Measurement methods with higher accuracy are needed both for the concentricity of the core of an optical fiber with the outer periphery of the fiber and for the ellipticity of that core. New measurement methods are needed for determining the geometric properties of the waveguides embedded in optical integrated circuits.

Noise

For laser diodes, improved measurement capability is needed for determining the noise in emitted light. Noise is unwanted variation in the light. Noise may reflect variation in the intensity of the light (intensity noise) or in the wavelength of the light (phase noise). Intensity noise (also called relative intensity noise) can affect the accuracy of measurements made with intensity, phase, or wavelength mechanisms. Phase noise can affect the accuracy of measurements made with phase mechanisms. Noise places a limit on the sensitivity of optical-fiber sensors for a wide range of measured quantities. Measurement needs for lasers diodes more broadly are discussed in Chapter 8, Lasers, beginning on page 203.

Wavelength and Linewidth

The wavelength of a laser diode varies with temperature. In fact, controlled changes in temperature can be used as a method for controlling the wavelength of laser light. Unfortunately, as the temperature is changed, the modal composition of the light from a laser diode may also change, affecting not only its wavelength but also its linewidth. A standardized measurement method is needed for determining the changes that occur in wavelength and linewidth with changes in temperature.

ENDNOTES

1. A high degree of *repeatability* is important for a sensor. A sensor is highly repeatable if its electrical or optical output has very nearly the same value each time is exposed to a given value of the quantity, such as pressure, to be measured.

2. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation (Mountain View, California), p. IV-4 (1991).

3. *Sensor and Instrumentation Markets, 1992-1995*, Richard K. Miller & Associates, Inc., pp. 6, 8 (1992).

4. Experimental evaluation of silicone rubber fibers is underway. They are readily embedded in other materials and show a useful sensitivity to a variety of measured quantities of interest including pressure, relative humidity, and strain, among others. Jeffrey D. Muhs, "Silicone Rubber Fiber Optic Sensors", *Photonics Spectra*, p. 99 (July 1992).

5. Michael P. Kovacs, "Fiberoptic Sensors Approach Commercial Success", *Laser Focus World*, p. 175 (March 1991).

6. A. A. Boiarski and N. E. Nilsson, "New Fiber Sensors Take Power Plant's Temperature", *Photonics Spectra*, p. 93 (September 1991).

7. Kausar Talat, "Fiber Sensors Take Wing in Smart-Skin Applications", *Photonics Spectra*, pp. 85-88 (April 1991).

8. Richard O. Claus, 'Fiber Sensors as Nerves for "Smart Materials"', *Photonics Spectra*, p. 75 (April 1991).

9. This definition is based on the emerging definitions under development by the Institute of Electrical and Electronics Engineers and the Telecommunications Industry Association.

10. As described in Chapter 9, Optical-Fiber Communications, an optical fiber is composed of a core surrounded by a cladding. The core is concentric with the cladding. The optical properties of both the core and the cladding are essential to the transmission of light down the fiber. The core has a higher index of refraction (and thus a slower speed of light) than the cladding. The interface of the cladding with the core reflects light back to the core, thus guiding the light along the core.

11. A coating is a layer of material applied outside of the cladding that normally surrounds the core of an optical fiber.

12. Industry is experimenting with a variation in which a single fiber contains two cores, that is, two pathways for light, one for each direction of travel to the sensing element.

13. Refractive index is explained on page 328.

14. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. V-2 (1991).

15. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. V-2 (1991).

16. Robert S. Mellberg, *Fiber-Optic Sensors*, Stanford Research International, Research Report No. 684, p. 3 (Summer, 1983).

17. *Photonics: Maintaining Competitiveness in the Information Era*, National Research Council, National Academy Press, Washington, DC, p. 51 (1988). The quoted market for automated controls is assumed to be the U.S. market, also.

18. *Sensor and Instrumentation Markets, 1992-1995*, Richard K. Miller & Associates, Inc., p. 7 (1992).

19. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. III-3, III-12, and III-15 (1990).

20. According to *Fiber Optic Sensors Market (Europe)*, Report No. E1204, Frost & Sullivan (1989) as reported in *Sensor and Instrumentation Markets, 1992-1995*, Richard K. Miller & Associates, Inc., p. 7 (1992).

21. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. III-3, III-12, and III-15 (1990).

22. According to Business Communications Company, as reported in *Sensor and Instrumentation Markets, 1992-1995,* Richard K. Miller & Associates, Inc., p. 86 (1992).

23. According to Corporate Strategic Intelligence, as reported in *Sensor and Instrumentation Markets, 1992-1995*, Richard K. Miller & Associates, Inc., p. 87 (1992).

24. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. III-9 (1991).

25. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. III-16 (1991).

26. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. III-6 (1991).

27. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. XIV-4, -5, -11, -18, -28, and -33 (1991).

28. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. XIV-9 and XIV-10 (1991).

29. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. XIV-31 (1991).

30. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. V-8, VI-9, VII-9, VIII-8, IX-11, X-10, XI-11, and XII-7 (1991).

31. The number of blocks shown was determined by rounding the market share to the nearest 10 percent and then dividing by 10. No block is shown if the relative share is less than 5 percent for a given end-use industry.

32. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. I-2 and I-3 (1991).

33. *OITDA Activity Report, Vol. 5, For Fiscal Year Ended March 31, 1992*, Optoelectronic Industry and Technology Development Association, pp. 5-6 (1992). The datum is for the 1991 fiscal year used in Japan (ending March 31, 1992) rather than for the calendar year 1991. The datum was published in yen and was converted to dollars by dividing by 134.59 yen per dollar. This conversion factor is the average annual exchange rate for calendar 1991, based on a Federal Reserve Statistical Release of January 24, 1992. An average annual exchange rate for Japan's fiscal year was not available.

34. According to Corporate Strategic Intelligence (Middlebush, NJ), as reported in *Sensor and Instrumentation Markets, 1992-1995*, Richard K. Miller & Associates, Inc., p. 88 (1992).

35. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. I-2 and I-3 (1991).

36. According to Corporate Strategic Intelligence (Middlebush, NJ), as reported in *Sensor and Instrumentation Markets, 1992-1995*, Richard K. Miller & Associates, Inc., p. 88 (1992).

37. Michael Shimazu, "The Ball Game is Bigger Than OEICs", *Photonics Spectra*, p. 75 (October 1992). The information provided comes from the book *The New Optoelectronics Ball Game: The Policy Struggle Between the U.S. and Japan for the Competitive Edge* by Philip Seidenberg of Georgia Tech, published by IEEE [Institute of Electrical and Electronics Engineers] Press in 1992. Ernest Sternberg, "Should We Have a Policy for Optoelectronics?", *Laser Focus World*, p. 59 (September 1992).

38. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. VI-6 and VI-7 (1991).

39. Michael P. Kovacs, "Fiberoptic Sensors Approach Commercial Success", *Laser Focus World*, p. 183 (March 1991).

40. Otis Port, "Materials That Think for Themselves", *Business Week*, pp. 166-167 (December 5, 1988).

41. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. VII-8 (1991).

42. Otis Port, "Now Fiber Optics Can Hear and Feel, Too", *Business Week*, p. 168A (December 5, 1988).

43. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. X-7 to X-8 (1991).

44. George A. Pavlath, "Fiber Sensors Demand More From Components", *Photonics Spectra*, p. 106 (June 1992).

45. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. XIV-16 to XIV-25 (1991).

46. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. VIII-7 and IX-12 (1991).

47. Michael P. Kovacs, "Fiberoptic Sensors Approach Commercial Success", *Laser Focus World*, p. 182 (March 1991).

48. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. XI-4, XI-7, and XI-8 (1991).

49. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. X-4 (1991).

50. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. XI-7 (1991).

51. Michael P. Kovacs, "Fiberoptic Sensors Approach Commercial Success", *Laser Focus World*, p. 181 (March 1991).

52. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. XI-9 and XI-10 (1991).

53. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. XIV-26 (1991).

54. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. XI-9 (1991).

55. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. XII-5 (1991).

56. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. XIV-31 (1991).

57. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. VIII-6 (1991).

58. Michael P. Kovacs, "Fiberoptic Sensors Approach Commercial Success", *Laser Focus World*, p. 175 (March 1991).

59. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. V-6 (1991).

60. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. X-7 (1991).

61. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. V-6 (1991).

62. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. V-3 to V-4 (1991).

63. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. VIII-5 (1991).

64. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. XII-6 (1991).

65. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. VII-8 (1991).

66. *Fibre Optic Sensors: Technology Developments in the USA*, ERA Technology Ltd., p. A.13 (April 1989).

67. Otis Port, "Now Fiber Optics Can Hear and Feel, Too", *Business Week*, p. 168A (December 5, 1988).

68. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. V-6 and V-7 (1991).

69. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. VIII-7, IX-9, and IX-10 (1991).

70. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. XI-8 (1991). Michael P. Kovacs, "Fiberoptic Sensors Approach Commercial Success", *Laser Focus World*, p. 181 (March 1991).

71. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. VIII-5 (1991).

72. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, pp. V-5, VI-5, VII-6, VIII-3, IX-6, X-3, XI-5, and XII-4 (1991).

73. *World Markets for Fiber Optic Sensors*, Report No. 477-40 of Market Intelligence Research Corporation, p. X-5 (1991).

CHAPTER 11

VIDEO

January 1993

Chapter 11

VIDEO

SUMMARY

In the previous decade, television, computers, and telecommunications were considered largely separate technologies. Now the promise of advanced video systems is fostering a merger of these technologies to a provide a powerful "window" to the information age. This merger will provide new capabilities for diverse applications in education, engineering, manufacturing, robotics, entertainment, medicine, defense, security, transportation, publishing, advertising, banking, and government.

The hallmark of advanced video systems is the simultaneous presentation of high-resolution images, full color depth, and smooth motion to take full advantage of the capabilities of the human eye. To date, computer monitors have provided high resolution but without smooth motion and rarely with full color depth. Similarly, conventional television has provided smooth motion and full color depth but without high resolution. The marriage of these three capabilities, plus a merger with networked telecommunications services, particularly on-demand services, promises users access to libraries of video, audio, and text information whenever they wish.

The emergence of advanced video systems is being made possible by advances in five supporting technologies:

(1) high-resolution image generation that gives high-quality video
(2) real-time signal processing that prepares video images for transmission, storage, and display
(3) high-data-rate transmission that includes optical-fiber and microwave communications networks for video images
(4) high-density information-storage systems that record and store video images
(5) high-resolution displays that reproduce high-quality video images

The expanded markets that advanced video systems will create for the products of these five supporting technologies are vast. Already, world markets for the products of these technologies are about $154 billion dollars per year (1990). Even modest percentage increases attributable to advanced video systems will yield large markets.

International competition to meet the demand for advanced video systems is intense. The winners in this competition will penetrate markets of high diversity. At the present time, the manufacturing of advanced video systems is dominated by Japan. Japan now broadcasts 8 hours per day of high-definition television (HDTV) programs. The U.S. is a lesser player in the manufacturing of advanced video systems, but the U.S. will certainly be a major consumer. However, the U.S. leads in computers and in the digital technology used to process signals at high speeds, both of which reflect strength in the supporting technologies required for advanced video systems. Building on this base, the U.S. is beginning to make significant contributions to advanced video systems.

More broadly, the U.S. industry is attempting to improve its competitiveness in advanced video systems by pursuing key goals in developing, manufacturing, and applying advanced video products. Industry's goals, common to all five of the supporting technologies, are these:

(1) improve the infrastructure among suppliers, manufacturers, and users of advanced video systems
(2) reduce costs of components for advanced video systems by increasing yield, quality, and reliability of such components
(3) increase the performance of individual components, including the number of functions performed per component

To assure that U.S. industry can accomplish these goals and compete successfully in the development of advanced video systems, and to assure that the U.S. can consume video technology intelligently, greatly improved measurement capability is needed. Specifically, new measurement capability is needed to support:

(1) research and development of new advanced video systems for diverse applications
(2) design of, and control of quality during, manufacturing processes for components with improved performance and lower cost
(3) marketplace entry for U.S. video products, and evaluation of components for purchase
(4) installation and maintenance of advanced video systems

The competitive manufacture and application of advanced video products demand a sophisticated, underlying infrastructure that provides high-performance communications networks, protocol and other standards, computers, and other hardware and software. The protocols are standard practices for controlling and exchanging data on the network. Protocols affect both the physical architecture and the controlling software of network systems. Critical to this infrastructure will be the numerous suppliers and vendors of component parts that must produce high-quality and reliable materials, products, and services for those enterprises associated with the information age.

ABOUT THIS CHAPTER

Organization

This chapter begins with a description of advanced video technology. World markets for products related to advanced video technology and the U.S. competitive position are discussed. Technological alternatives for implementing advanced video technology are reviewed. This discussion and the rest of the chapter are organized around the five technologies that support advanced video technology:

(1) high-resolution image generation that gives high quality video
(2) real-time signal processing that prepares video images for transmission, storage, and display
(3) high-data-rate transmission that includes optical-fiber and microwave communications networks for video images
(4) high-density information-storage systems that record and store video images
(5) high-resolution displays that reproduce high-quality video images

The discussion continues with an assessment of the major technical challenges that U.S. industry faces in pursuit of its goals for competitiveness in advanced video technology. Finally, the measurement needs that these technical challenges give rise to are discussed. The purpose of the order of this discussion is to show how the measurement needs derive from the goals that industry must pursue for competitiveness.

Emphases

The measurement needs for high-resolution flat-panel displays are discussed in considerable detail in this chapter. Also, an assessment is provided of the measurement needs for real-time signal processing. Further, a preliminary assessment is provided for high-resolution image generation; a more complete analysis of this area is contemplated for a future edition of this document. The measurement needs for high-density information-storage systems are discussed in detail in Chapter 5, Magnetics, beginning on page 95. The measurement needs of high-data-rate transmission systems are discussed in detail in Chapter 7, Microwaves, beginning on page 147 and in Chapter 9, Optical-Fiber Communications, beginning on page 217.

The emphasis given in this chapter to flat-panel displays is driven largely by the intense competition already evident in this area. Nations, especially Japan, are investing very large amounts of resources to manufacture flat-panel displays since the markets for products containing displays are expected to be enormous. Japanese companies are investing over $2.0 billion in flat-panel-display plants costing from $45 million to $300 million each for the first stage.[1] Additional stages may bring the total investment to over $900 million for some individual plants. The realization of the key role that displays play, and the investments that Japan and the U.S. have already made, have produced much interest in display technologies.

OVERVIEW OF ADVANCED VIDEO TECHNOLOGY

Television, computers, and communications have long been viewed as separate services, using different technologies, and requiring different business approaches. However, as technology advances, these services are gradually merging.[2] For example, the outputs from supercomputers are frequently presented as video images that are sent to displays at remote sites. The term "advanced video technology" is used here to summarize this common technology. The term includes the technology needed to generate, process, store, transmit, and display video images with the high resolution required for clear presentation of publication-grade text and graphics, with the full color depth required for life-like images, and with the high speed required for smooth motion. Advanced video technology has the potential to be implemented with networking to interconnect multiple users, on an international scale, and to provide access to on-demand services for video, audio, text, graphics, and data.

Advanced video technology is also one of the most demanding technologies that the world has attempted to commercialize. For example, some have stated that the large-area active-matrix liquid-crystal display, one of several proposed display methods, is among the most complex devices that man is attempting to mass produce with high yields and low costs.[3] Others have stated that the active-matrix liquid-crystal display is too complex and will not succeed as a large-area display. The success of advanced video technology demands collaborative efforts that span numerous disciplines such as mathematics, chemistry, physics, materials science, networks, computer science, electronics, mechanical engineering, and manufacturing science. It requires sophisticated, high-performance, and reliable electronic and optical components and networking strategies. All of these must be implemented at affordable costs.

Advanced video technology can serve so many fields in so many ways that its impact is difficult to predict. At a minimum, advanced video technology can serve education, engineering, manufacturing, robotics, entertainment, medicine, defense, security, transportation, publishing, advertising, and banking. More generally, advanced video technology can serve business and government broadly through improved office automation, electronic mail, and teleconferencing. Advances in technology can make these applications possible. If the technology base is in place, the mix of available services will be determined by market, regulatory, and other factors.

Some specific examples of the services that advanced video technology is beginning to provide are these:

> *Education* -- Systems are being developed to permit individuals to gain access through networks to giant, multi-media and interactive libraries of information in text, audio, advanced-graphic, and video forms. For example, students are able to read text, hear music, and see still or moving pictures of historical events.
>
> *Entertainment* -- High-resolution video programming, accompanied by superb sound quality, are becoming available through high-definition television (HDTV).
>
> *Medicine* -- Improved consultations with medical experts at distant locations are becoming possible using pictures of outstanding clarity and color accuracy for surgical procedures, live sonograms, x-ray images, and the like. Physicians are increasingly able to display simultaneously the medical images from different diagnostics procedures for improved and timely health care.
>
> *Business and Government* -- High performance and widely available teleconferencing capabilities could become as common as the telephone today, furthering collaborative efforts and fostering creativity among individuals working across great distances. Telecommuting could change family life, business, and traffic.
>
> *Home* -- Family members at different locations could share holiday events through giant flat-screen displays and superb sound systems. Individuals could gain immediate electronic access to the day's newspapers and newsreels, displayed with sharp text and color pictures, all easily searchable by busy people for topics of interest.

Such new services will require high levels of information-handling capacity, provided at affordable cost. Advanced video systems will demand more information capacity than any other products yet contemplated for broad use in both commercial and consumer markets. The development of this new technology will require the development of a new measurement infrastructure simultaneous with the development of new or improved equipment. Representative equipment that will be required by advanced video systems is shown in Table 136.

Table 136
Representative Equipment Needed by Advanced Video Systems

Function	Equipment Category	Representative Equipment and Specific Functions
Generation of High-Resolution Video Images	video cameras	arrays of semiconductor sensors such as charged-coupled devices and scanners or detectors, operating in the visible or infrared regions, for robots in manufacturing, multi-media systems, and television
	software	computer-generated images, computer-aided design, computer-aided manufacturing, and computer-aided engineering
	diagnostic equipment	magnetic-resonance imaging, positron-emission tomography, computer-assisted tomography, X-ray, scanning-electron microscopes and atomic-force microscopes
Real-Time Signal Processing	video processing equipment	video processors with limited programming, video-supercomputers with extensive and user-friendly programming, and compression-decompression circuits
	application-specific integrated circuits and semiconductor chip sets	encode and decode multimedia digital signals, implement compression/decompression protocols according to standards such as those recommended by Motion Picture Experts Group and Joint Photographic Experts Group
	electronic circuits	display drivers, control of individual pixels, amorphous and crystalline silicon, polysilicon, compound semiconductors
High-Data-Rate Transmission	lightwave	optical-fiber cable
	radio frequency, including microwave	terrestrial broadcasting equipment, direct broadcast satellite, mobile communications, coaxial cable
	software	narrow-band integrated-services digital network, broadband integrated-services digital network, fiber distributed data interface, and high-performance parallel interface
High-Density Information Storage	magnetic	high-speed, high-density heads and materials, for tapes and disks
	optical	lasers, read-write technology
Display	cathode-ray tube	brightness, field reduction
	flat-panel	active-matrix liquid-crystal, front-projection, rear-projection, plasma, electroluminescent displays
	projection	transmissive and reflective liquid crystals, deformable micromirrors

WORLD MARKETS AND U.S COMPETITIVENESS

Proponents of information technology, a topic which includes advanced video technology, predict that the information technologies will produce the same sweeping economic impact in coming years that the railroads and interstate highways have in the past. Thus, it is not surprising that advanced video technology has been targeted for development by foreign governments and firms as well as by U.S. firms.

The economic impact of advanced video technology will come from at least three factors:

(1) the sale of components and equipment for advanced video technology;
(2) the sale of the services performed with the new technology; and
(3) economic stimulation within the countries exploiting those services for business and commerce.

The increased productivity of these services attributable to advanced video technology stimulates further improvements in design, manufacturing, financial services, and other areas of endeavor.

The predicted sweeping impact has a precedent in the introduction of television. Since the experimental broadcasts of television in the U.S. and Europe in the 1930s, television has spread explosively throughout the world. Today, there are an estimated 630 million television receivers (290 million monochrome and 340 million color sets) in the world. According to United Nations statistics the world's population is about 5 billion. This means that on average 1 out of every 8 people in the world owns a television receiver[4] and represents a large potential market for higher resolution television receivers. All of the components needed for consumer distribution of high-definition television receivers are now available with the exception of large displays at reasonable cost. Most of the cost of a consumer high-definition-television installation is for the display because of its large size and high resolution. Reasonably priced consumer high-definition television receivers with large displays are expected by 2000 to 2003 and considerable growth in the market for consumer high-definition television should occur then. In the nearer term, the use of high-definition displays for theaters, conference presentations, education, and manufacturing will provide the required public exposure to high-resolution systems and motivate the public to become the future consumers of high-definition television.[5]

The markets for components and equipment alone for advanced video systems is predicted to be enormous. Even though published market surveys do not present data in terms of the subcategories for the five supporting technologies for advanced video systems, estimates are possible. Table 137 shows estimates of the world market for 1990 for those subcategories of the five supporting technologies for which some market data are available.[6] More than $154 billion dollars per year of components and equipment are already sold in the subcategories. Thus, it is not surprising that the world market for the equipment required for just one application of advanced video technology, high-definition television, has been projected at rather large values, typically $5 billion to $20 billion per year beginning between 2000 and 2010.[7] Also, the world market for displays is just for devices. The value increases substantially when the supporting electronics and cabinet costs are added.

Japan, Europe, and the U.S. are all very active in communications networks. Japan is experimenting with a very high-speed version of the Integrated Services Digital Network that operates at 1.2 gigabits per second.[8] This rate is about twice that of the so-called Broadband Integrated Services Digital

Table 137
World Equipment Markets Affected by Video Technology (1990)

Function	Equipment	World Market ($billions)
Generation and Transmission	broadcast and studio equipment[6(a)]	4
	microwave, optical fiber, and satellite systems[6(b)]	35
	telephone and telegraph equipment[6(c)]	38
Signal Processing	memory and application-specific integrated circuits[6(d)]	28
Storage	magnetic[6(e)]	37
	optical[6(f)]	1
Display	cathode-ray, flat panel[6(g)]	11
Total		154

Network. The first standards for Broadband ISDN were published by the International Consultative Committee for Telephone and Telegraph in 1990.

In Japan, development of the Integrated Services Digital Network is part of a government-sponsored national technology project called the Information Network System. It is part of the Japanese Government's goal to take Japan into the information age as quickly and as coherently as possible. Similarly, in Europe, most of the Integrated Services Digital Network efforts are government controlled. For example, West Germany has been running, since the mid 1980s, a pilot fiber-optic broadband network that operates at 140 megabits per second.[9] Although trailing Europe, the U.S. is now realizing services from a lower-speed version, the Narrowband Integrated Services Digital Network, in major cities and business centers in the U.S.[10]

The United States had in 1991 an international trade surplus in optical fiber and the cables made from that fiber. This trade surplus amounted to $185 million and was a 37 percent increase from 1990. However, many foreign-based companies produce fiber, cable, and fiber-optic components in facilities located in the U.S. Also, this trade figure only includes optical fiber and cable; the trade figure does not include optoelectronic products, connectors, splicing equipment, and other products of the optical-fiber communications industry. The Japanese and Europeans continue to be effective competitors. The world market for fiber optics was about $4.25 billion in 1991, of which the U.S. market was $1.7 billion or 40 percent.[11]

Microwave integrated circuits promise major increases in performance per unit cost, compared to existing technology, and contribute to enabling the use of small direct-broadcast satellite receiving antennas used at some sites.[12] The countries first implementing them will gain major competitive advantages, not only in broadcast technology for video, but also in the world market for microwave equipment.

Processing video signals will require increased demands for semiconductor memory and application-specific integrated circuits. For example, today's typical editing video cassette recorder with split-screen capability contains a few hundred kilobytes of memory whereas tomorrow's high-definition television receivers are expected to contain about eight megabytes of memory. Hence, more than a factor of ten increase in the demand by video technology for semiconductor memory and application specific integrated circuits may occur.[13] The world market for semiconductor memory and application specific integrated circuits was about $28 billion in 1990.[14]

For magnetic storage systems, the world market for all types of magnetic storage devices and systems was about $52 billion in 1990.[15] About $37 billion of this market was devoted to the high-density information-storage systems, as explained in endnote 6; these are the systems most closely related to those that would support advanced video systems. In the computer market for rigid disk drives, the U.S. held a strong 77 percent share of the world market of $26 billion in 1990. This share had increased from 73 percent in 1987. However, the U.S. has essentially lost the computer market for floppy disk drives to Japan. The U.S. market share dropped from 5.8 percent in 1987 to 3.3 percent in 1990. Japan dominates the market for video cassette recorders. U.S. production of video cassette recorders accounted for less than 10 percent of the total U.S. consumption in 1990. Reliable data on U.S. competitiveness for all products associated with optical storage systems are lacking. However, the U.S. is one of the major consumers of these products.[16]

For displays, the U.S. is a manufacturer but not a leader. For example, factory sales in the U.S. of black and white and color television picture tubes was about 15 million units valued at $1.54 billion in 1991. In 1991, the U.S. imported television picture tubes valued at $74.3 million and exported television picture tubes valued at $342.6 million.[17] The Japanese lead in producing flat-panel displays with a market share of about 98 percent. Presently, liquid-crystal displays dominate the flat-panel display market, and most manufacturing experience and economic data pertain to liquid-crystal displays. For example, the Nomura Research Institute estimates that the liquid-crystal flat-panel display market will grow from $1 billion in 1990 to about $7 billion in 1995.[18] The Semiconductor Equipment Manufacturers Institute also estimates that the flat-panel display market will grow to about $7 billion in 1995. The active-matrix flat-panel display market is expected to represent at least 10 percent of the value of the entire integrated-circuit market this year. Other display types, in descending order of market share, are vacuum-fluorescent, plasma gas-discharge, light-emitting diodes, and electroluminescent displays. Vacuum microelectronic displays are still in the research stage.[19] According to Department of Commerce data, companies located in the U.S. shipped $21 million worth of liquid-crystal displays in 1990. This is an increase from $17 million in 1989. There are about 10 companies located in the U.S. that ship liquid-crystal displays.[20]

Both Japan and Europe have taken specific actions that indicate their seriousness about advanced video technology. Japan broadcasted the Seoul Olympic Games in high-definition television to selected public locations. Japan offers a full range of products for high-definition television. One example is a magnetic tape recorder that employs eight parallel tracks and has heads that travel at high speeds (184 kilometers per hour) relative to the tape in order to produce the required record and playback rates.[21] Japan is also pursuing development of programmatic material and of its domestic market through its "Hi-Vision" programs. Both Europe and Japan have selected direct-broadcast satellites as their initial transmission method.

Either analog or digital methods may be used to transmit high-resolution advanced television. Each method has its advantages and disadvantages. However, results after more than a year of testing at the Advanced Television Test Center indicate that digital transmission systems generally are better than analog systems.[22]

The emergence of advanced video technology offers the possibility of dramatic changes in competitive positions. As the functions of computers, telecommunication systems, and television begin to merge under the influence of advanced video technology, new markets for "integrated" products will replace former "single-function" products. For example, as computers and television sets merge, U.S. companies will have an opportunity to use their strength in computer technology to penetrate television markets. Similarly, companies with strength in television technology will have an opportunity to penetrate computer markets.

IMPLEMENTING ADVANCED VIDEO SYSTEMS

Advanced video systems are likely to be implemented in a variety of ways to address different applications or to accommodate different goals in capability, compatibility, or cost. Some of these approaches are summarized in Table 138 and are discussed in the following sections.

Encoding

Encoding refers to the method used to prepare information for transmission and storage. Present television pictures are transmitted and stored with analog encoding. Analog encoding is accomplished by converting the video information to a continuously varying electronic signal. For example, a video camera in a broadcast studio generates smoothly varying signals with voltages that are related to the brightness and colors in the picture. These signals are then used to modulate carrier waves; that is, they typically change either or both the amplitude or frequency of the carrier wave. The exact modulation technique depends on how the video picture is to be transmitted or stored. It may be transmitted through air, space, copper wire, or optical fiber. It may be stored on magnetic tape or optical disks.

Another approach to encoding a signal is digital encoding. Digital encoding means converting the analog information signal into a series of numbers (zeros and ones) like those computers use. The numbers are stored or transmitted to a desired receiving location. When received, or recovered from storage, the numbers are used to reconstruct the original analog information signal.

Digital encoding has a number of advantages over analog encoding. It enables reconstruction of the original signal with a predetermined level of quality, even in the face of considerable degradation during transmission and storage. That degradation may take the form of reduced signal strength, introduction of noise, introduction of distortion, or other effects. Digital encoding also facilitates processing the signals in powerful ways, including compression. Finally, digital systems are directly compatible with computer data, with emerging telephone networks (including the several different speed levels of the Integrated Services Digital Network), and with printers to produce paper copies.

Table 138
Alternatives for Implementing Advanced Video Systems

Function	Alternatives
Encoding	analog, digital, or a combination of the two
Picture quality	continuum of choices
Compression	degree of compression (compression ratio) usually based on trigonometric functions or other mathematical constructs such as fractals
Transmission	broadcast (one-way) air (terrestrial and satellite) cable (coax and optical fiber) network (two-way) cable (coax and optical fiber)
Storage	magnetic, optical, and ferroelectric in two or three dimensions
Display	cathode ray tubes or flat screens liquid crystal vacuum fluorescent plasma gas discharge light-emitting diodes electroluminescent vacuum microelectronic
Compatibility	with today's TV, and among multi-media systems and computers

Digital encoding has disadvantages, too. The fastest digital circuits are not as fast as the fastest analog circuits and may cost more for a given speed. Therefore, the use of digital encoding may raise the cost of a system and, if major speed advances are required, can delay implementation. In the past, digitally encoded signals generally required more transmission capacity or storage capacity than their analog counterparts. However, with effective digital compression algorithms, it is now often possible to develop digital channels that use less bandwidth than their analog equivalents. Therefore, digital encoding is used in digital cellular telephone systems, military voice radio systems, and high-definition television to reduce bandwidth requirements.[23]

A third alternative is a combination of analog and digital encoding. Such combinations can reduce the total amount of digital information that can be processed and can facilitate achieving compatibility with present analog systems. Some digital processing techniques are used in present-day television receivers, video cassette recorders, and audio receivers to provide performance enhancements, such as picture-in-picture formats and surround sound with many audio channels.

Picture Quality

Picture quality is determined by a number of factors including:

1. the number of lines in the picture
2. the number of identifiable picture elements (pixels) in each line, where a pixel is the smallest element in a display that either passively or actively emits light
3. the number of brightness levels for each color in a pixel
4. the number of pictures created per second
5. the "scanning" technique used to create and send each picture
6. the degree of compression

A scanning technique is required because all video pictures are sent line by line. That is, the information across the area of the picture must be reduced to a steady stream, or line, of data before it can be sent. In progressive scanning, every line of a picture is sent in succession at a rate of 60 complete pictures per second. In interlaced scanning, which is used in today’s NTSC television, all odd-numbered lines are sent in succession, followed by all even-numbered lines, producing 30 complete pictures per second. Compromises exist between picture quality (resolution) and data rate (bandwidth) for these two scanning methods. Progressive scanning produces the highest quality picture for a given number of lines. To achieve the same *perceived* resolution with interlaced scanning, the number of interlaced lines has to be increased by a factor of about 1.3. Hence, for the same perceived resolution, progressive scanning requires 1.5 times the bandwidth of interlaced scanning. However, progressive scanning makes it much easier and cheaper to interface with multimedia systems based on computers and to do post-production editing videos for the final version shown to viewers.

A combination of the above and other factors determine the quality of the picture. The picture quality for advanced video systems can be selected from numerous possibilities. Improvements of factors of two to more than ten, relative to present television, are generally contemplated for various implementations. A factor of ten will be assumed as the basis for the rest of the discussion in this Section; however, lower factors are also under consideration. A factor of ten improvement would mean that a digitally encoded advanced video signal would contain about 1 billion bits per second of information, prior to compression, which is discussed below. A bit is the smallest unit of digital information and is represented by either a zero or a one. In comparison, a conventional television signal, if encoded in digital form rather than its usual analog form, would require about 100 million bits per second of information.[24]

Compression

The large amount of information associated with high-quality pictures is particularly difficult to accommodate during transmission and storage. For this reason, considerable thought is being given to methods of reducing the amount of information that must be transmitted or stored to achieve a picture with a given level of "perceived" quality. By taking advantage of the characteristics of human vision when viewing moving images, reduction, or compression, is possible. Compression is done before storage or before transmission; and the complementary process, decompression, is done after playback from storage or after reception, respectively.

Compression techniques have disadvantages, too. They can increase the cost of the receivers by requiring them to have complex signal processing circuits to decompress and decode the incoming signal. Also, highly compressed video data are highly sensitive to errors introduced by transmission or storage. Tolerable error rates may be limited to 1 bit wrong out of every 10 billion to 100 billion transmitted or stored bits.[25] This is about 10 to 100 times more demanding performance than the raw performance of most networking hardware used with computers. Such hardware is usually designed for an error rate no greater than 1 bit wrong out of 1 billion bits.[26] In contrast, uncompressed video data are relatively insensitive to errors because of the type of material transmitted and because it passes through fewer electronic circuits.

Compression is not unique to video systems and is possible because much of the data being sent may not be necessary. Compression is already used for video-like systems such as FAX machines, teleconferencing, and robotic vision. Compressing is also used for speech. In fact, modern digital speech systems, such as those in new digital telephone lines, use compression to reduce data requirements from 64 to 8 kilobits per second, or even less with loss of some naturalness. Data compression will become an important tool in future low-speed video applications, including security systems for passport and fingerprint storage, retrieval, and identification.

Transmission

Advanced video systems may employ one or more of a number of transmission methods. Some of these are broadcast, or one-way, methods; others are network, or two-way, methods. The broadcast alternatives are summarized in Table 139. They are arranged in order of increasing information handling capacity. That capacity increases with the frequency of operation of the transmission method.

Table 139
Video Broadcast Transmission Alternatives

Transmission Method	Frequency of Operation
terrestrial broadcast	submicrowave frequencies (below 1 gigahertz)
direct-broadcast satellite	microwave frequencies (above 1 gigahertz) with principal frequencies of interest above 10 gigahertz
coaxial cable	submicrowave frequencies (below 1 gigahertz)
optical-fiber cable	optical frequencies (200,000 gigahertz)

Terrestrial Broadcasts

Terrestrial broadcasts, the first over-the-air broadcast approach listed in Table 3, have some key advantages. They offer the prospect of compatibility with present television. They also operate on relatively low frequencies (below 1 gigahertz) for which receiver electronics are relatively inexpensive. Terrestrial broadcasts have disadvantages, too. They have limited information capacity because of their relatively low frequency of operation. Further, they cannot operate on frequencies much above 1 gigahertz because those frequencies do not propagate well along the surface of the earth. However, with new compression techniques, advanced video signals can be sent in existing channel widths. The tradeoffs between the number of channels and performance are technically and politically very complex.

Direct-Broadcast Satellites

Direct-broadcast satellites -- the second over-the-air broadcast approach -- operate somewhat differently. The satellites are positioned in high, geosynchronous orbits above the equator where they orbit exactly with the rotation of the earth. Thus, they remain always over the same locations. They send signals directly to the antennas of individual recipients, along a line of sight. Such satellites for advanced video technology will likely operate at frequencies above 10 gigahertz, for two reasons: these frequencies offer high information-carrying capacity, and they work well with small antennas. During the 1988 Olympic games, the Japanese demonstrated antennas as small as 45 centimeters for receiving direct broadcasts from satellites.[27] Another approach is to use low-orbit satellites, flying at altitudes of typically 1000 kilometers[28] (600 miles) that communicate with each other and phased-array receiving and transmitting antennas for two-way, multimedia communications on earth.

Direct-broadcast satellites have the advantage that they are relatively easy to implement. A single satellite can serve a huge geographical area, even the entire Unites States. They are also readily reconfigurable; that is, they can be redirected to serve different areas, especially if they are equipped with electronically steerable antennas (antennas that can be directed by electronic commands alone, without physical movement). Direct-broadcast satellites have disadvantages, too. The receiving equipment that individual users must have is more expensive than receivers for the lower-frequency terrestrial broadcasts. The high cost is attributable in part to the high frequency of operation and in part to the high sensitivity required to receive relatively weak satellite signals from high satellite orbits at 29,000 kilometers (18,000 miles) above the surface of the earth. Much less sensitivity is needed to receive more powerful terrestrial broadcast stations from distances of only a few tens of kilometers.

The basic concept of direct-broadcast satellites is not novel, even if the performance levels required for advanced video systems are extremely high. The U.S. presently uses satellites to deliver television programs to local television stations for terrestrial rebroadcast. Those transmissions can be received directly by users in a manner entirely similar to direct-broadcast satellites. In fact, 1.5 million such receivers are already in use in the U.S., particularly in rural areas with limited access to terrestrial broadcast stations or cable television.

Broadcasting by Cable

Broadcasting can also be accomplished over cable, either coaxial cable, which is employed by nearly all present cable television systems, or optical-fiber cable. Coaxial cable is capable of delivering advanced video signals. However, because of the physics by which a signal moves in a guiding structure such as coaxial cable, there is a maximum frequency above which the signal cannot propagate in the guiding structure. This maximum frequency is called the cut-off frequency. It depends on the size, shape, and construction materials of the guiding structure. For commonly used coaxial cables, the cut-off frequencies limit operation to relatively low frequencies, typically below 550 megahertz, which is equivalent to 0.550 gigahertz.

Since the amount of information that a transmission system can carry increases with its frequency of operation, coaxial cable can carry somewhat less information than those transmission systems operating at higher frequencies. Optical fiber cable, which operates at much higher frequencies (200,000 gigahertz), offers hundreds of times lower losses, and has potentially thousands of times higher information capacity. But realization of that potential will require development of higher performance optoelectronic components that can exploit more fully the natural capacity of the fiber itself. Since the cost of optical-fiber communications systems is decreasing, those systems can soon contribute to growth of the advanced video industry.[29] The cost of providing optical-fiber cable to the home is rapidly approaching the cost of providing conventional twisted-pair copper wiring to the home. In fact, fiber-to-the-home is being considered actively in a number of states.[30]

Comparison of Broadcasting Alternatives

Comparing the over-the-air broadcast approaches, as a group, to the cable broadcast approaches, the over-the-air approaches have the advantage that they are easier to implement, since they do not require wiring every recipient, resolving difficult right-of-way issues, or navigating difficult terrain. However, the cable approaches have advantages, too. Television cable and telephone companies already can provide service to nearly every home and business in the U.S. The entire information capacity of the cable can be dedicated to a single purpose, since the cable system, unlike the over-the-air system, does not have to share available frequency space with highly diverse services such as FM radio, mobile radio, aircraft and satellite communications systems, police and fire services, radar systems, and many others. Figure 7 shows how the electromagnetic spectrum is shared among some major users. Also, the cable system does not require users to cope with interference or with antenna installations that must accommodate varying local terrain and distances from transmitting locations.

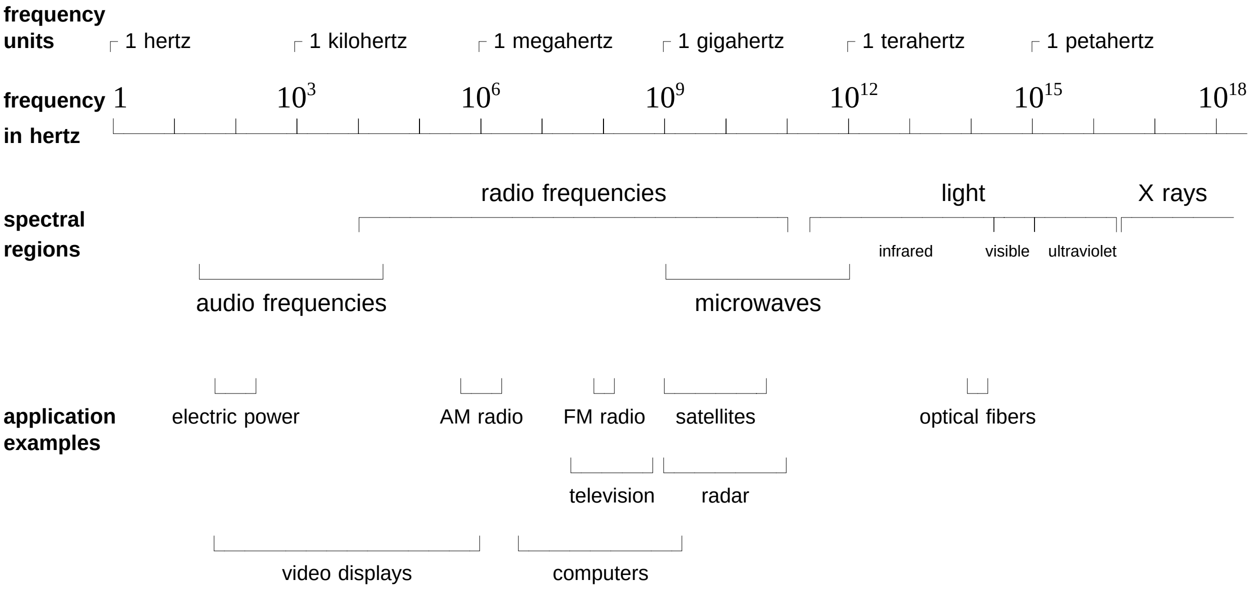
At present, 61 percent or 55.8 million households of the 91 million U.S. households with television receivers are served by cable. This is 64 percent of the 87.4 million homes that are passed by cable. The remaining 31.6 million homes that could have access to cable rely solely on over-the-air broadcasts.[31] The 1990 Census will report that of the 92 million occupied households in the U.S., 5.2 percent or 4.8 million occupied households do not have telephones.[32] In many ways, the separation between cable based services such as television and telephones is a matter of public policy and not technology.

Networking

An alternative to the broadcast, or one-way approach, is the networking, or two-way approach. Networks route transmissions flexibly between arbitrary points of transmission and reception. Networks for advanced video technology require so much information-handling capacity that over-the-air approaches are unlikely. They could support only very small numbers of users in the spectral space accessible with technology available in the near term. Cable approaches will likely be necessary.

Among the cable approaches, coaxial cable presents some difficulties for networking for advanced video technology because of its limited information capacity. Coaxial cable is not likely to provide adequate service to large numbers of users in a networking environment.

Figure 7
Overview of the Frequency Spectrum



Optical-fiber cable, however, will be able to serve large numbers of users in advanced video networks, if high performance components that can exploit the cable's capacity fully can be developed and deployed. In fact, optical-fiber cable systems should eventually be able to carry at least 1000 uncompressed advanced video signals simultaneously and tens of times more compressed signals, perhaps enough to provide on-demand video services.

Storage

The term "Storage" as used here refers to long-term storage rather than to the momentary storage provided electronically by most semiconductor memory. Because advanced video systems produce so much information every second, storage becomes a major challenge. Information from advanced video systems may potentially be stored by magnetic, optical, or ferroelectric storage systems with either analog or digital encoding. Today, the primary means for storing electronic or video images are optical disks such as the video laser disk, magnetic tape such as the magnetic tape used in video cassette recorders and in camcorders, and magnetic disks such as those used in personal computers.

Storage with analog encoding is presently more compact than storage with digital encoding, although this may change in the future. However, analog encoding is also more susceptible to degradation of image quality. Storage in optical form enables writing data to, and reading data from, the storage media without physical contact with the media, thus reducing wear. But optical storage systems are presently limited in capacity by the relatively small amount of surface area that they can offer for storage on spinning disks. Much of the equipment for optical storage that is sold today permits the user to write once and to read many times. However, optical storage systems that enable the user to both write and read many times are also widely available and have an increasing market share in the popular 90 millimeter and 130 millimeter disk sizes. These rewritable systems use magneto-optic media for which erasing is slow and causes writing delays. Work is being done on other media that do not use the magneto-optic effect and that would not have the write-delay problem.[33]

Storage in magnetic form on spinning disks is also limited by surface area, whereas, storage on magnetic tape offers much more surface area per unit volume. However, use of magnetic tape requires physical contact for writing and reading data, giving rise to wear. Most storage technologies, such as magnetic tape, typically cannot handle much more than 150 megabits per second of data per track.[34] This is to be compared with high-resolution cameras that generate digital data well in excess of 1 gigabit per second and with display and transmission systems that handle digital data also well in excess of 1 gigabit per second. This means that, based on advanced prototype video systems, in the near term at least, some forms of compression and decompression will be required in order to accommodate the limitations of both storage media and terrestrial broadcast systems when used.

However, different levels of compression are required for different purposes. For example, designers of advanced video systems for terrestrial broadcast have proposed relatively low data rates of 22 megabits per second to maintain the appropriate bandwidth for individual channels. In contrast, relatively high data rates of 300 megabits per second are appropriate for today’s high-speed digital magnetic tape recorders intended for broad use.[35] Without the benefit of emerging compression techniques, the limited storage capacity of magnetic tape used in home video recorders may limit future picture quality to below broadcast levels.

The need for different levels of compression presents an important technical challenge. Most compression algorithms assumed a fully decompressed input. That means that broadcasters would have to decompress, or decode, signals stored on magnetic tape at 300 megabits per second to 1 gigabit per second before they could compress them again to 22 megabits per second. Thus broadcasters would have to support the fully decompressed data rate in at least some of their equipment. Better schemes would permit "piggybacking" the compression (and decompression) stages; that is, they would permit moving directly from one level of compression (or decompression) to another without having to fully decompress in between. Broadcasters could then compress signals from magnetic tape at 300 megabits per second directly to broadcast levels of about 22 megabits per second.

Displays

Many promising alternatives are emerging for display technology. Cathode-ray tubes are used in current television sets and in most personal computers. They produce light by scanning an electron beam (the "cathode ray") across a phosphor screen to stimulate the emission of light. Cathode-ray tubes have the advantages of high brightness, good color reproduction, and reasonably wide viewing angles. However, cathode-ray tubes must be made from strong materials to withstand the pressure of the atmosphere because they are evacuated devices. The tubes become increasingly difficult to implement in large screen sizes.

In contrast, emerging flat-panel displays are considerably smaller and lighter for a given screen size. They are presently more expensive than cathode-ray tubes of the same screen size, but their cost is decreasing and their performance is rising. Color liquid-crystal displays that have diagonal sizes greater than 25 centimeters (10 inches) are being made by several manufacturers, installed, and sold in the tens of thousands per month. A joint venture of a U.S. subsidiary in Japan and a Japanese company produces in Japan a 25-centimeter diagonal, full-color, active-matrix, liquid-crystal display for portable personal computers.[36] A few Japanese companies have demonstrated prototype liquid-crystal displays as large 43 centimeters (17 inches).

Another promising new version of liquid-crystal display is based on ferroelectric materials. Ferroelectric displays have much wider viewing angles than active-matrix liquid-crystal displays. Unfortunately, ferroelectric materials used in displays have a comparative disadvantage, too; they can have only two brightness levels. To produce intermediate gray levels, spatial dithering or temporal dithering, or a combination of the two, must be used. For spatial dithering, a fraction of the pixels in a local region is turned on to produce the desired gray level. For this approach, the fact that ferroelectrics will stay in the assigned on or off state without refreshing is an advantage. For temporal dithering, each ferroelectric pixel is switched on and off rapidly in such a way that its fractional on time gives the desired gray level. For this approach, the high switching rates in microseconds of ferroelectrics are an advantage. For video-rate applications, the driving electronics for temporal dithering will have to be faster than the driving electronics for spatial dithering. Ferroelectric displays have two other disadvantages. First, because the ferroelectric materials must be used in very thin layers enclosed between glass panels with close and uniform spacing, they are more difficult to manufacture. Second, because ferroelectric materials tend to be photosensitive, they may complicate processing steps that involve light.

A Japanese company has demonstrated a 38-centimeter (15-inch) very high-resolution, full-color ferroelectric liquid-crystal display for desktop publishing of Japanese text.[37] Key portions of this ferroelectric technology are based on U.S. research and licensed patents that pertain to surface-stabilized ferroelectric liquid crystals. Processing Japanese text requires much higher resolution, memory size, and data rates than processing the text of most non-Asian languages. In fact, many cite this need as one reason why Japanese companies were motivated to excel in making dense, high-speed, semiconductor memories.

Among the technologies for dot-matrix flat-panel emissive displays, plasma displays are the most widely produced in the world.[38] Plasma displays are candidates for large-area high-definition displays, but only laboratory prototypes are available in full color at video frame rates.[39] These plasma displays use photoluminescent phosphors that are excited by an ultraviolet gas discharge. Such displays are expected to be commercially available very soon. For example, a 48-centimeter diagonal (19-inch), 640-by-480-pixel, full-color plasma display that has 64 gray levels and improved-definition video capabilities was presented at the Display Symposium of the Society for Information Display in May 1992.[40] This display operates at 60 frames per second with progressive scanning; and its luminosity, chromaticity, contrast, uniformity, and viewing angle approach those for cathode-ray tubes.

Compatibility

Compatibility has at least three dimensions: compatibility among variations of advanced video systems; compatibility between these advanced video systems and present television; and compatibility among advanced video systems, multi-media communications systems, and computers for diverse applications.

Compatibility among the variations of advanced video systems offers several advantages. It reduces the cost of the systems by providing high commonality of components, large markets for these components, and thereby economies of scale. It furthers efficient use of advanced video systems by enabling them to serve in multiple applications.[41] It increases the chances that new systems can communicate through a common network, further broadening applications. Compatibility can have a downside, too. It can force systems to be more or less capable than optimal for some applications. However, discussions among computer manufacturers, electronics engineers, and selected broadcasters have suggested ways to prevent this disadvantage from occurring.

Compatibility between advanced video systems and conventional television would enable utilization of the many existing television sets in the U.S. Ninety-five percent of U.S. households currently own a color television set. The Federal Communications Commission has already taken steps to assure that programs broadcast over the air using advanced video technology will be accessible to present television sets, either through compatible implementations of advanced video technology, or through simultaneous broadcast of advanced and present formats.

Once data are in a digital form, processing depends very little, if at all, on whether the data represent video, audio, graphics, text, medical images, or inputs and outputs for computers. Engineers[42] have proposed preceding digital data streams with a descriptor or header that provides instructions to equipment (for example, a high-definition television receiver, an National-Television-Standards-Committee receiver, an engineering work station, or a medical diagnostic display) on how it should process the incoming data. Such broad compatibility permits great economies of scale.

One factor favoring broad compatibility would be the development of image production standards that are suitable for not only high-definition-television images, but also for images generated by computers, engineering work stations, and medical diagnostic and scientific equipment. A single production standard would be preferred because: (1) It would promote efficient international communications. (2) The U.S. is very strong in producing and exporting video sources for entertainment, and this would assist the U.S. to maintain its market share.[43] (3) Personal computers may come to outnumber television receivers and this would facilitate combining computers and television. (4) The post production phase now involves combining images from many sources such as films at 24 frames per second, video recordings at 30 frames per second with interlaced scanning, and computer generated images at 72 frames per second with progressive scanning. At present, combining images without compatible[44] standards is costly and frequently gives lower-quality results.

Several production standards that offer compatibility across numerous applications have been proposed. However, a consensus does not exist now as to the most acceptable standards. The most broadly based production standards suggested and suitable for diverse applications have these attributes: (1) Full digitization, progressive scanning, and square pixels will make it easier and cheaper to interface video with computers, add digital effects, rotate images, alter color, and use other touch-up effects. For example, rotating images is most easily accomplished with square or circular pixels. "Square" and "circular" pixels are those with square and circular areas, respectively. The use of non-square or non-circular pixels requires many more high speed silicon circuits and therefore is more costly and less reliable. (2) The number of frames per second is a multiple of the present 24 frames per second for film cameras. This is required in order to avoid problems associated with converting from 24 frames per second for film to the 30 frames per second for the National Television Standards Committee video standard. (3) The number of pixels in the rows and columns each is a reasonably large power of 2 in order to make it less costly to interface video data with computers. The operation of computers is based on expressing data in terms of powers of 2. (4) The aspect ratio, width divided by height, is close to 16/9.

Telecommunications systems and computers do not use the current rate of 30 frames per second of U.S. television and may not in the future. Frame rates that are multiples of 24 frames per second will eliminate conversion problems that are now present in all post-production work. This problem arises from the incompatibility between the 24 frames per second of film cameras and the 30 frames per second of today's domestic and Japanese television systems and the 25 frames per second of European television systems. Different frame rates exist due to both historical and economic factors. Countries such as those in Europe with electricity that alternates 50 times per second use 25 frames per second for television. Countries such as Japan and the U.S. with electricity that alternates 60 times per second use 30 frames per second. Those production houses that form joint ventures with computer companies in order to offer movies through computers are not likely to use either the 30 or 25 frames per second.[45]

Among the many transmission systems [terrestrial broadcast and cable (coaxial and optical fiber)], only terrestrial broadcast is losing market share. Ten years ago, broadcast (terrestrial) accounted for 80 percent of the viewers. Today it is 55 percent and continues to decline. Some predict that it will not go much lower than 40 percent because there is not a significant product differentiation between television programs available on cable and terrestrial broadcasts. Even though cable brings more channels into the home, it is similar to broadcast in terms of programming material.

In the past, profits from films and television entertainment were used to develop new products for this U.S. industry. But, now, as terrestrial broadcasters lose their share of the market, their ability to finance products of high technical quality is lessened. The cost of development and equipment for high-resolution video systems is so great that a larger market than just the market for high-definition television as entertainment is needed. Changing over to a new system that is compatible for many applications and markets will reduce by economies of scale the cost for each user-application.

At present, there are few alternatives to the high-definition television cameras that offer 1125 lines, 30 frames per second, and interlaced scanning; they dominate the high-end market. The manufacturers of these 1125-line cameras propose to sell them with attached converters ("black boxes") to interpolate to other standards. But, interpolation among standards may degrade resolution and performance and requires additional equipment and complexities that lower overall reliability and increase the costs of advanced video systems.

High-definition television cameras and other production equipment such as that for recording and storage are expected to cost about three times as much as conventional cameras and equipment. One way to lower the prices is by selecting standards applicable to the greatest number of users. However, some companies now offer production equipment based on a production standard that is not compatible with several applications. This standard has 1125 lines, 30 frames per second, and interlaced scanning. Such equipment will require expensive and perhaps less reliable converters to interface with computers, robots, and other multi-media systems.

GOALS OF U.S. INDUSTRY FOR COMPETITIVENESS

U.S. industry is attempting to improve its competitiveness in advanced video systems by pursuing key goals in developing, manufacturing, and applying advanced video products. Industry's goals, common to all five of the supporting technologies, are these:

(1) Improve the infrastructure among suppliers, manufacturers, and users of advanced video systems.

(2) Reduce costs of components for advanced video systems by increasing yield, quality, and reliability of such components.

(3) Increase the performance of individual components, including the number of functions performed per component.

Pursuit of these goals required addressing a number of important technical challenges, each of which is giving rise to a number of critical measurement needs.

TECHNICAL CHALLENGES AND MEASUREMENT NEEDS

The previously mentioned five technologies that support advanced video systems all have technical challenges. Addressing these technical challenges reveals the critical measurement needs for advanced video systems.

The following descriptions of those technical challenges and the measurement needs stress the high performance levels and the new components needed to meet the requirements of advanced video technology. Many of these challenges are not unique to video systems but occur in other technologies such as optoelectronics. One example is the flat-panel display, which technologically is a very large-area optoelectronic device with pixels that embody concepts and technology of direct use in optical computers. The pixel in a display has for some cases, such as the ferroelectric display, two states, i.e., is bistable. Pixels are also very tiny, on the order of tens of micrometers, in many cases. These attributes could be the basis for an optical computer, for switching optical signals at high speeds, or for three-dimensional storage. Also, the electronic circuits that make displays work present considerable demands on the semiconductor industry to increase their performance.

Generation of High-Resolution Images

High-resolution vision systems must capture the signals used for advanced video systems. The standards selected for picture quality will specify the acceptable resolution, but performance levels that are perhaps ten times those of present television are contemplated. Examples of the vision systems referred to here include: video cameras; scanners that read documents such as text, graphics, and photographs; computers that generate still and video images, and robots used for automated manufacturing. Key among the demands placed on such video systems will be their performance with moving images, and in particular, their ability to deliver adequate levels of resolution, sensitivity, and color balance in the presence of moving subjects. Many groups are developing future professional video cameras based on semiconductor circuits that are sensitive to light. These light-sensitive circuits are arranged in a flat array. The image in the camera is focussed on this electronic array and thereby is converted into electronic signals. Video cameras that work on this principle are called "charge-coupled devices" or CCD cameras. Most consumer camcorders also have similar light sensitive circuits but their resolution and sensitivity to light are less than those required for professional video cameras.

The output from high-definition television CCD cameras is digital with a data rate of about 1 gigabit per second. In order to use this data rate with today's post-production digital recorders, operating at 300 megabits per second, compression of about 4:1 is required. The compression may be done in the camera or in the recorder. Greater interoperability (scope of compatibility) results when the compression is done in the recorder. These cameras can be used both for entertainment and for most of the other applications listed above. The camera designers, who address interoperability and extensibility, consider this and assume that the output is scanned progressively with horizontal lines that number from about 1024 to about 2048 lines, 16:9 aspect ratio, and square or circular pixels. As in displays, these pixels have light sensitive areas with square or circular shapes, respectively. The more expensive frame-rate converters will then only be needed when the same source material is released into markets with different frame rates. With today's lower cost solid-state power supplies, the frame rates could become independent of the frequency of the electricity in use. But, today the countries of the world have chosen for economic reasons to continue these differences in frame rates.

Changing the frame rate of a camera simply involves changing the clock rate associated with the electronic circuits in the camera. This is not a technical barrier. However, one major technical challenge is to build progressively scanned solid-state cameras that have acceptable output levels and adequately low noise at high resolution when the subject is dimly lit.

There are some common technical challenges in making high-yield arrays of charge-coupled devices for cameras and in making high-resolution flat-panel displays. These challenges are outlined in Table 140.

Table 140
Technical Challenges for High-Resolution Charge-Coupled Arrays for Imaging Cameras

defect reduction	reducing the defects caused by particles and contamination
color filter quality	producing stable, reliable, and uniform color filters that go over the pixels in cameras and in flat-panel displays
dielectric quality	improving insulation or dielectric properties of very thin films that occur in the semiconductor circuits for video cameras, in the transistors that drive and address the pixels in flat-panel displays, and in the pixels of flat-panel displays themselves

Technical challenges that are more specific to solid-state imaging cameras for high-resolution systems are outlined in Table 141. Time constants limit the speed at which the camera can operate; they, in turn, are controlled by the materials used to make the devices and the device geometry. High data rates are needed since, in advanced systems, more than two million pixels must be accessed at least

30 times per second. Integration is needed not only to achieve high levels of performance but also to achieve high reliability and low cost.

Table 141
Technical Challenges for High-Resolution Solid-State Imaging Cameras

fast response	reduction of the time constants of poly-silicon gates used in charge-coupled devices
high data rates	increasing the data-handling capability of the entire device
integration	integrating onto a single chip the light sensors (e.g., charge-coupled devices) and output amplifiers to produce low noise, high bandwidth, and high light sensitivity

The important measurement needs that must be met to address the technical challenges are shown in Table 142. Today's measurements of sensitivity of cameras to light and to noise are not adequate for many applications. Different divisions in the same company, and different distributors to the same market, usually do not use the same methods for making measurements of camera noise and sensitivity.[46] Measurements made with the different methods are likely to disagree. This part of U.S. industry would benefit substantially from agreed upon, universal ways to measure quantities such as camera sensitivity, noise, and resolution, particularly for moving images. Industry needs such standards to minimize transactions costs in the procurement of systems as well as to design and market improved systems.[47]

Table 142
Measurement Needs for Cameras

noise
sensitivity
resolution for moving images

Real-Time Signal Processing

Most of the information sent in advanced video systems will be processed in "real time". That is, it will be processed as fast as it is generated so that a backlog of unprocessed information does not occur. This is obviously essential for live television, but it is also necessary for playing back recorded material which is to be viewed or listened to at natural speeds. Real-time signal processing is principally relevant to digitally encoded forms of advanced video technology, since digital techniques facilitate the manipulation of all data, including video data.

Important technical challenges facing the industry include developing efficient, inexpensive hardware to perform the functions shown in Table 143. When digital encoding is used, analog-to-digital data converters are needed at the transmission end to translate the video information into digital form. Similarly, digital-to-analog data converters are needed at the receiving end to translate the digital data back into analog form for those displays that cannot accept digital data. These converters may have to operate at microwave speeds (above 1 gigahertz). Present converters rarely operate much above a few hundred megahertz. Analog-to-digital and digital-to-analog data converters are presently used in test equipment, such as voltmeters and oscilloscopes, and in entertainment equipment, such as audio recording equipment, audio playback equipment, laser compact disk players, and in video teleconferencing systems.

Table 143
Technical Challenges for Signal Processing

conversion between analog and digital forms
compression and decompression of an image
conversion among alternative digital formats
manipulation of an image

When digital techniques for compression are used, then digital processors, data compression techniques, and semiconductor memories are required. Some or all of these may operate at microwave speeds (above 1 gigahertz). Each of these is discussed further below.

Digital processors are the circuits that constitute the "brain" of a computer. In advanced video systems, they can be used to compress and decompress transmitted information, to correct errors, to reduce noise, and to create special effects. Digital signal processing is needed for editing and special effects and for combining and manipulating images from different sources, such as video cameras, computers, scientific and medical equipment, and robots and for whatever encoding scheme is used in transmission and storage. Signal processing will also be needed for decoding in receivers, terminals, and displays.

Reducing standards, such as the compression standards of the Joint Photographic Experts Group and the Motion Pictures Expert Group, to practice requires the evaluation of mathematical expressions in hardware. These expressions frequently contain terms that cannot be represented with perfect accuracy by any real implementation. As a result, a single standard may be realized by a variety of implementations leading to different performance levels. Further, different implementations may be needed to achieve optimal performance in different architectures or on different platforms (for example a general-purpose central processor versus a programmable digital signal processor) with different processing capabilities. As a result, tests are needed to quantify the differences and to assess compliance with the standard.

Further compression is often achieved by representing the output with no greater precision than is necessary to achieve the desired image quality. This is called quantization. The goal of quantization is to discard information which is not visually significant. Whether the quantization selected is acceptable depends on source image characteristics (e.g., landscapes, sports, or scrolling text), display characteristics (such as pixel aspect ratio, size, and viewing angle), and viewing distance. Psychovisual experiments are generally needed to determine the best quantization scheme.

Table 144 lists some of the measurement needs for signal processing. This is not an all-inclusive list. Because there is not a mature measurement infrastructure in place, one technical challenge is to design and build the measurement equipment necessary for analyzing the accuracy and precision of competing compression algorithms for given multi-media source material (video, graphics visualization, audio, text, and computer data). Another challenge is that measurements, such as resolution, are complicated by the non-linear nature of compressed systems. Also, independently designed physical implementations of the same compression algorithm will produce different outputs from identical inputs. Those physical implementations include general-purpose central-processing units, programmable digital signal-processing systems, or dedicated hardware. The different outputs occur because each implementation differs in the precision with which mathematical expressions are calculated. The correlations among different algorithms for the same standard, their implementations, and the perceived image quality will be essential measurements. For example, new measurement capability is needed in the digital domain for comparing competing algorithms and their implementations, over a variety of source material, in order to establish objectively the differences among their outputs when their inputs are identical. Also, special testing equipment is needed to evaluate the equipment used to convert among different compression/decompression methods, frame rates, and scanning methods.

Table 144
Measurement Needs for Signal Processing

Feature	Parameter
Interoperability	spatial and temporal resolution field of view and viewing distance aspect ratio image acquisition rate display refresh rate dynamic range conversion artifacts compression frame rate scanning method
Image architecture	data stream characterization testing for compatibility with existing systems image coding for color and intensity coding algorithm colorimetry of source transfer characteristics from one medium to another
Tuning to human vision	flicker/fusion frequency flicker threshold pixel persistence objective correlations for visual effects contrast ratio

Objective physical measurements associated with determining the psychovisual thresholds for efficient quantization would provide flexible, cost-effective means to screen various quantization implementations. For example, artifacts in the methods that will likely be employed in digital video could fatigue some viewers. Objective ways to correlate perceived eye fatigue with quantization parameters would avoid this problem.

Another challenge is to provide video source materials for testing both "lossy" and "lossless" compression (encoding) and decompression (decoding) procedures.[48] These source materials should be objectively correlated with human vision and could be in the format of laser disks or video tapes. One need would be to provide text standards (somewhat like the ticker tapes for stock exchanges) that would scroll in the vertical, horizontal, and diagonal directions. Scrolling text offers a way to have the source material contain known and controlled spatial and temporal frequency components.[49] For example, diagonal scrolling would reveal one kind of artifact; that is, pixels on or near the diagonal that should be either black or white, but instead are gray. When the video system has sufficient resolution, one practical solution to this kind of problem would be to defocus the video camera or to use a camera with less resolution. Note that even for such solutions as these, the eye plays an essential role.

High-Data-Rate Transmission

For implementation of broadcast approaches, the capabilities of existing terrestrial television broadcast systems and existing coaxial-cable systems will have to be expanded to reach the required information capacity. This likely would be accomplished by increasing the efficiency of use of existing channels. With the increasing demand for spectrum use, it is unlikely that new frequencies will be used for television.

High-performance broadcast satellites and ground-based receiving equipment will be required for the implementation of direct broadcasts from satellites. These satellites and ground-based equipment will operate at microwave frequencies and will be capable of delivering signals with high information capacity. Satellite antennas with superb control of beam shape and direction will likely be needed to enable direct broadcast satellites to concentrate their limited energy on specific geographic regions. Higher power microwave sources for satellites will also be needed to reduce the sensitivity required by the multitude of receivers. Microwave integrated circuits will be needed to reduce the cost, size, and weight of microwave satellite electronics and ground-receiver electronics.

For implementation of optical-fiber cable in broadcast mode with digital encoding, higher information capacity will be required for optical-fiber communications systems. New components will be needed that can exploit the information capacity of the fibers. In particular, high performance modulators, multiplexers, switches, and amplifiers. The promise of coherent systems will have to be explored; coherent systems will support many more channels and will enable "tuning" among them, much as present television does for radio waves. Further, optical-fiber lines will have to be installed for local delivery to offices and homes, not just for the cross-country and undersea routing. This requirement will increase the market for optical-fiber system components. These are all part of the essential infrastructure for high performance computing and communications and for which a competitive domestic manufacturing base in optoelectronics will be very advantageous.

To implement networking on optical-fiber systems, high-performance components will again be required. Protocols will also be needed. Ideally, these protocols should be accepted internationally to permit interconnecting information systems worldwide. The most significant international protocols currently under development are those for the Integrated Services Digital Network.

The Integrated Services Digital Network integrates support for text, audio, video, graphics, and computer data. Two levels of Integrated Services Digital Network (ISDN) are presently the focus of international attention. The Narrowband ISDN, or just ISDN for short, can support data rates up to 1.544 megabits per second for an individual user. The Broadband ISDN can support data rates up to 622 megabits per second for an individual user. An international protocol for narrowband ISDN was adopted by an international standards body, the International Consultative Committee for Telephone and Telegraph, in December 1988. A whole spectrum of proposed protocols and standards exist for Broadband ISDN. Some portions of international protocols for Broadband ISDN were voted on by the International Consultative Committee for Telephone and Telegraph in 1992. For example, Study Group XVIII of the Committee adopted the Synchronized Optical Network standard and defined header formats for the asynchronous transfer mode cells. Narrowband ISDN, with a maximum delivered data rate of 1.544 megabits per second, does not require optical fibers and does not have sufficient capacity for advanced video signals. However, it can send video information at low resolution or slow frame rates that might be acceptable for teleconferencing and VHS-quality entertainment.[50] Broadband ISDN will employ optical fibers and will be able to accommodate tens of highly compressed advanced video signals. Variations of ISDN with even greater capacity are the subject of international research.

Other approaches to networking are also the subject of international research. One of these is the Passive Optical Network. It provides very high information capacity for two-way communications. It works much like a two-way broadcasting system; every party on the network has the capability both to broadcast and to receive. This approach exploits the very high information capacity of optical fiber systems to reduce network complexity by simplifying switching and network protocols. It is based primarily on optical couplers called combiners and splitters and on optical amplifiers that amplify without converting light to electrical signals and vice versa.[51]

Specialized testing and performance analysis equipment is needed to evaluate hardware adapters and translators among networking protocols for linking video processors, computers, and supercomputers to long-haul networks that serve advanced multi-media systems. Such equipment will be necessary for handling the high data rates required for transferring full-motion, color images among video-supercomputers, other supercomputers, engineering work stations, personal computers, and advanced video stations or other sources in global networks.

The combining of different protocols and services for transmission over the same network is receiving much attention. For example, multimedia commercial standards, incorporating video services, are being developed. Two U.S. companies have announced standards for motion video in a limited field of a computer display.[52] This would enable motion video in one or more window-environments of an engineering workstation. A second example pertains to the development of video-conferencing products that run well with speed in the range of 56 to 128 kilobits per second on Narrowband ISDN. The International Consultative Committee for Telephone and Telegraph has issued video compression standards for video-conferencing, but many vendors of video-conferencing products have proprietary coding schemes that work better than the standards.

The measurement challenges for telecommunications networks are covered in detail in Chapter 7, Microwaves, beginning on page 147, and in Chapter 9, Optical-Fiber Communications, beginning on page 217.

High-Density Information Storage

To implement optical or magnetic information storage, advanced video systems will require major increases in the capacity of storage systems and in the rate at which data can be stored or played back. Capacity increases might be obtained with an increase in the amount of information stored per unit area (the information density) or with an increase in the area used for storage, or both.

Depending on the increases obtained in density, playback-rate increases might have to be achieved by operating the systems at higher mechanical speeds. For magnetic tape systems, this would mean higher rates of wear and more critical tolerances during manufacturing.

The relative importance of optical and magnetic storage systems for advanced video technology is not entirely clear at the moment. Magnetic tape is attractive because is permits major increases in area rather easily, by just using longer (if not also wider) tapes. Magnetic disks and optical disks do not lend themselves as easily to large increases in area.

Today's computer disks provide a reference point for understanding the limitations of disk-based magnetic and optical storage for systems employing digital encoding. The performance of both types of drives is changing so rapidly with new technological developments that the comparison presented here will be valid only momentarily. Today's single-platter high-capacity magnetic and optical storage systems for computers hold typically 1 to 10 gigabytes per platter. The use of automatic cartridge handlers gives much greater storage capacity. Optical media now generally store about 10 times more information in a given area, but the magnetic media have the higher theoretical storage density and thus may eventually surpass the optical-media density. Without data compression, 4 gigabytes of information would provide for only 32 seconds of uncompressed advanced video playback (at 1 gigabit per second), if that information could be played back fast enough. Unfortunately, play back rates for even the fastest magnetic disk drives are about 20 to 40 times too slow to achieve this rate. Playback rates for optical drives are even slower, by a factor of 4 typically and by a factor of 2 at best. Even with a high level of data compression, by a factor of 8 for example, only about 4 minutes of advanced video information could be accommodated on a 4-gigabyte disk. Further, playback rates would still be about 5 times too slow. Two other comparisons may be of interest. Optical drives have the advantage of removable media; in contrast, the magnetic drives are generally sealed because of their high sensitivity to contamination. The magnetic drives have faster access times (times to reach the desired program material on the disk) by a factor of 4 typically.[53] Clearly, the challenges facing disk-storage technology in service to advanced video technology are formidable.

Magnetic tape can store more data -- up to 165 gigabytes on the new 19-millimeter helical-scan tapes and up to 5 gigabytes on the 8-millimeter helical-scan tapes -- than optical disks, about 10 gigabytes on the 356 millimeter disks. The 19-millimeter tape systems are much more expensive than systems for other magnetic tape media. If random access is needed for say post-production work or editing, magnetic tape will have a much longer access time than a mounted disk. However, data in the 100-gigabyte or higher range on optical disks would require the use of automatic cartridge handlers that would give delays of several seconds for mounting the optical disks.[54]

The measurement challenges for magnetic information storage are covered in detail in Chapter 5, Magnetics, beginning on page 95. The measurement needs for optical information storage are not covered in this edition of this document.

New storage technologies that are only in the research phase now may emerge to serve advanced video technology. One approach may be to exploit the three-dimensional nature of matter to enable storing much more information than possible with the surface-based storage approaches in present use. Three-dimensional optical memories for digital video data could increase the capacity of today's laser disks by 100 times. Challenges in accomplishing this are to mass produce low-cost, reliable, high-power lasers and to develop new, more sensitive materials.

High-Resolution Displays

Among the challenges of bandwidth limitations, signal processing, and image capture, the performance and economic constraints of available displays are currently the greatest pacing factor for progress in advanced video systems.[55] Applications that are in the best position to prosper are those for which the constraints imposed by present displays are not significant.

Cathode-ray tubes can already be made with the performance required for advanced video technology, but they become more expensive and heavier in the larger screen sizes. Smaller cathode-ray tubes with higher intensity can be used in projection displays. Flat-panel displays are in use today, but at present their images are limited in size and quality in comparison to cathode-ray tubes. The flat-panel displays are also more expensive for a given screen size. However, because the flat-panel displays are a new technology, significant progress in improving performance and in reducing cost is expected.

Full-color flat-screen displays, made from liquid crystals and other organic materials, are available for computers and for miniature television applications. Major technological challenges for flat-screen displays include achieving adequate brightness levels, proper color rendering, and low cost. The displays will likely contain embedded electronic driving circuits throughout. Manufacturing flat-panel displays presents many of the problems of integrated circuits, but those problems are greatly exacerbated by the large sizes of the displays, which may be one to two square meters in area. Also, displays are complex structures. For example, liquid-crystal displays combine liquid crystals, phosphors, microelectronics, and optoelectronics.

The production yields of flat-panel displays are low and their costs are high. The technical problems in making flat-panel displays are formidable but solvable. All technologies for flat-panel displays have technical barriers that must be overcome. Most of the experience today is with liquid-crystal displays. The other technologies that compete with liquid-crystal displays for some markets are: plasma, field emission, electroluminescence, semiconductor laser-diode arrays, and semiconductor light-emitting diode arrays. Present yields for 8-centimeter, passive, liquid-crystal displays, such as those in portable televisions, are around 70 percent with 20-micrometer feature sizes. The yields for 25-centimeter active-matrix liquid-crystal displays for personal computers are thought to vary anywhere from a few percent to above 50 percent with feature sizes of 5 to 10 micrometers. These latter yields are beginning to rise.

Anticipated improvements in flat-panel displays include: greater resolution (especially for projection displays), larger size, wider viewing angles without display fade out, faster response times to input signals, greater contrast ratios between bright and dim portions of the display, and more levels of light intensity between the brightest and darkest regions of the display.

In the determination of quality, the human eye is the final arbitrator about processed and displayed video images. Knowledge of the response of the eye, its perception of resolution and the dependence of perceived resolution on color, geometry, object size, and solid-angle subtended at the eye by the display or its pixels are areas that will be essential in designing acceptable advanced video systems. Such attributes of the eye played a crucial role in the present National Television Standards Committee standards. These attributes of the eye are now playing essential roles in the standards proposed by the Society of Motion Picture and Television Engineers, the Joint Photographic Experts Group, and the Motion Picture Experts Group. Since the processing of the associated audio in the video images for many applications (education, teleconferencing, and cataloging) cannot be done independently of the video processing, audio-visual synchronization is being addressed by the Motion Picture Experts Group. The Joint Photographic Experts Group works mainly on still images. To support this effort more data are needed on the responses of the eye. Particularly important will be data on the distribution of responses that arise from variations in eye sensitivities.

The psychovisual characteristics of the eye are expected to influence video compression techniques to an equal or greater degree than the psychoacoustic characteristics of the ear have influenced the development of efficient techniques for recording audio. For example, because most people do not hear sounds much above 20,000 hertz, audio systems usually process only frequencies below 20,000 hertz. This in itself reduces the amount of audio data that must be processed. The common compact disk stores audio information in a digital form based on sampling the original audio source at a rate of about two times the upper frequency for hearing, namely, 44,000 hertz. A second example concerns surround-sound encoding for selected movies and television programs. Some approaches to surround sound encoding reduce the amount of audio information that is processed by increasing the number of effective audio channels on two-channel compact disks, video tapes, and films to more than four channels in some cases. Phase shifts, amplitude changes, and times delays of the order of 20 milliseconds are used to accomplish this. The time delay is based on a phenomenon known as the Haas effect in which the direction of sound is perceived according to what is heard first relative to the front, side, and rear channels. A third example is based on hearing dynamic thresholds of the human ear.[56] One can hear sound only above a certain amplitude or level that is called the hearing threshold. Also, loud sounds hide or mask softer sounds occurring in their vicinity and thereby dynamically change the threshold of hearing. These two aspects of human hearing enable engineers to code audio signals by computing the masking threshold and by transferring only those audio signal components that are above the dynamic masking threshold. Such adaptive encoding may decrease considerably the amount of audio data that are sent.

The major technical challenges in making flat-panel displays with profitable yields are shown in Table 145. Process-induced contamination and defects, such as those from particles, are believed to be a major reason for low yields,[57] so their elimination is paramount. For example, if a transistor controlling a pixel in an active-matrix liquid-crystal display is shorted by a defect, then the pixel is on all of the time, which produces a dark spot. Or if the defect produces an open circuit, then the pixel is off all of the time, which produces a white spot. Only a few defects of these two types are acceptable. Further, the front and back panels of a flat-panel display are subjected to different processing steps and therefore to different thermal histories, making their alignment and registration difficult and complicated undertakings. Also, there are many mechanical difficulties associated with the alignment of patterns and the controlling of pattern shifts during each processing step. Repair of defective flat-panel displays is important because the displays are so complex that improvements in process control may not be able to eliminate all defects. Other technical challenges exist but they are not as critical as the three just mentioned. Avoiding electrostatic damage to very large flat-panel displays requires extremely precise environmental control, such as humidity, and careful handling of processed materials. Avoiding ionic contamination and free radicals on the glass surfaces requires care in designing, building, and cleaning reactors and in fabricating glass panels. Such problems will be greatly reduced in significance by maintaining excellent clean-room conditions for all steps in the fabrication of flat-panel displays and are less of a problem for plasma and electroluminescent displays than they are for active-matrix liquid-crystal displays.[58]

These technical challenges give rise to a number of measurement challenges. To reduce defects, for example, real-time, contactless process monitoring is needed. Such monitoring enables control of the

Table 145
Technical Challenges for Flat-Panel Displays

process-induced contamination and defects
alignment and registration of front and back panels
repair of defect displays
electrostatic damage
ionic and free-radical contamination

manufacturing process by providing rapid feedback when defects are observed. Computers will be used to monitor, to simulate, and to control each process step. Monitoring demands real-time sensing of process status for fast feedback control of the gas-phase reactors used in many flat-panel display processing steps. Real-time means that the measurements and controls are completed in a time that is less than the time it takes for the process parameter being monitored to change. Simulating processes requires sophisticated computer programs to predict the deposition, doping, and patterning. Controlling processes requires the design of computer-controlled actuators and reactor valves. Chemical-vapor-deposition methods are dominant in fabricating flat-panel displays. The understanding of the physics and chemistry for reproducible control of gas-phase reactors will have to be done initially for specific process equipment in order to verify the process-control simulations and to demonstrate feasibility.

The alignment of the front and back panel also depends on measurement technology. Typical measurement tolerances for areas measuring 0.5 meters on a side are better than 0.5 micrometers, or of order one part per million. The processing equipment "looks" for index marks on a panel-by-panel basis. This increases the time for fabrication. The glass substrates change size each time they undergo a thermal cycle. Conventional glass substrates for flat-panel displays shrink ("compaction") with each thermal cycle. When yields are low, as they are today, the more expensive substrates, such as partially preshrunk glass, are too expensive. But, as yields increase, more expensive substrates are an option. Glass compaction will be a problem for poly-silicon thin-film-transistors and monocrystalline thin-film-transistors on glass.

Panel repair using both ablative and deposition methods requires new testing, inspection, and repair equipment for flat-panel displays. Because a flat-panel display contains from 800,000 to 3 million transistors, automated inspection is required. The repair of 100-centimeter flat-panel displays presents many technical challenges.

The automated inspection also stimulates the need for new ways to handle data. The handling of vast amounts of data for monitoring and controlling each of the several processing steps requires new approaches to "smart" manufacturing or intelligent processing of highly structured devices. Such data must be processed at very high speeds to achieve real-time control. Information sharing among the equipment used to overcome the first three challenges in Table 145 above will require standards for data formats and transfers. Data from detection of defects, from process analyses, from linewidth determinations, from test structures for process monitoring, and from patterning all must be interrelated. This will also be critical for replacing the present and costly process of "testing for reliability" with the much more efficient but more difficult process of "designing for reliability".[59]

New objective measurements and tests to evaluate completed displays and the effects of signal processing on the perceived images are also needed. These will include ways to measure uniformity

Table 146
Measurement Needs for Flat-Panel Displays

Needs Category	Measured Quantity
materials characterization	defect density resistivity flatness birefringence carrier mobility
manufacturing yield	defect density chemical identification layer thickness layer uniformity
display performance	pixel response time gray-scale linearity brightness contrast ratio field of view resolution for color uniformity resolution for intensity aliasing fixed-pattern noise artifacts color and brightness non-uniformity
interfacing to computers	figures of merit for resource allocation semiconductor memory CPU time

of luminosity of each pixel and gray levels. In the final analysis, the eye determines the merit of any flat-panel display. Understanding of human visual perception so that others can design high-performance display systems is a key challenge that has technical, psychological, and physiological aspects. Objectively evaluated sources to test display performance for scientific and engineering purposes, for medical and robotics imaging, for entertainment, and for teleconferencing will be required. The establishment of agreed upon methods for measurements related to human vision are a necessity for acceptance in the market place. Each user group will need different resolution, response time, frame rate, and number of gray levels. Aspects such as eye fatigue, flicker, resolution, color, artifacts caused by signal processing or display technology, and the like will have to be addressed.

The most important measurement needs cutting across the spectrum of candidate display technologies are summarized in Table 146. The measurement needs are organized in groups that focus, respectively, on materials, manufacturing, performance of displays, and interfacing of displays with computers. The entries in the table reflect a variety of measurement-related needs. These include needs for higher accuracy, for measurement reference standards, for accepted definitions, for accepted and standardized measurement methods, and for comparisons among measurement methods. Two

Table 147
Technical Concerns in the Manufacturing of Active-Matrix Liquid-Crystal Displays

Technical Concern	Impact		Measurement Needs/Issues
	Cost/Yield	Performance	
High Priority			
substrate particle inspection for uncoated, coated, and patterned layers	primary	secondary	defect density, chemical identification, and index of refraction
less expensive color filters with less absorption of light	primary	secondary	transmission efficiency
process materials development for active matrices	secondary	primary	fast in-situ measurements of process parameters such as deposition rates, thickness, and chemical stoichiometry; statistical process control; and reduced processing time
new liquid crystals for reflective mode operation (long-term and high-risk challenge)		primary	pixel response time, gray-scale linearity, viewing angle, resolution, birefringence, and layer thickness
Medium Priority			
substrate cleaning	primary		surface density of contaminants
resist and polyimide coatings	primary		speed of application, amount that is waste, layer uniformity, and thickness in real-time
alignment layer	secondary	primary	relative rubbing speeds, pressure, spacing between ridges, cross-section of ridges, and uniformity of these quantities across large areas
packaging	secondary	primary	crosstalk among drivers, speed with which rows and columns may be addressed, pitch, viscosity of liquid crystal with temperature, and thickness and size of substrates

examples follow. (1) *Standardization of measurement methods*: Industry has not agreed upon ways to specify the field of view and contrast ratio of flat-panel displays. Users are thus left to set up their own measurement and test facilities to verify that products meet specifications. This duplication is costly and cannot lead to uniform results. Factors such as measurement sensitivity to ambient conditions (temperature, illumination), which are important for these and other measured quantities, may not be systematically treated by the different approaches of the users. (2) *Development of*

Table 148
Technical Concerns in the Manufacturing of Plasma Displays

Technical Concern	Impact		Measurement Needs/Issues
	Cost/Yield	Performance	
High Priority			
substrate inspection	primary	secondary	defect density, chemical identification, index of refraction, shape, flatness, and warpage
phosphors		primary	efficiency, stability, match to color filters
packaging	primary	secondary	strength, electrical properties, leakage currents, lifetime, failure mechanisms
Medium Priority			
inspection of coatings	primary		defect density, chemical identification, index of refraction
driver chips and IC circuits coatings		primary	current-voltage characteristic, uniformity of threshold voltages, leakage currents, optical properties, reliability, thermal management
contact exposure tools	primary		reproducibility, throughput, number of reticle changes, down-time, resolution

accepted definitions: In the interfacing area, definitions are lacking for figures of merit for such fundamental factors as optimal allocation of resources, including semiconductor memory and central-processing-unit time, between the display and the supporting computer circuits.

Consensus building for making flat-panel displays[60] and for interfacing them with computers is occurring.[61] Present standards for interfacing computers and displays tend to be optimized for the physical characteristics of cathode-ray tubes and tend not to exploit the physical characteristics of other display technologies. The development of new interfacing approaches will benefit from measurements of the dynamic resource application among memory, buffers, the display drivers, etc.

As illustrative examples, the technical concerns expressed by some representatives of the U.S. display manufacturers at a meeting of the Display Working Group of the Photonics Materials Workshop are shown in Table 147, Table 148, and Table 149 for each of three display technologies.[62] The first column of each table shows the principal technical concerns, grouped by priority (either high or medium) for a given display technology. The second column of each table shows the relative impact of the technical concerns on the achievement of two goals (improvements in cost/yield, or improvements in performance). The third column of each table lists the measurement needs, in the form of measured quantities or measurement-related issues, that are inhibiting addressing the concerns.

Table 149
Technical Concerns in the Manufacturing of Electroluminescent Displays

Technical Concern	Impact		Measurement Needs/Issues
	Cost/Yield	Performance	
High Priority			
inspection of incoming materials		primary	defect density, chemical identification of any contaminants, and strength
blue phosphors	secondary	primary	efficiency, stability, match to other key parameters in display
drivers and integrated circuits		primary	current-voltage characteristics, uniformity of threshold voltages, leakage currents, optical properties, reliability, thermal management
Medium Priority			
packaging and chip-on-glass	primary	secondary	strength, electrical properties, leakage currents, lifetime, failure mechanisms, strength of polymer adhesives, pressure, temperature, temperature history
coating	primary		uniformity, defect density, repair-site identification and positioning for repair, green-state optical and electronic properties, thickness
contact proximity exposure tools	primary		reproducibility, throughput, number of reticle changes, down-time, resolution

The deliberation of the Display Working Group concerned manufacturing processes primarily for active-matrix liquid-crystal, plasma, and electroluminescent displays. The weighting of time that was devoted to the various topics reflected the interests and experiences of those who attended the Display Working Group. High on the lists for each display technology was the inspection of incoming materials and substrates. Such inspection would involve measurements to determine the density of defects, their chemical and physical properties, and, when appropriate, their optical and electronic properties. All of these measurements should be non-destructive and fast.

ENDNOTES

1. *Japanese Evaluation Technology Center Panel Report on Display Technologies in Japan*, National Technical Information Service Report Number PB92-100247, Table 1.5, p. 14 (1992).

2. "Proceedings of the 1990 International Conference on Consumer Electronics", *IEEE Transactions on Consumer Electronics*, Vol. 36, No. 3 (1991).

3. According to the item entitled "FPD Trends in Japan" in *Channel*, p. 19 (September 1992), the total number of parts required to make an active-matrix liquid-crystal display is about 50 times the number of parts required to make today's integrated circuits.

4. According to the *1991 TV and Cable Factbook*, as cited by T. Yamazaki, "The Impact of HDTV on Society", *Proceedings of the 12th International Display Research Conference, Japan Display 92,* p. 7 (Hiroshima, Japan: October 12-14, 1992).

5. William E. Glenn, "High Definition Television Display Research - Present and Future", *Proceedings of the 12th International Display Research Conference, Japan Display 92*, p. 17 (Hiroshima, Japan: October 12-14, 1992).

6. The superscripts in Table 137 refer to the notes that follow: (a) U.S. shipments of broadcast, studio and related electronic equipment were about $1.9 billion in 1990, according to the *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 31-2 (1991). For any product category, the U.S. market equals U.S. shipments minus the balance of trade. For the parent product category of "radio communications and detection equipment", the balance of trade is small relative to U.S. shipments for 1990 (about 3 percent). Thus, to a first approximation the U.S. market for broadcast, studio, and related electronic equipment, should be roughly equal to U.S. shipments. The world market for electronic products is often 2.0 to 2.5 times greater than the U.S. market, suggesting a world market of about $4 billion for broadcast, studio, and related electronic equipment. (b) U.S. shipments of communications systems and equipment, except broadcast, were about $15 billion in 1990, according to the *1992 U.S. Industrial Outlook*, p. 30-1 (1992). Since these shipments are part of the same parent product category of "radio communications and detection equipment", applying the reasoning in (a) above suggests a world market of about $35 billion. (c) U.S. shipments of telephone and telegraph equipment were about $14.7 billion in 1990 and the U.S. market (U.S. shipments - balance of trade) was about $16.5 billion in the same year. These figures come from the *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 29-2. Using the multiplier of 2.3 introduced in (a) above, suggests a world market of about $38 billion. (d) The world market for bipolar and MOS digital memory is about $18 billion in 1990, according to Table 26 on page 65 in Chapter 4, Semiconductors. The world consumption of application-specific integrated circuits is about $10 billion in 1990 according to Table 28, on page 67, in Chapter 4. (e) The world market for high-density magnetic information-storage equipment (rigid disk drives, digital tape drives, and video cassette recorders, and camcorders) in 1990 was $37 billion, according to Table 48 on page 108 in Chapter 5, Magnetics. (f) The world market for optical disk drives was $1.2 billion in 1990 according to *1990 Disk/Trend Report: Optical Disk Drives*, Disk/Trend, Inc., p. SUM-4 (July 1990). (g) According to "Technology Profile: Flat Panel Displays", SRI International, p. 27 (January 1992), the 1991 world market for flat-panel displays is $3.8 billion and is expected to rise to $6.5 billion in 1996 and $10 billion in 2001. These market data are to be compared with market projections by Japanese companies. Specifically, the report *Flat Panel Displays in Japan 1992* by TechSearch International, p. v (May 1992) contains translations from *Nikkei Business Publications*; Japanese companies such as Sharp, Toshiba, and NEC, predict the market to be somewhere between $4 billion and $8 billion in 1995 and between $8 billion and $15 billion in 2000. The world market for cathode-ray tube displays is about $6.9 billion according to 1989 data from the Japanese Ministry of International Trade and

Industry that were presented by A. Firester in his talk at the International Flat Panel Display Conference, San Francisco, June 17, 1992. Firester also reports that the 1989 markets for flat-panels displays (active and passive liquid-crystal, florescent, plasma, and electroluminescent) add up to about $3.6 billion. Averaging the above 1991 ($3.8 billion) and 1989 ($3.6 billion) world markets for flat-panel displays produces an estimate of $3.7 for intervening year, 1990, and adding this figure to the figure of $6.9 billion (1989, as an estimate of 1990) for cathode-ray tubes produces the estimate of $11 billion (1990) shown in Table 137.

7. An even more optimistic estimate of $28.5 billion per year is provided in "Source: Dataquest; Announcing High Definition Video Technology; The Collision Between Television and Computers", Dataquest Incorporated (San Jose, California), p. 1 (1989).

8. A gigabit is 1 billion bits.

9. A megabit is 1 million bits.

10. James Martin, "The Development of ISDN is Proceeding Worldwide", *PC Week*, p. 68 (April 10, 1989). Also, private communication, Shukri Wakid, NIST (June 1992).

11. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, pp. 30-12 to 30-14 (1992).

12. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, pp. 30-11 and 30-12 (1992). H. S. Bennett, "Summary Foreign Trip Report", Number 727-199, p. 9 (May 8, 1988).

13. Private communication, Bruce Field, NIST (December 2, 1992).

14. U.S. shipments of telephone and telegraph equipment was $15.2 billion in 1990. This figure is from the *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 29-2 (1992). If the U.S. trade deficit of $1.8 billion in 1990 (p. 28-2) is added to this figure, then the U.S. market can be estimated at $17 billion for 1990. To estimate world markets from the U.S. market, a multiplier of about 2.0 to 2.5 is usually about right for these market segments. If the value of 2.3 is used, then world market for these video technology related products would be about $39 billion.

15. See Table 48 on page 108 in Chapter 6, Magnetics.

16. See page 109 in Chapter 6, Magnetics.

17. *1992 Electronic Market Data Book*, Electronic Industries Association, Tables 4-6, pp. 77-79 (1992). These data do not state what percentage of the television picture tubes consumed in the U.S. are made in the U.S. by foreign companies. When commodity-like products are large and when their shipping is an appreciable fraction of their manufacturing costs, many companies choose to make such products in the countries in which they will be consumed.

18. "Electronics Evolution - Through a Fuzzy Crystal Ball", *Microelectronics Manufacturing Technology*, p. 16 (January 1992).

19. Lawrence E. Tannas, Jr. "Flat-Panel Displays Displace Large, Heavy, Power-Hungry CRTs", *IEEE Spectrum*, p. 34 (September 1989). Spencer Chin, "A Varied Brew of Display", *Electronic Products*, pp. 19-20 (October 1989).

20. "Semiconductor, Printed Circuit Boards, and Other Electronic Components", *Current Industrial Reports*, Bureau of the Census, U.S. Department of Commerce, Issue MA36Q(90)-1, p. 7 (October 1991).

21. Sony High-Definition Digital VTR System, Model HDD-1000.

22. "High-Definition Decision Hangs in the Balance", *Video Technology News*, Vol. 6, No. 4, page 1 (February 15, 1993).

23. Private communication, William Burr, NIST (December 14, 1992).

24. Dennis Roddy and John Coolen, *Electronic Communications: Third Edition*, p. 600 (1984).

25. Hitomi Murakami, Hideo Hashimito, and Yoshinori Hatori, "Quality of Band-Compressed TV Services", *IEEE Communications*, p. 65 (October 1988).

26. These error rates are those provided by the basic transmission hardware, prior to the execution of software programs that correct errors. These software programs are highly sophisticated and can reduce the errors to insignificance at the expense of additional computing time and slowed throughput.

27. Herbert S. Bennett, "Summary Foreign Trip Report", Number 727-199, p. 9, (May 8, 1988).

28. Edward E. Reinhart, "Mobile Communications", *IEEE Spectrum*, p. 28 (February 1992).

29. *1992 Electronic Market Data Book*, Electronic Industries Association, p. 49 (1992).

30. *High Definition Information Systems*, a report prepared by the Subcommittee on Technology and Competitiveness, transmitted to the Committee on Science, Space, and Technology of the House of Representatives, 102nd Congress, Second Session, p. 16 (July 1992). Note that it is difficult to compare the installations costs of copper and optical-fiber cables because the costs depend considerably on data rates, number of channels for communications, and whether the data are analog, digital, or both.

31. *1992 Electronic Market Data Book*, Electronic Industries Association, Table 2-7, p. 48 (1992).

32. *1990 Census of Population and Housing: U.S. Summary Social, Economic, and Housing Characteristics*, Bureau of the Census, U.S. Department of Commerce, to be published early 1993.

33. Private communication, Dana Grubb, NIST (December 15, 1992).

34. Sony HDD-1000 recorder, now a commercial product. Information provided by the Sony Corporation.

35. Private communication, Herbert Ohlandt, National Video Center (1992).

36. *JTEC Panel Report on Display Technologies in Japan*, Japan Technology Evaluation Center sponsored by the National Science Foundation, Defense Advanced Research Projects Agency, and the Air Force Systems Command, and coordinated by Loyola College (Maryland), p. 256 (June 1992).

37. Barry Fox, "Flat Screen Future for HDTV", *New Scientist*, Vol. 134, No. 1823, p. 18 (May 30, 1992).

38. Peter S. Friedman, Donald K. Wedding, and Ray A. Stoller, "Color AC Plasma Displays for High Definition Systems", Paper No. 1664-29, Society of Photo-Optical Engineers Symposium on Electronic Imaging Science and Technology (San Jose, CA: February 12, 1992).

39. Peter S. Friedman, Ray A. Stoller, and Donald K. Wedding, "An Improved Definition 19-Inch-Diagonal Full Color AC-PDP Video Monitor", paper 3 given at the Society for Information Display Symposium, (Boston, Massachusetts: May 1992); and Peter S. Friedman, Ray A. Stoller, and Donald K. Wedding, "An Analysis of Large-Area HDTV Display Technology: CRT, LCD, and PDP", *Proceedings of the Society for Information Display*, Vol. 32, Issue 2, p. 99 (1991).

40. Peter S. Friedman, Ray A. Stoller, and Donald K. Wedding, "An Improved Definition 19-Inch-Diagonal Full Color AC-PDP Video Monitor", paper 3 given at the Society for Information Display Symposium, (Boston, Massachusetts: May 1992).

41. "Digital Systems Information Exchange IV" (Washington, DC: May 1992) and "Digital Systems Information Exchange V" (Washington, DC: October 1992). These are one-day conferences.

42. "Digital Systems Information Exchange IV" (Washington, DC: May 1992) and "Digital Systems Information Exchange V" (Washington, DC: October 1992). These are one-day conferences.

43. *1992 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 31-2 (1992).

44. Many workers in this field also use the term "interoperable and extensive." An interoperable standard or protocol is one that can serve many of today's applications. An extensive standard or protocol is sufficiently broadly based that it will serve many of tomorrow's applications.

45. Private communication, John V. Weaver, Liberty Television (1992).

46. Private communication, Herbert Ohlandt, National Video Center (1992). In the catalog numbered C2919 and entitled "Sony Consumer Products for Business", printed August 1992 by Sony Corporation of America, Park Ridge, NJ, the following statement appears on page 46: "Minimum illumination for Camcorders: There is no industry-standardized procedure for testing minimum illumination; specifications should only be compared with models of the same manufacturer." *The Complete Crutchfield Winter/Spring 1993 Catalog*, printed by Crutchfield, Charlottesville, Virginia, contains on page 123 the statement: "Minimum Illumination - The minimum amount of light a camcorder needs to produce an acceptable picture is expressed in lux, a metric measure of the intensity of light. Each camcorder manufacturer has a slightly different technical definition for acceptable picture. Thus, lux ratings are comparable only between models from the same manufacturer. [A sunny day at noon is about 100,000 lux, candlelight from 8 inches away is less than 10 lux.]" Lux is a measurement of the intensity of light.

47. Private communication, William Glenn, Florida Atlantic University (1992).

48. In a "lossless" compression or decompression procedure, all information and data are preserved and processed. In a "lossy" compression or decompression procedure, some of the information and data are not processed and the amount and kind of data that are processed depend on the intended application.

49. Private communication, Bruce Field, NIST (March 1992).

50. Private communication, Glen Reitmeier, Sarnoff Research Center (February 10, 1993).

51. Private communication, William Burr, NIST (September 1992).

52. Private communication, William Burr, NIST (December 14, 1992).

53. See page 103 in Chapter 6, Magnetics and associated endnotes for information that supports the comparisons in this paragraph.

54. Private communication, Dana Grubb, NIST (December 15, 1992).

55. *Report of the Task Force on Digital Image Architecture*, Society of Motion Picture and Television Engineers, p. 34 (August 1992).

56. G. C. P. Lokhoff, "DCC-Digital Compact Cassette", *IEEE Transactions on Consumer Electronics*, Vol. 37, No. 3, pp. 702-706 (August 1991).

57. *Flat Panel Displays in Japan 1992*, TechSearch International, Inc., p. 78 (Austin, Texas: May 1992). Translated abstracts of articles in *Flat Panel Display 1992* from *Nikkei Business Publications*.

58. Private communication, Walter Worobey, Sandia National Laboratory (January 1993).

59. Private communication, Harry Schafft, NIST (May 1992).

60. Herbert S. Bennett, "Report of the Display Working Group of the Photonics Materials Workshop, August 26-27, 1992". This Workshop and the deliberations of its Display Working Group were held at NIST, Gaithersburg, Maryland on the dates shown. Also, Flat Panel Display Infrastructure Workshop, September 9-10, 1992. This Workshop was held by the American Display Consortium at the Marriott Hotel in Gaithersburg, Maryland.

61. Private communications, Daniel Stokesberry, NIST (August 1992).

62. Herbert S. Bennett, "Report of the Display Working Group of the Photonics Materials Workshop, August 26-27, 1992". This workshop and the deliberations of its Display Working Group were held at NIST, Gaithersburg, Maryland on the date shown. Because other display technologies, such as passive-matrix liquid-crystal, ferroelectric liquid-crystal, and field-emission technologies, were discussed only briefly by those contributing to the deliberations of the Display Working Group, tables showing technical concerns for these three technologies are not provided here.

CHAPTER 12

ELECTROMAGNETIC COMPATIBILITY

June 1992

Chapter 12

ELECTROMAGNETIC COMPATIBILITY

SUMMARY

Electromagnetic compatibility has become an important requirement for all electronic systems. To achieve electromagnetic compatibility, manufacturers of electronic systems generally perform various tests to ensure that their products have sufficient immunity to electromagnetic interference and that they do not create electromagnetic interference. In spite of these efforts, electromagnetic interference does occur frequently and typically impacts one or more of four important areas: economics and competitiveness, national security, health, and safety.

Increasing efforts are required to maintain electromagnetic compatibility because of several trends in our society and in the electronics industry. The number of emitters of electromagnetic fields, such as communication satellites, mobile telephones, radar systems, computers, and electric motors, is increasing rapidly. Many of these emitters are operating at higher frequencies and are filling the microwave spectrum. Some of these emitters employ very high power levels. Some of the unintentional emitters, such as computers, are using faster digital electronics and are becoming more efficient radiators of electromagnetic energy at higher frequencies. The increasing complexity of integrated circuits requires more external interconnections that provide additional points of entry for electromagnetic interference. The lower levels of operating power required by integrated circuits make them more susceptible to electromagnetic interference. The increasing use of nonmetal casings, such as plastics and composites, can reduce shielding resulting in increased emissions and reduced immunity for electronic devices. Further, sensitive electronic devices are being used increasingly to control more functions in large systems, such as automobiles and airplanes, where safety considerations are so important. Finally, concern is increasing that low frequency sources of electromagnetic fields, such as power lines, household wiring, and home appliances, may pose potential health hazards. Public concerns about power-frequency fields alone have given rise to costs approaching $1 billion per year.

In addition to the general issues above, key U.S. industries face electromagnetic compatibility problems that are specific to their products and systems. The aerospace industry, with U.S. shipments of $120 billion in 1990, faces a difficult challenge in measuring the immunity of aircraft electronic systems to high-intensity radiated fields produced by, for example, commercial broadcast stations and airport radar systems. The computer industry, with U.S. shipments of $71 billion in 1990, routinely performs emissions measurements, but will soon be required to perform immunity measurements in order to market computer equipment in the European Community. Such immunity measurements are not yet required in the U.S. European standards are already in place for video display terminals, and U.S. standards have been drafted. The automobile industry, with $243 billion of shipments of vehicles and parts in 1990, is steadily increasing the electronic content of its products; that content is expected to reach $2000 per vehicle in North America by 1995. This increase is likely to require more immunity testing because of safety and reliability concerns, and new European immunity standards are likely to make it more difficult for U.S. manufacturers to market in Europe. In 1990

U.S. shipments of medical and dental equipment were $28 billion. Electromagnetic compatibility of medical equipment is particularly important for hospital environments where electromagnetic interference can be life threatening. For communications and radar, with U.S. shipments of $60 billion in 1990, the rapid increase in the numbers of transmitters and sensitive receivers is leading to spectrum crowding and to increased potential for electromagnetic interference. The U.S. electric power industry, which delivered $176 billion of electricity in 1990, struggles to avoid surges and sags in delivered power, plus other irregularities in delivered power that can give rise to electromagnetic interference, in the face of a greater diversity of power sources and loads.

To maintain competitiveness in the electronics industry and in other industries that depend on electronic systems, U.S. industry must produce electronic systems that satisfy electromagnetic compatibility standards where those systems are to be marketed and used. It is equally important to achieve electromagnetic compatibility in a cost-effective and timely manner in order to maintain cost competitiveness and to avoid delays in marketing. Electromagnetic compatibility is achieved most efficiently by employing sound principles in the initial design stage, for example by using appropriate wiring geometries to reduce radiation and pickup at the circuit level and by employing components with high inherent immunity. It is also important to perform electromagnetic compatibility measurements throughout the design stage as well as in the final product stage.

The three groups of electromagnetic compatibility standards of most importance to U.S. industry are U.S. commercial standards, U.S. military standards, and international standards. Because better electromagnetic compatibility standards and measurement methods are needed, all three groups of standards are currently undergoing change. U.S. industry faces a formidable challenge in meeting these changing and multiple electromagnetic compatibility standards. Some of the standards groups are attempting to harmonize national and international standards to avoid multiple testing, but they face a difficult task in a rapidly changing electromagnetic environment. In many cases, the measurement capabilities required to meet present and proposed standards do not yet exist. The seriousness of the measurement problem was evident in the results of a survey of industry completed by NIST in 1991. While the survey addressed current and future problem areas in electromagnetic compatibility broadly, rather than measurement needs specifically, 83 percent of those responding cited the need for improved measurement capability.

Improved measurement capability is needed for three key applications: immunity testing, emissions testing, and characterization of the electromagnetic environment and related standards setting. A variety of generic improvements are needed. Measurement capabilities must be extended to cover the entire frequency range from 0 hertz (direct current) to microwave frequencies of at least 300 gigahertz. Present measurement capability is very limited above 1 gigahertz and almost non-existent above 20 gigahertz. A key application of this new measurement capability will be the characterization of the electromagnetic environment to support the setting of proper immunity standards. Test facilities must be developed that will accommodate larger test objects, such as commercial aircraft. Facilities for immunity testing must be developed to generate higher field strengths (above 200 volts per meter) and higher energy conducted disturbances to simulate an increasingly hostile electromagnetic environment. The accuracy and repeatability of electromagnetic compatibility measurements must be improved, particularly among different types of test facilities. Measurement methods for pulsed electromagnetic radiation must be developed to accommodate modern digital electronic systems. In particular, measurement methods using very short pulses (durations less than 1 nanosecond) with correspondingly broader bandwidths are required.

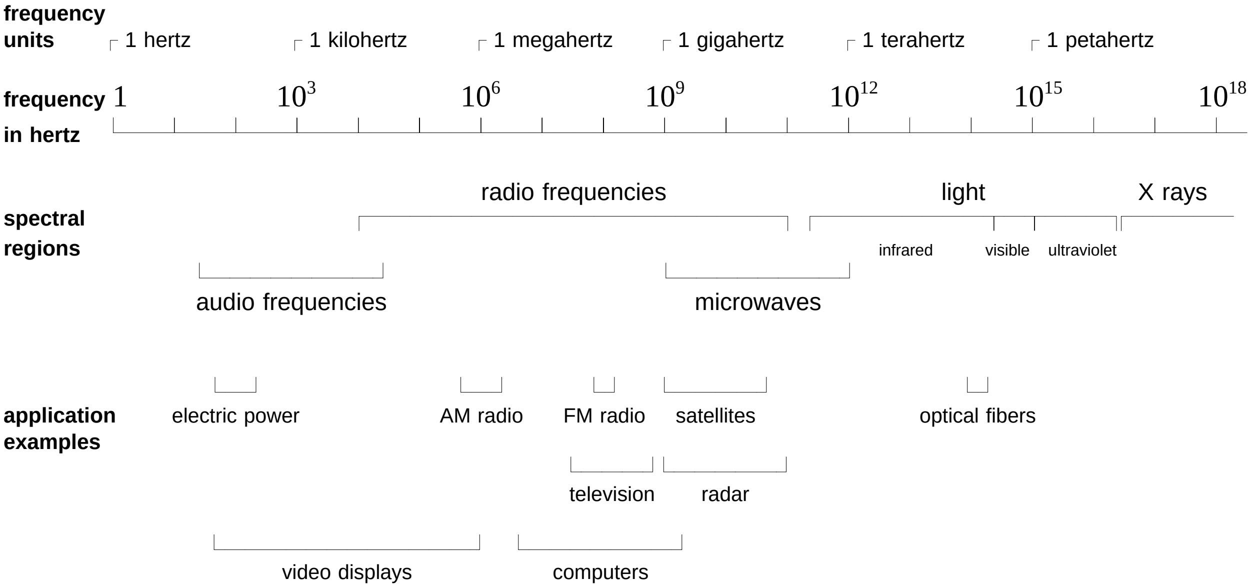
INTRODUCTION TO ELECTROMAGNETIC COMPATIBILITY

We live in a world that increasingly depends on electronic and electrical systems in virtually every arena of life. Throughout the world, thousands of engineers and technicians annually spend billions of dollars testing these systems for electromagnetic compatibility (EMC). This chapter examines why this massive effort is necessary, why EMC test methods are so numerous and diverse, and what progress in test and measurement methods is necessary to maintain EMC for present and future electronic and electrical systems.

Definitions

Electronic systems must satisfy two criteria to maintain electromagnetic compatibility: (1) they must operate satisfactorily in the electromagnetic environment in which they are to be used, and (2) they must not introduce intolerable disturbances in the normal operation of other equipment. The first criterion relates to *immunity*, and the second relates to *emissions*. When electromagnetic emissions are high, equipment without sufficient immunity may experience electromagnetic interference. Electromagnetic interference (EMI) is the disruption of the normal operation of electronic equipment or systems by an electromagnetic disturbance.[1]

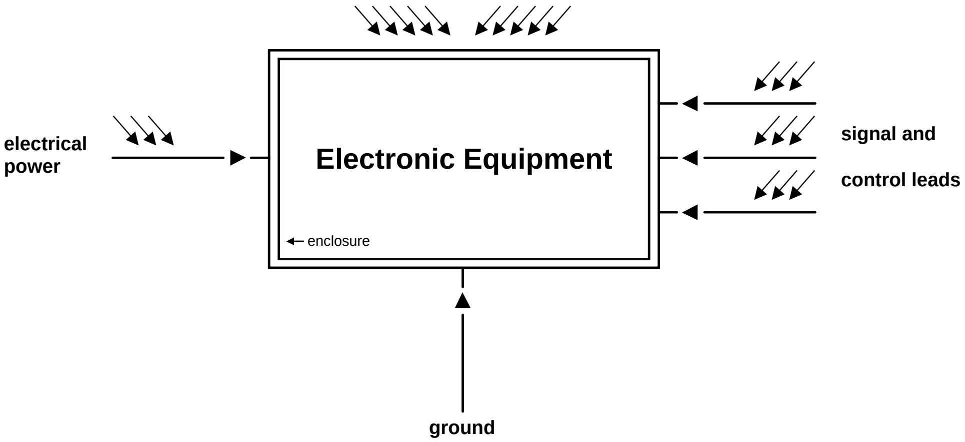
Figure 8
Overview of the Frequency Spectrum



An *electromagnetic wave* is characterized by variations of *electric* and *magnetic fields*. These variations occur over a wide range of frequencies illustrated in the overview in Figure 8. The units of measure of frequency are shown along the "ruler" at the top of the figure.[2] Key spectral regions are defined just under the "ruler" in terms of its divisions. They include audio frequencies, radio frequencies, microwaves, light (infrared, visible, and ultraviolet), and X rays. The frequencies at which familiar applications, such as electric power systems and television broadcast stations, operate are also shown. Finally, the wide frequency range of interest for EMC concerns is illustrated at the bottom of the figure. It extends all the way from 0 hertz (direct current), which is off-scale to the left of the entire figure, to the microwave frequencies used in satellite communications and radar. This wide frequency range of interest is one of the main reasons for the wide variety of tests and measurements that are required to maintain EMC in our society.

Unwanted electromagnetic energy may enter electronic equipment through several paths, as illustrated in Figure 9. The term *conducted interference* refers to interfering electromagnetic energy that enters electronic equipment through conducting leads but that *is not attributable* to picking up energy from electromagnetic fields in the immediate vicinity of the equipment. The conductor may be a power lead, signal lead, control lead, or ground lead. Conducted interference may be caused by disturbances in the supply of electrical power (such as switching transients) or by other disturbances on any of the conductors. The term *radiated interference* refers to interfering electromagnetic energy arriving through the air, whether from a nearby source or from a distant source. Electromagnetic fields from a nearby source generally have a more complex nature than those from a distant source.[3] Radiated interference includes the interference that enters electronic equipment directly through its enclosure or on conducting leads entering that enclosure if that interference *is attributable* to pick up near the enclosure. Radiated interference is indicated by the arrows (↘↘) in Figure 9. Alternatively, electromagnetic energy may be picked up by power, signal, or control leads at a greater distance from electronic equipment, and may be carried into the enclosure. Such energy is considered conducted interference because it is farther removed from its radiated origins by the time it enters the enclosure.

Figure 9
Paths of Entry of Electromagnetic Interference



This chapter addresses all of these types of interference except the subset of conducted interference that is attributable to disturbances in the supply of electrical power itself. Such disturbances are associated with the issue of *power quality* and will be addressed separately by a chapter on power networks in a future edition.

Motivation for EMC Concerns

EMI problems can be classified as impacting four general areas:

economics and competitiveness
national security
health
safety

Economic impact can result from several causes related to EMC. Products may fail to meet EMC regulations which can prevent market access. Products may be less competitive in the marketplace because of degraded performance. Products may be less competitive because of unnecessarily high costs resulting from overdesign to meet immunity requirements or from inefficient EMC testing. National security is affected when electronic systems fail to function properly in the severe electromagnetic environments often found on aircraft and ships or in battlefield conditions. Health hazards of electromagnetic fields have been studied for some time, and present[4] or developing standards cover the frequency range of 50-60 hertz[5] and 3 kilohertz to 300 gigahertz.[6] More recently, the concern about possible health hazards of lower frequencies has received considerable publicity.[7] Safety problems occur when electronic systems (such as aircraft altimeters) perform abnormally.

In practice most EMC problems impact more than one of the four major areas listed above. For example, concern about possible health hazards of electromagnetic fields from video display terminals can create an economic/competitiveness problem if the consumer decides not to purchase a particular model because of its electromagnetic fields. EMC concerns more generally can lead to a variety of costs. For example, public concerns about the health effects of power-frequency fields alone have given rise to estimated costs approaching $1 billion per year for "political, legal, and marketing reactions".[8] The question of whether low-frequency magnetic fields are a health hazard continues to be studied. In the U.S., for example, the Electric Power Research Institute has embarked on a major study to characterize these fields and their possible hazards.

In spite of past and present efforts to control EMI, a wide variety of types of EMI incidents have occurred. The following examples of incidents and broader concerns suggest this variety.

These examples are primarily related to economics and competitiveness:

> Personal computers have interfered with television reception.[9]
>
> Data, voice, or video information has been lost or degraded in telecommunication systems.[10]
>
> Electronically controlled automobiles have failed to start due to EMI.

These examples affect national security:

> The electromagnetic pulse from a nuclear burst (NEMP) can disrupt electronic and communication systems over a very large area.[11]

> Army helicopters have experienced malfunctions possibly linked to electromagnetic interference that resulted in loss of life.[12]
>
> A British warship had to turn off its protective radar in order to send a signal to the Admiralty in London. As a result, an undetected incoming Exocet missile destroyed the ship.

These examples are primarily health related:

> Electromagnetic fields from video display terminals[13] and from police radars have raised concerns about possible health hazards.
>
> Electromagnetic fields from 60-hertz power lines have also raised concerns about possible health hazards.[14]

These examples are primarily safety related:

> Electronic cash register systems have interfered with aeronautical instrumented landing systems at 113 megahertz and with police radio at 155 megahertz.[15]
>
> Communications with Air Force One (with the President aboard) caused garage-door openers to activate as the aircraft flew overhead.
>
> The inadvertent signals produced by relays in home appliances caused an electronic furnace control to demand full heat; a secondary failure in furnace operation burned down the house.
>
> Signals associated with on-the-spot television coverage of a Space Shuttle landing interfered with communications systems on the vehicle. The Shuttle gets only one chance to land.
>
> Patients with early-model heart pacemakers experienced traumatic disruption of pacemaker's operation because of EMI caused by microwave ovens.
>
> The driver of an eighteen-wheel tractor-trailer rig lost control when the trailer's electronically controlled brakes were prevented from operating as a result of citizens' band radio operation in nearby vehicles.

Current and Future Issues in Electromagnetic Compatibility

Current and future increases in the number of electronic devices in our society will lead to greater potential for EMI problems. Most electronic devices are both potential emitters of EMI and potential victims of EMI.

The numbers of both intentional and unintentional emitters of electromagnetic energy are increasing rapidly. Examples of intentional emitters are communication satellites, mobile telephones, and various radars. Many of these emitters are operating at higher frequencies and are filling the microwave spectrum. Very high power densities are produced in the main beams of highly directional microwave antennas. Large numbers of sensitive, low-power transmitters and receivers, such as mobile telephones, create a high potential for EMI. The increase in unintentional emitters is even greater as electronic systems are added to automobiles, airplanes, and other equipment. All computers are unintentional emitters, and faster clock rates make them even more efficient radiators of higher frequencies. Power lines, house wiring, and electric motors are other unintentional sources of potentially interfering fields. The increased use of non-metal enclosures, including plastics and ceramics, for electronic components and equipment can result in decreased *shielding* and more radiation. The result of all of these trends is that the electromagnetic environment sees more sources spread over a broader frequency range.

At the same time that the sources of potential interference are increasing in number, the potential victims are increasing in number, too. Many electronic systems are shrinking in size, becoming more complex, and using lower power levels, particularly as they rely increasingly on sophisticated integrated circuits. Two of these three changes -- the increasing complexity and the lower power levels -- make them more susceptible to EMI. Further, electronic devices are being used more frequently to control large systems, such as automobiles and airplanes, where safety considerations are so important. These large systems may acquire the vulnerability of their electronic controls to EMI. Further, the same nonmetal enclosures that allow emissions also make devices more susceptible to EMI because of their reduced shielding.

Because of the greater potential for EMI, increased EMC testing and compliance is being called for throughout the world. The European Community was expected to issue very stringent EMC standards in 1992, but the new standards have been delayed to January 1, 1995. These new standards will call for both emissions and immunity testing, whereas the U.S. generally requires only emissions testing. Thus the new European Community standards have the potential to affect sales of U.S. products in Europe and elsewhere. The new European Community standards are developed and issued by the European Commission in which the U.S. is not a member. The standards developed by international standards bodies may or may not be reflected in the new European Community standards.

General Types of EMC Testing

EMC testing is performed to determine whether the equipment under test is compatible with the *electromagnetic environment* in which it is designed to operate.[16] The two general requirements for maintaining EMC are:

(1) The equipment must be immune to having its operation degraded by levels of electromagnetic energy already present or permitted in its intended operating environment.

(2) The equipment must not emit levels of electromagnetic energy that could deteriorate the operation of electronic systems already present or permitted in its intended operating environment.

EMI immunity testing is performed to determine to what extent the equipment satisfies the first requirement. EMI emissions testing is performed to determine to what extent the equipment satisfies the second requirement. Frequently the term *susceptibility* is used as the complement to the term *immunity*. That is, the susceptibility of a system is its sensitivity to the electromagnetic energy that causes electromagnetic interference.

As noted earlier two basic routes exist for transmission of EMI into or out of electronic systems: *radiation* and *conduction*. Radiation occurs through free space, and conduction occurs through connecting wires or cables. For example, public power lines often provide a direct, uninterrupted conduction path for EMI to be transmitted between electronic systems. Because of the two distinct types of EMI transmission, EMI emissions tests are frequently divided into tests for radiated emissions and conducted emissions. Similarly, immunity tests address sensitivity to both radiated and conducted emissions.

An important and especially difficult aspect of EMC testing is determining and specifying the electromagnetic environment in which the equipment is to be operated. Predicting the nature of the sources and the sensitivities of the potential victims (both electronic and biological) in the intended operating environment is difficult at best and sometimes impossible with available knowledge. The terrestrial electromagnetic environment[17] is a result of natural sources and man-made sources. A good example of natural radio noise is broadband (wide range of frequencies) atmospheric noise which originates in thunderstorms throughout the world. Man-made sources are more numerous and varied and include intentional radiators, such as radio and television broadcast stations, radio navigation, and radar, as well as unintentional radiators, such as radio and television receivers (local-oscillator radiation), automobiles (ignition-system radiation), appliances with electric motors, fluorescent lamps, high-voltage power lines, and personal computers.

The limits or requirements set forth for EMC testing are established in part by regulatory agencies, such as the Federal Communications Commission (FCC). In addition to mandatory requirements set forth by regulatory agencies, there are a number of industrial voluntary EMC standards available for use. These voluntary standards are usually published by professional organizations and standards committees. EMC regulations and standards and the measurement methods required for making EMC tests will be covered in more detail in later sections.

EMC ISSUES FOR U.S. INDUSTRIES

U.S. industries that produce electronic equipment or other equipment that contains electronic components are being forced to pay more attention to EMC. These industries face the general issues of more emitters that cause EMI, more potential victims of EMI, and increased susceptibility of sensitive electronics. In addition, individual industries must solve additional EMI problems that are more specific to their products and systems.

Aerospace

The aerospace industry remains one of the leading contributors to the U.S. trade balance. In 1990 the trade surplus for the aerospace industry hit an all-time high of $28 billion, and in 1991 the surplus is estimated at $32 billion, mostly because of the growing exports of widebody jets. U.S. shipments in aerospace products were $120 billion in 1990.[18] Although the aerospace industry as a whole continues to grow, declining defense budgets both in the United States and in other developed countries have reduced purchases of military aircraft and missile systems from U.S. suppliers. This decline has been accompanied by a strong shift in demand from the military to the civilian sector.

The aircraft industry currently faces difficult measurement challenges for establishing EMC because of concerns regarding the susceptibility of aircraft electronic systems to *high intensity radiated fields (HIRFs)*. HIRFs are electromagnetic fields generated by powerful emitters, such as over-the-horizon and shipboard radars, airport surveillance radars, AM and FM radio stations, and VHF television transmitters. The increased concerns are a result of the following: (1) greater dependence on electronic systems for performing functions required for continued safe flight and landing; (2) reduced electromagnetic shielding afforded by the use of composite materials in place of metals; (3) increasing susceptibility of electronic components due to their greater sensitivity; (4) adverse experience with some aircraft; and (5) increasingly severe external HIRF environment due to an increase in the number and power of HIRF emitters.

To model the HIRF environment in the U.S., the Federal Aviation Administration (FAA) has employed databases that contain parameters for all transmitters approved to operate in the U.S. Similar efforts were undertaken by a number of European nations, and the data from these nations were combined with data from the U.S. to produce a composite multinational electromagnetic profile. The resulting HIRF envelope is a representation of electromagnetic field strength over a frequency range of 10 kilohertz to 40 gigahertz. Generally speaking, the HIRF envelope remains below 200 volts per meter for frequencies below 400 megahertz, but contains higher level fields at selected frequencies above 400 megahertz representing narrow radar beams. Study of the HIRF envelope is continuing.

The FAA has set interim EMC requirements for electronic equipment in new aircraft. That equipment must be shown through testing to withstand potentially interfering HIRF with a very high electric-field strength of up to 200 volts per meter. The FAA's criteria are designed to ensure that if an aircraft flies close to an emitter, HIRF effects will not jeopardize normal operation of on-board electronic equipment critical to the safety of the flight. A subcommittee of the Society of Automotive Engineers (SAE) is working on a certification standard for operation in the HIRF environment.[19] Whatever certification standard the FAA chooses to adopt will require difficult EMC immunity measurements because of the high field strengths, large aircraft test volumes, and wide frequency ranges required. Several of the respondents to a recent NIST survey of industry problems with EMC stated that they have no capability to perform HIRF tests and that new facilities and measurement techniques are required. There have already been several cases of new aircraft that have been required to undergo additional HIRF testing resulting in costly delays in marketing.[20] The NIST industry survey is discussed in detail on page 394.

Computer and Peripherals

In 1990 U.S. shipments of computer and peripheral equipment were $71 billion.[21] The world market in computer equipment in 1990 was $151 billion. The U.S. represented more than one third of this total market, and the combined Western European market was approximately one third. The Japanese market was about one fifth ($29 billion).

The computer industry does a great deal of EMC testing. For some years now, the Federal Communications Commission has required the testing of electromagnetic emissions from most new electronic data processing equipment, as described in Part 15 of the FCC Regulations, principally to ensure non-interference with communications equipment.[22] Currently, the FCC is working through the American National Standards Institute (ANSI) to revise the emissions test procedures. The ANSI C63 Electromagnetic Compatibility Committee has produced a new standard, ANSI C63.4-1991. This new standard is generally thought to be a substantial improvement over prior versions and over prior FCC measurement proposals.[23] The test site for this standard must be an open site or an absorber-lined chamber that is equivalent to an open site (because of its lack of internal reflections) over the test frequency range of 30 megahertz to 1 gigahertz. Absorber-lined chambers normally perform well in the upper part of this frequency range, but not in the lower part down to 30 megahertz where absorbers become ineffective. Many of the respondents to the NIST industry survey stated that the low-frequency range was a major area of concern to them and that better absorber and more accurate measurement methods were needed in this frequency range.

Another area of concern to many respondents to the NIST industry survey is that the new EMC standards in the European Community will require additional EMC testing beyond that already required in the U.S. The European Community market in computer equipment is particularly important since it is as large as the U.S. market. The European Community directive on EMC applies to all equipment which is likely to "cause electromagnetic disturbance" or whose performance "is liable to be affected by such disturbance".[24] The directive requires that equipment be designed so as not to cause undue interference to radio and telecommunications equipment and, at the same time, to have an immunity level adequate to ensure that it will operate as intended. The details of what constitutes adequate performance are not contained in the directive itself. Rather, the European Community Commission will publish standards for compliance. Although the details of these standards are not yet known, at least some new immunity testing will be required for computer manufacturers since immunity testing is not required in the U.S. by the FCC. In effect the European standards are transferring the responsibility for immunity testing from the buyer to the seller. In this sense European buyers may receive more service than U.S. buyers when purchasing the products of U.S. manufacturers.

The Swedish National Board for Metrology and Testing has developed guidelines for acceptable levels of low-frequency fields from video display terminals, and for the methods for measuring those fields. The U.S. and the European Community are studying these guidelines,[25] and the U.S. has prepared a draft guideline already.[26] Measurements of display fields now appear in product reviews, and claims of low field levels are increasingly common in product advertising.

Motor Vehicles

The motor vehicles and parts industry is one of the largest sectors in the U.S. economy, with shipments of $243 billion of vehicles and parts in 1990. The motor vehicle industry accounted for 4.1 percent of the nation’s $5.2 trillion GNP in 1989. The U.S. recorded a negative automotive trade balance of $51.4 billion in 1989, 80 percent of which was due to trade with Japan and Canada. "Exports of both vehicles and parts are expected to increase measurably over the next few years, while the rate of growth in imports may be slowed by more competitive U.S. products and by additional local production by foreign manufacturers."[27]

The number of electronic features in U.S. vehicles is increasing rapidly, and this increase is causing the U.S. automotive industry to do an increasing number of EMC tests requiring new measurement techniques. It has been predicted that by 1995 the electronic content in automobiles in North America will reach $2000 per vehicle, compared to $395 in 1987. Although some of this increase will be in audio equipment, the majority will be in safety and security products.[28]

The use of plastic parts in the automotive industry has increased over the past 5 years, due to their lighter weight, lower production cost, and resistance to rusting and denting. Federal corporate average fuel economy requirements have encouraged original-equipment manufacturers to use plastics to reduce vehicle weight, which helps increase fuel efficiency. Plastics, ceramics, and composites will be used increasingly in the 1990s.[29] The use of these materials increases the need for EMC tests and measurements because of their reduced electromagnetic shielding effectiveness compared to metals. Improvements have been made in measuring the shielding effectiveness of these materials, but further improvements in accuracy and frequency range are needed.

The U.S. automotive industry has no EMC regulations, but has followed the voluntary standards developed by the Society of Automotive Engineers (SAE). The SAE J551 Standard[30] covers electromagnetic emissions from vehicles over the frequency range of 30 to 1000 megahertz. The SAE J1113 Standard[31] covers EMC exposure methods, measurements methods, and suggested levels for immunity. Immunity testing is becoming more important because electromagnetic interference to some new automobile systems, such as an anti-lock braking systems or fuel metering systems, can result in unsafe operation of the vehicle. Both of the SAE standards are under revision because of the changing EMC needs of the industry.

The International Special Committee on Radio Interference (CISPR) has a standard on vehicle emissions that is fairly consistent with SAE J551. However, there is a concern in the U.S. automotive industry that the European Automobile Constructors (ACEA) will develop new EMC standards for use by the European Community as part of its new European Community standards rather than adopting the CISPR standards. New European Community standards could exclude U.S. automobile manufacturers from the European Community market or at least make competition difficult.

Medical Equipment

In 1990 U.S. shipments of medical and dental equipment were $28 billion. International trade is playing an increasingly important role in medical equipment sales, and the U.S. trade surplus was $2.4 billion in 1990.[32] U.S. firms active in medical electronics have a very high research-and-development investment rate, but are facing strong foreign competition.

U.S. manufacturers of medical equipment must adhere to FCC regulations that limit electromagnetic emissions. The Food and Drug Administration (FDA) has a voluntary EMC standard, MDS-201-004, that covers both emissions and immunity.[33] The emissions limits are of critical importance in preventing EMI in hospitals and in other medical environments where it can be directly life threatening.[34] The immunity standard is equally important for ensuring that medical equipment will function properly in the expected electromagnetic environment. The current FDA immunity standard covers frequencies up to 1 gigahertz, but new standards and measurement methods are needed for higher frequencies because of the increasing use of microwaves in our society. Even though conformance to the FDA standard is voluntary, compliance becomes mandatory when the standard is referenced by Federal, state, or local regulatory agencies.

The FDA is currently working with the International Electrotechnical Commission (IEC) in an effort to harmonize U.S. and international standards. The European Community is developing directives that will determine the EMC requirements for medical devices for that community, and it is not known whether these requirements will be harmonized with IEC standards. Consequently, this is another area of concern for U.S. manufacturers regarding possible trade barriers of the European Community. The Western European market for medical devices is a very important one, accounting for 40 percent of U.S. exports of medical devices in 1990.[35]

Consumer Electronics

U.S. shipments of consumer electronics, excluding home computers, were $6.7 billion in 1990. U.S. exports of consumer electronics rose to $1.8 billion in 1990, while imports fell to $13.1, still leaving a consumer-electronics deficit of $11.3 billion.[36] High-definition television (HDTV) is considered by many to be the next major achievement in video technology and will be subject to the same type of intense international competition evident in the above balance of trade.

U.S. manufacturers of consumer electronics must adhere to FCC regulations to limit electromagnetic emissions, but EMC immunity standards are voluntary. However, a consensus has been reached among ANSI, SAE, and IEC that electronic equipment should withstand an upper limit of 200 volts per meter.[37] This high field level is difficult to reach in a measurement facility at many frequencies. Further, reaching this level over the entire desired frequency range of 10 kilohertz to 1 gigahertz will require several different measurement facilities. Further development of standards by standards bodies and further improvements in measurements capabilities are needed in this area.

Another area of increasing importance is *electrostatic discharge (ESD)*.[38] The most common example of electrostatic discharge is the spark that occurs when one walks across a carpeted floor and reaches for a metal door knob. Modern electronic products are more vulnerable to damage or interference from ESD because of plastic casings and smaller, more sensitive electronic components. Improved, well-characterized exposure techniques are needed for ESD immunity testing. Some standards work is underway and growing. The Electrostatic Overstress/Electric Discharge (EOS/ESD) Association has applied for membership in ANSI, and its Standards Committee is seeking accreditation as an ANSI standards producing organization. The approach of the European Community has generated a demand for a common set of EOS/ESD standards for European nations, and a European draft standard has been released for industry comment.

Communications and Radar

If navigation systems are included with communication and radar equipment, U.S. shipments in 1990 were $60 billion. More than 60 percent of these shipments were in the general area of radar and sonar systems, navigation systems, reconnaissance and surveillance systems, and electronic warfare equipment. Traditionally, this market has been dominated by government and defense applications, but in 1990 demand in this market stagnated due to budgetary pressures and the changing international political climate. Increasing civilian and international demand helped compensate. The communications area makes up 27 percent of this industry and includes fiber optic systems, microwave and satellite communication systems, paging equipment, and mobile communications systems. The most rapid growth is occurring in cellular, mobile, and satellite communications.[39]

Important technical advances have occurred in advanced phased array radars and in laser radar. Defense communication systems are being upgraded, and numerous EMC problems need to be solved. The military EMC standards[40] are extensive and more demanding than civilian standards because military equipment is expected to function in a more severe operating environment. Many of the older military EMC standards and test methods have been shown to be inadequate, and new standards and measurement methods are currently being developed.

The evolution of digital cellular communications is important for the expansion of the civilian market. Personal communication networks using "microcells" about 200 meters in radius are being proposed, and the FCC is receiving a continuing flow of requests to establish such a system. One innovative satellite-cellular telephone system proposed for 1996 would include as many as 77 satellites in low earth orbit to provide digital switching for worldwide mobile communications.[41] Direct broadcast service for television is already widely available in Japan and parts of Europe, and it is in the planning stages in the U.S. All of these new systems raise new spectrum and EMC issues that must be resolved.

Because of spectrum crowding, satellite and short-haul systems are moving to higher microwave frequencies above 11 gigahertz. The FCC has granted nine U.S. companies conditional permits for high-power direct broadcast satellite (DBS) services at frequencies between 12.2 and 12.7 gigahertz dedicated solely for DBS. Overcrowding of the most popular frequencies, 2.4 to 6 gigahertz, by common carriers has increased interest in frequencies as high as 38 to 60 gigahertz. The use of these higher frequencies will extend the need for EMC test facilities and measurements into these same frequency bands. An unresolved question is whether the strong shift to digital systems will require new types of EMC tests beyond the traditional continuous-wave tests. Some of the respondents to the NIST industry survey suggested the need for further research in the use of pulsed signals for EMC testing.

GOALS OF U.S. INDUSTRY FOR COMPETITIVENESS

U.S. industry can enhance its competitiveness in the electronics industry and in other industries that depend on electronic systems by achieving key goals in EMC. The primary goals are to produce electronic systems that perform their intended function in the buyer’s electromagnetic environment and to gain access to the buyer’s market. To accomplish these goals, electronic products must satisfy the emissions and immunity standards where they are to be used. Ideally, products should have sufficiently low emissions and sufficiently high immunity to operate successfully in a wide range of environments and applications. Further, emissions and immunity standards must be satisfied in a cost-effective and timely manner in order to maintain cost competitiveness and to avoid delays in marketing. Well conceived standards can help make this possible.

An important approach to achieving EMC efficiently is to employ sound EMC principles[42] in the initial design stage of electronic systems, that is, to design EMC into a product rather than to pursue EMC on a remedial basis. This approach typically involves using appropriate wiring geometries to reduce radiation (emissions) and to reduce pickup (improve immunity) at the circuit level. Computers and other digital devices are now generating higher clock frequencies that are requiring even more attention to reducing radiation at the circuit level.[43] Some of the benefits of employing EMC principles in the initial design at the circuit or component level are these: (1) cost savings by avoiding later design modifications, (2) time savings by avoiding multiple EMC compliance tests, (3) cost and weight savings by reducing the need for extra shielding (particularly metal), and (4) better product performance by reducing EMI effects within the total system. Several respondents to the NIST industry survey stated that EMC design and education are important goals for U.S. industry to pursue.

EMC test failures can usually be avoided by considering EMC in early design stages and by performing preliminary EMC measurements throughout the design and construction process. Some of these preliminary measurements are simple and can be performed without elaborate test facilities.

For example, measurements of currents on wiring when combined with calculations can lead to good estimates of the emissions or immunity of the final product.[44]

Pursuit of EMC goals is dependent on the use of a wide variety of EMC measurements both during the design stage and during final product testing. There are currently major deficiencies in EMC measurement capability, and the changing electromagnetic environment will call for even better EMC measurement capability in the future.

EVIDENCE OF MEASUREMENT NEEDS

NIST Industry Survey

In 1991 NIST completed a survey of U.S. industry in order to collect information on the impact of measurement capability provided by NIST and to identify current and future problem areas in EMC. A five-page questionnaire was sent to approximately 200 individuals currently working in EMC for a wide variety of U.S. companies. The questionnaire was not directed specifically at new measurement needs, but 39 of the 47 questionnaires that were returned listed the need for improved EMC measurement techniques.

One of the most frequently mentioned needs was for the capability to perform immunity measurements for large systems over wide frequency ranges at high field levels. More specific needs for measurement methods and measurement facilities for HIRF testing of aircraft were also frequently mentioned.

In addition to wideband EMC measurement needs, the need for measurements capable of characterizing electromagnetic pulses in the time domain directly was frequently stated. The main reason given was that most modern electronics are digital and susceptible to pulses. NIST involvement in pulse EMC measurements was specifically requested.

The need for improved *repeatability* in EMC measurements both within and especially between organizations was frequently stated. Part of the problem is due to the different types of sites and facilities that are used in EMC measurements. The desirability of being able to make repeatable EMC measurements in indoor anechoic chambers was frequently stated.

A number of comments had to do with more basic EMC research in the initial design stage and at the component level, particularly emission from and susceptibility of microprocessors. Other comments called for more basic research on electrostatic discharge and on shielding effectiveness measurements of new materials, such as composites. Shielding effectiveness measurements of cable shields and connectors was another area where improvement was considered important.[45]

Even in areas where many of the technical measurement problems are solved, the variations in EMC test regulations throughout the world can increase manufacturing costs unnecessarily if the differences are based on artificial differences in intended operating environments. Agreement among Japan, Europe, and the U.S. on EMC standards was frequently asked for so that unnecessary multiple testing to different standards could be avoided. The same concerns dominated the comments from U.S. industry at a recent EMC workshop sponsored by NIST.[46]

EMC Publications

The numerous journals, symposia, and recent books on EMC measurements are additional evidence of interest in EMC measurements. The increase in EMC publications was explained in a recent paper: "Although there has been much activity in EMC for many years, there has been a recent explosion of interest in this field as evidenced by the number of conferences and publications devoted to the subject. There are three main reasons for this: (1) the extensive use of electronic devices for control, information processing, and other functions, (2) the extensive use of semiconductors which have relatively low immunity thresholds in those devices, and (3) the fact that devices are capable of emitting and responding to higher and higher frequencies."[47]

Many of the papers in the Institute of Electronic and Electrical Engineers (IEEE) Electromagnetic Compatibility Transactions and Technical Symposia are devoted to new EMC measurement techniques. In addition, there are several trade journals[48] published on EMC measurements. The measurement topics covered in these publications are primarily the same ones that respondents to the NIST survey cited as needing attention.

EMC Standards

The three groups of EMC standards of most importance to U.S. industry are: (1) U.S. commercial standards,[49] (2) U.S. military standards,[50] and (3) international standards.[51] The fact that standards in all three areas are currently undergoing revision reflects both the need for new EMC standards to accommodate the changes in the electromagnetic environment and the difficulty of implementing some existing standards with present measurement capability.

U.S. Commercial Standards

Part 15 of the FCC Rules and Regulations sets emissions standards, and Subpart B applies specifically to computing devices. In an effort to improve the current standards and test methods, the FCC is in the process of replacing its current test methods and measurement requirements with the American National Standards Institute’s ANSI C63.4.[52] This new standard will have the advantage of being compatible with an international measurement standard published by CISPR, but will be more detailed. This new standard will allow the use of absorber-lined chambers as well as open sites, but the problems of showing equivalence and repeatability for measurements in the two types of facilities will have to be solved.

The new FAA requirements for aircraft immunity to HIRF were mentioned earlier, and these requirements present measurement challenges relating to large volumes, high field strengths, and wide bandwidths that have not yet been resolved.

U.S. interest in standards for magnetic fields emitted by video display terminals is growing. The Institute of Electrical and Electronics Engineers has already drafted a standard for the measurement of such fields at frequencies from 5 hertz to 400 kilohertz.[53]

Many voluntary industry standards have been useful in the U.S., but these may not be adequate for future marketing in foreign countries.

U.S. Military Standards

The most important U.S. military standards are MIL-STD-461 which specifies both emissions and immunity limits and MIL-STD-462 which identifies the measurement methods. These standards address the vast majority of military equipment. They are currently being rewritten[54] because of problems with repeatability, lack of correlation of laboratory measurements with platform measurements, incompatibilities between the procurement system and the standards, and technical deficiencies in some areas. Even though some of these problems are to be corrected in the revisions, there will still be measurement difficulties in the areas of repeatability, high power levels, and large test volumes.

International Standards

International EMC standards from international organizations such as CISPR and IEC are extensive, but the forthcoming European EMC standards[55] for the European Community are of particular importance to U.S. industry. The details of these standards are not yet known, but they are expected to include requirements in several areas (immunity, ESD, and transients) not included in U.S. standards. This means that new measurements methods and facilities will be required by U.S. companies who market in Europe. The potential problem of having to use European laboratories for final compliance testing is a concern for U.S. companies unless a new agreement can be reached. The concern of U.S. manufacturers about European EMC standards was evident at the recent EMC workshop for industry and Government participants, sponsored by NIST.[56] The workshop participants urged the U.S. Government to play a strong role in representing U.S. interests in negotiations with the European Community and in assisting U.S. manufacturers in satisfying directives of the European Community. In particular NIST's assistance in disseminating information relevant to EMC was solicited.

For human exposure, the standards for electromagnetic fields from video displays terminals from the Swedish National Board for Metrology and Testing for magnetic fields have been noted above.[57] Those standards are cited with increasing frequency. Also, the first interim international standard for exposure to electric and magnetic fields at 50 to 60 hertz has been issued.[58] This was also referenced earlier. The U.S. has no standard for such exposure at this time.

MEASUREMENT NEEDS

New and improved measurement capability is needed to meet the challenge of maintaining EMC in an electromagnetic environment containing more emitters over an expanding frequency range and more potential victims with increasing susceptibility. The time is approaching when virtually all electronic and electrical devices will have to be tested thoroughly for emissions and immunity.

Applications of EMC Measurements

EMC measurement capability serves three principal categories of application: (1) immunity measurements, (2) emissions measurements, and (3) characterization of the electromagnetic environment. Immunity and emissions measurements are required for demonstrating compliance with EMC regulations and standards. Measurements of the electromagnetic environment are required so that EMC standards can be set at the appropriate field-strength levels. For example, the FAA used a combination of measurements and mathematical modeling to determine the HIRF environment in order to set immunity standards for aircraft.

Generic Improvements Needed

The principal improvements needed in measurement capability for the three different categories of application are outlined in Table 150. Some of the improvements are needed now while others will be needed within the next ten years.

Table 150
Improvements Needed in EMC Measurement Capability

Improvement	Specifics
broader frequency coverage	up to 300 gigahertz
larger test volumes	aircraft size
higher field strengths	up to 200 volts per meter
higher accuracy	±1 decibel
improved repeatability	±1 decibel
pulse measurements	less than 1 nanosecond duration
improved spatial resolution	less than 1 micrometer

Here are several important examples of the improvements needed. The FAA currently requires aircraft immunity measurements for frequencies up to 40 gigahertz. Measurements at higher frequencies will be needed in the next ten years because radar and communication systems will make increasing use of those frequencies. Present measurement capability is very limited above 1 gigahertz and almost non-existent above 20 gigahertz. Further, because of the complexity of modern systems employing electronics, it is already necessary to test entire systems (such as aircraft) at once, rather than testing components separately. This requires larger test facilities. To cope with HIRF, new measurement capability is needed now for electric fields with strengths up to 200 volts per meter for the commercial environment; for military environments, new measurement capability above 200 volts per meter will be needed within the next ten years. [Although electromagnetic waves consist of both electric and magnetic fields, *the field strength* is usually specified in terms of the electric field in volts per meter.] Improved accuracy and repeatability to a level of about ±1 decibel[59] (+26, -21 percent) will be needed within the next ten years, but improvement in site-to-site repeatability is particularly important and is needed now. Currently, some EMC measurements give variations as great as 30 decibels (a factor of 1000) from one site to another, and variations of ±5 decibels (+216, -68 percent) are typical. Improved accuracy and repeatability are needed for EMC measurements at both the component level (filters, gaskets, etc.) and the system level. It is more difficult to attain high accuracy for system measurements where the goal of ±1 decibel is very ambitious. Because so many electronic systems are now *digital*, pulse EMC measurements are needed now for pulse widths of 1 nanosecond or less. Presently, pulse measurements are limited to broader pulses which have less high-frequency content. Finally, there is emerging interest in measurements for detecting interfering currents and fields with a spatial resolution less than 1 micrometer. Such a capability would support the study of interference effects within integrated circuits and within biological systems where the response of individual cells would be examined.

Types of EMC Measurement Capability

To serve the above applications, seven types of measurement-related capability are needed. They are listed in Table 151 along with the EMC applications they support. The seven types are these: (1) special antennas for the measurement of electromagnetic fields for emissions testing and environment characterization; (2) special antennas and facilities for generating fields needed for making immunity measurements (and for supporting the development of the antennas for measuring electromagnetic fields); (3) special facilities for determining the shielding properties of materials, components (e.g., connectors and coaxial cable), and enclosures for electronic equipment; (4) measurement methods for employing the preceding measurement tools in a meaningful way; (5) measurement reference standards to support the accuracy of the measurement methods; (6) measured reference data on the shielding effectiveness of materials; and (7) mathematical models to characterize complex fields from a limited number of actual measurements. Note that the antennas and facilities for field measurement and generation support all three applications. This is true because the development of either measurement or generation capabilities requires the other. Each of the types of measurement-related capability is discussed below for the applications to which it applies. Because the facilities are so important to all three of the applications, an overview of those facilities is provided first.

Table 151
Types of Measurement Capability Supportive of the Three EMC Applications

	Supported EMC Applications		
Types of Measurement Capability	Immunity Measurements	Emissions Measurements	Environment Characterization
measurement antennas and facilities			
field measurement antennas	✓	✓	✓
field generation antennas and facilities	✓	✓	✓
shielding properties facilities	✓	✓	
measurement methods	✓	✓	✓
measurement reference standards	✓	✓	✓
reference data on materials shielding properties	✓	✓	
mathematical models for characterizing complex fields			✓

Supporting Facilities for EMC Measurements

Since electromagnetic radiation travels and is reflected, absorbed, or transmitted by nearby objects, repeatable measurements require a controlled environment. Measurements made to support emissions testing, immunity testing, and environment characterization employ a variety of special facilities that provide this controlled environment. They are summarized in Table 152, along with the applications they support, the frequency ranges they serve, and their relative costs. The cost depends greatly on the size and details of the particular facility. For Table 152 the cost ranges are defined as follows: *low* (below $100,000), *moderate* ($100,000 to $1,000,000), and *high* (above $1,000,000). To put the facilities into service, special techniques are required to create fields within them and to measure fields within them.

Table 152
Characteristics of EMC Facilities

Facility	Applications Supported	Frequency Range	Cost
TEM cell	immunity, emissions, environment	below 100 megahertz	low
open site	emissions, environment	30 kilohertz to 3 gigahertz	moderate
anechoic chamber	immunity, emissions, environment	above 200 megahertz	high
reverberation chamber	immunity, emissions	above 200 megahertz	moderate

TEM Cells and Anechoic Chambers

TEM cells, or transverse electromagnetic cells, are enclosed metallic chambers that can be connected directly to metallic waveguides that provide the needed incoming electromagnetic energy. The TEM cells preserve the stable pattern of the fields arriving in the guides. An anechoic chamber is also an enclosed metallic chamber, but its inner wall surfaces are lined with special materials that absorb electromagnetic radiation so that no reflections occur. Anechoic chambers simulate free space by assuring that no reflected radiation returns to the object being tested. Unlike TEM cells, the fields in anechoic chambers must be created by specially designed radiating devices placed within the chambers.

TEM cells have been made in a wide variety of sizes from 8 centimeters to room size. Anechoic chambers can be made in all sizes, too, but are generally room size. TEM cells are useful only below a specified frequency that decreases with increasing cell size. Anechoic chambers are useful only between two frequencies that depend on the properties of the absorber. The absorber tends to pass electromagnetic radiation of too low a frequency, allowing the chamber walls to reflect, and to reflect radiation of too high a frequency.[60] At the low-frequency end, anechoic chambers are useful above about 200 megahertz. At the high-frequency end, anechoic chambers presently reach about 100 gigahertz. However, with improved absorbers, they should be able to reach down as low as 30 megahertz and as high as 300 gigahertz to meet the goal of Table 150. Both TEM cells and the anechoic chambers are entirely shielded from the external electromagnetic environment and are suitable for accurate measurements.

Reverberation Chambers

Reverberation chambers are also enclosed metallic chambers that are fully shielded from the outside electromagnetic environment. But the similarity to TEM cells and anechoic chambers ends there. Reverberation chambers produce highly complex fields by design. A field is generated within the chamber by special radiators. That field is literally stirred by a giant rotating reflector to produce as much of a random (statistically uniform) character as possible. In immunity testing, this stirred field bombards the electronic device under test from every angle. Reverberation chambers of useful sizes operate above about 200 megahertz and have no theoretical upper frequency limit. With proper supporting instrumentation, they should also be able to reach the goal of 300 gigahertz listed in Table 150.

Open Sites

Open sites are outside test facilities set up at ground level. Ideally, in order to minimize reflections, no other structures should be anywhere near these facilities. The base of an open site is usually a flat surface made of a conducting wire mesh (a *ground plane*). While the ground plane will produce reflections, its simple geometry makes accounting for those reflections systematic. The lower frequency limit for an open site depends on the size of the ground plane. The upper frequency limit depends on the characteristics of the wire mesh used for the ground plane. Open sites are typically used from 30 kilohertz to 3 gigahertz. This is the widest frequency range supported by any of the test facilities discussed here (five orders of magnitude, or a factor of 10^5 or 100,000). The relatively low cost of the open site facility makes it financially accessible for testing, but its openness to the surrounding physical and electromagnetic environment makes its use problematical.

Measurement Needs by Application

Table 150 summarized the generic improvements needed in measurement capability for EMC broadly. Against this background the specific measurement needs associated with the individual applications can be described. Table 153 provides an overview of the areas of principal measurement need, arranged by application area.

Immunity Measurements

The main requirement in performing measurements of radiated immunity is to expose the electronic equipment under test to a uniform electromagnetic field of known strength while monitoring the performance of the equipment. Generally, the methods of exposing the equipment can be standardized across broad product lines, while the methods of detecting interference must be product specific. For example, a computer might be tested for interference by repeated checks for errors in a lengthy calculation during exposure to a potentially interfering field. This method of detection would be inapplicable to many other types of electronic products. This discussion focuses principally on the part of immunity testing that can be standardized across broad product lines -- the process of exposure to the interfering fields.

Broadband Testing

Both TEM cells and anechoic chambers are used to expose electronic equipment for immunity measurements. Because of the broad frequency range required, no single facility is capable of generating the required test field over the entire range. The transverse electromagnetic (TEM) cell[61] provides a fairly uniform field for immunity measurements for frequencies below about 100 megahertz. For frequencies above 100 megahertz, it is difficult to establish a uniform, high-intensity field over a large test volume. A gigahertz version of the TEM cell, the Gigahertz TEM (GTEM) cell,[62] uses broadband terminations and absorbers in an attempt to reach frequencies above 1 gigahertz, but further improvement is required before GTEM cells can provide adequate accuracy and repeatability.

For both anechoic-chamber measurements and open-site measurements, a single antenna is sometimes used to create the test field. Antennas that are presently available can produce a weak uniform field far from the antenna or a strong nonuniform field near the antenna. A field that is strong and uniform near the antenna and that extends over large volumes is needed. The development of *near-field*

Table 153
Areas of Measurement Need by Application

IMMUNITY MEASUREMENTS	
Exposure Techniques	broadband testing pulse testing large-system testing current-injection testing high-spatial-resolution measurements electrostatic discharge
Shielding Effectiveness	planar materials cables, connectors, and enclosures artifact immunity reference standards
EMISSIONS MEASUREMENTS	
Radiated Fields	antenna calibration broadband and pulse field measurements standard field radiators
Total Radiated Power	reverberation chamber characterization hybrid chamber characterization
ELECTROMAGNETIC ENVIRONMENT CHARACTERIZATION MEASUREMENTS	
Field Probe Development	electric field probes magnetic field probes photonic probes electric and magnetic field (Poynting vector) sensors
Probe Characterization	standard field generators anechoic chamber characterization and evaluation
Mathematical Models for Characterizing Complex Fields	deterministic space-time models statistical models

antenna arrays is necessary to enable producing such fields. A key challenge will be achieving good performance over the broad frequency ranges of interest.

The use of different facilities to cover different portions of the frequency spectrum for immunity testing is inconvenient and expensive. The use of a *hybrid chamber* to cover the entire frequency range of interest for immunity measurements is greatly needed. The hybrid chamber performs as a TEM cell at low frequencies and as a reverberation chamber at high frequencies. There is a transition frequency range where the characteristics of the hybrid chamber change and complicate measurement applications. This difficulty needs to be resolved if the chamber is to realize its full potential.

Pulse Testing

Most present measurement techniques for EMC, like those described above, employ continuous waves, that is, waves that are uninterrupted in time. However, pulse testing is becoming an important measurement need for modern digital electronics. The reason for this need is that digital electronic systems tend to radiate pulses and to be susceptible to pulses. For slower pulses, with frequency components below 100 megahertz, the TEM cell can be used to generate a uniform pulsed field for immunity testing. For faster pulses (sub-nanosecond) listed in Table 150, a *conical monopole antenna* over a ground plane can be used to generate a pulsed near field that is fairly uniform, or a *TEM horn antenna* can be used to generate a pulsed far field. However, neither radiator is satisfactory for generating a strong pulsed field that is uniform over a large volume. Development of such a radiator is needed to support improved immunity testing.

Large-System Testing

Because of the difficulties in generating strong uniform fields over a large test volume, the reverberation chamber[63] was developed to produce a *statistically uniform* field throughout a closed metal cavity. This facility has the advantage of producing strong fields in a shielded environment and is most effective above about 200 megahertz where a number of electromagnetic modes (field configurations) are supported by the cavity. The interpretation of data from reverberation chambers still needs further refinement to improve the accuracy and repeatability of immunity measurements, but the usefulness of the facility has been well demonstrated.

To address both large systems and broad bandwidths simultaneously, the hybrid chambers can be used. Generally, they can be made as large as necessary to accommodate large systems (such as aircraft), but the accuracy and repeatability of immunity measurements made in such chambers must yet be determined. Also, the feasibility of doing pulse measurements in hybrid or reverberation chambers needs to be established.

Current-Injection Testing

Measurements of immunity to conducted electromagnetic interference[64] involve the injection of currents into power lines or cables by use of a transformer or some appropriate circuitry. The frequency range of interest is generally lower than that of radiated immunity measurements because conducted EMI attenuates rapidly at high frequencies. However, some military standards[65] call for a frequency range of 50 kilohertz to 400 megahertz. The present measurement methods are probably adequate for present conducted-immunity measurement needs, but some refinements and improvements in accuracy and repeatability are required.

High-Spatial-Resolution Measurements

Interest is surfacing in measurements of interfering currents or fields with unusually high spatial resolution. Resolution below 1 micrometer will likely be of interest. Such measurement capability can be applied to both integrated circuits and biological systems.

Electrostatic Discharge

Damage to, or malfunction of, electronic equipment from electrostatic discharge (ESD) has become a serious issue in the electronics industry.[66] The problem of immunity to ESD involves both conduction paths and radiation paths. The measurement of the waveforms of electric-current pulses generated by commercial ESD simulators poses no special problem, but the measurement of field waveforms radiated by ESD requires special broadband antennas.[67] It is important to achieve better understanding of ESD effects because modern electronics are more susceptible to ESD and because new European EMC standards will call for ESD immunity testing.

Shielding Effectiveness

Measurement of the shielding effectiveness of modern materials is very important and useful to industry in achieving immunity in products. A number of methods are currently in use, but they tend to give different values of shielding effectiveness even for the same material sample. The various methods for measuring the shielding effectiveness of thin sheets of *planar materials* have tradeoffs in accuracy and frequency range,[68] and none of the methods have been proven useful for frequencies much above 1 gigahertz.

The problem of measuring the shielding effectiveness of *cable, connectors, and enclosures* is even more difficult, and no standard methods exist at this time. A section of coaxial line with a circular hole in the outer shield is currently being evaluated for use as an *artifact immunity reference standard*.[69] This special line provides an accurately known *weak* point for the entry of electromagnetic radiation. It can be used for determining the accuracy of shielding-effectiveness measurements made in reverberation chambers and anechoic chambers. Further comparisons of measurements and theory are needed to determine the usefulness of this reference standard.

Emissions Measurements

Emissions measurements may require determination of the radiated fields in specified directions from the equipment under test, or they may require determining the total radiated power in all directions coming from the equipment under test. Some measurements of radiated emissions can be done with the same facilities that are used for radiated immunity measurements, as indicated in Table 152.

Radiated Fields

In order to make accurate measurements of radiated fields for immunity and emissions testing broadly and for environment characterization as well, it is necessary to *calibrate antennas.* These antennas create fields and measure fields. The antennas that measure fields come in two principal forms: broadband antennas, often small in size (*field probes*); and larger receiving antennas with higher sensitivity and more limited bandwidth. Both of these must be calibrated over their entire frequency range of use. Such calibrations are typically performed in a TEM cell, over an open-site ground plane, or in an anechoic chamber. For frequencies up to 40 gigahertz,[70] calibrations can now be made with acceptable accuracy in most cases, but new methods are required to extend this frequency range to 300 gigahertz. TEM horns for broadband and pulse field measurements are currently calibrated with a conical monopole facility, but these calibrations need to be extended to greater bandwidth. In addition to calibrating antennas, it is also necessary to calibrate the test facilities themselves.

The FCC specifies that radiated-emission measurements for computing devices are to be made on an open site over a metal ground screen for the frequency range of 30 to 1000 megahertz. A receiving antenna is scanned in height (usually from 1 to 4 meters) to find the maximum field for both vertical and horizontal polarizations of the radiated field at a fixed horizontal distance from the equipment under test. The same facility could be used for radiated immunity tests, but this is not usually done because of FCC restrictions on transmitting at an open site.

The FCC permits the same emissions measurements to be made indoors in a semi-anechoic chamber if the chamber can be shown to perform in a manner equivalent to an open site. A semi-anechoic chamber has a reflecting floor, like an open site, and absorbing walls and ceiling which should return no reflections from their directions, analogous to an ideal open site. However, creating a successful indoor chamber is difficult because the absorber usually performs poorly at low frequencies near 30 megahertz, the lower end of the specified FCC range. This is an area where a better absorber would allow for more convenient, less expensive indoor measurements. Part of the difficulty is in measuring the performance of a radio-frequency absorber at low frequencies. A new pulsed method appears promising,[71] but further improvement is needed in this area.

Qualification of open sites or anechoic chambers for FCC measurements of radiated emissions involves measuring the insertion loss between a pair of standard dipole antennas.[72] The expected measurement result is often difficult to obtain because of reflections or an imperfect ground plane. This represents a serious problem in obtaining an acceptable site for radiated emissions measurements. Further improvement is needed, particularly for standard field radiators and anechoic chambers, so that industry will be able to make required emissions measurements.

Total Radiated Power

Total radiated power at relatively low frequencies (below 200 megahertz) can be measured in a TEM cell. In the cell the equipment under test creates the signals (emissions). These signals are received at the waveguide ports that were used for injecting signals for immunity tests. Emissions over a broader frequency range have been measured using a GTEM cell,[73] but the accuracy and repeatability of this method must yet be established. A complete characterization of the emissions requires rotating the equipment under test in the TEM cell,[74] and a more convenient method which places the equipment at the center of three orthogonal loop antennas[75] is under development. This method needs further development in order to make it practical and to establish its accuracy.

Measurements of total radiated power at higher frequencies (above 200 megahertz) are currently being made in reverberation chambers by using a receiving antenna rather than a transmitting antenna. The uncertainties are similar to those encountered in using reverberation chambers for immunity measurements. Further improvements are needed in both accuracy and repeatability. The situation is similar when using a hybrid chamber for total radiated power measurements.

Electromagnetic Environment Characterization Measurements

The electromagnetic environment must be well characterized to support setting immunity standards at the proper level in the environments where products are used. The electromagnetic environment must also be well characterized to support research on health effects of electromagnetic radiation. Since the electromagnetic environment varies with time and location, a variety of measurement methods and mathematical models are needed to obtain an adequate characterization.

Field Probe Development

Considerable effort has already gone into developing electric and magnetic field probes that are small and convenient and have a wide bandwidth.[76] Some of these sensors have the additional capability of measuring power density (magnitude of the Poynting vector). Some electric field probes are now operating at frequencies up to about 40 gigahertz, but broader bandwidth is needed to support characterization of complex electromagnetic fields with broader frequency content, especially pulsed fields.

Photonic probes are also well suited for pulsed field measurements. They employ optical-fiber signal leads, which, because of their non-conducting nature, minimize interference with the fields to be measured. Photonic probes also enable preserving phase information needed to determine directionality. Their potential broad bandwidth also makes them promising for measuring the shape of pulsed fields. But additional development is needed to exploit their full potential bandwidth, to realize higher sensitivity, and to achieve higher reliability.

A further need is for field probes that function at frequencies above 40 gigahertz and extending to 300 gigahertz. To characterize new field probes at frequencies above 40 gigahertz, new standard field generators will be needed that can create fields with values that are accurately known from first principles. Also, the anechoic chambers in which the probes are tested must be better characterized and evaluated to determine the level of unwanted reflections from their walls.

Both TEM cells and anechoic chambers can be used for testing the field probes. In this sense, these facilities support the characterization of the electromagnetic environment, as noted in Table 152.

For fields from dc to perhaps 500 kilohertz, the principal need is for test methods for characterizing the performance of versatile new measurement systems. In particular, measurement methods are needed for determining dynamic range, accuracy, and frequency response, especially in the presence of broadband (pulsed or transient) signals.

Mathematical Models for Characterizing Complex Fields

In any given electromagnetic environment, only a limited number of field measurements can be made. To characterize the spatial and temporal behavior of that environment, deterministic space-time models[77] have been used to estimate the field properties at locations away from actual measurement locations. As electronic devices proliferate and as the electromagnetic environment becomes more complex, deterministic approaches will not be practical and thus statistical models will be needed, along with better supporting field probes. These models and probes will provide the improved environment characterization necessary to support the development of sensible immunity standards that will accomplish their purpose without adding unnecessarily to product cost.

ENDNOTES

1. F. Jay, ed., *IEEE Standard Dictionary of Electrical and Electronics Terms, ANSI/IEEE Std 100-1984, Third Edition* (1984). The IEEE definition for EMC is "The capability of electronic equipments or systems to be operated in the intended operational electromagnetic environment at designed levels of efficiency." The IEEE definition of EMI is "Impairment of a wanted electromagnetic signal by an electromagnetic disturbance."

2. The ruler has a logarithmic scale. That is, factors of 10 increase in frequency are equally spaced. This type of scale makes it possible to illustrate a very wide range of frequencies in a single figure.

3. Components of the field near to the source may not be free to travel great distances, but they may be able to extend far enough to intercept electronic equipment and to interfere with its operation.

4. Jeffrey K. Eckert, *Commercial EMC Standards of the United States*, pp. 3.6-3.7 (Gainesville, Virginia: Interference Control Technologies, Inc.; 1988).

5. "Interim Guidelines on Limits of Exposure to 50/60 Hz Electric and Magnetic Fields", International Non-ionizing Radiation Committee of the International Radiation Protection Association, *Health Physics*, Vol. 58, No. 1, pp. 113-122 (January 1990). According to the preface in the guidelines: "The IRPA/INIRC, in cooperation with the Environmental Health Division of the World Health Organization (WHO), has undertaken responsibility for the development of health criteria documents on NIR [non-ionizing radiation]. These form part of the WHO Environmental Health Criteria Programme, which is sponsored by the United Nations Environment Programme (UNEP)."

6. "Standard Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz", Draft Standard C95.1 SCC28, Institute of Electrical and Electronics Engineers (November 27, 1991).

7. *Biological Effects of Power Frequency Electric and Magnetic Fields: Background Paper*, Office of Technology Assessment, Congress of the United States, p. 3 (May 1989).

8. H. Keith Florig, "Containing the Costs of the EMF Problem", working paper developed by Resources for the Future, Washington, DC, p. 2 (December 1991).

9. Art Wall, "FCC Measurements", *NIST Metrology Short Course* (June 19-22, 1990).

10. William G. Duff, *EMC in Telecommunications* (Gainesville, Virginia: Interference Control Technologies, Inc.; 1988).

11. *Special Issue on the Nuclear Electromagnetic Pulse*, *IEEE Transactions on Electromagnetic Compatibility*, pp. 1-187 (February 1978).

12. William von Achen, "The Apache Helicopter: An EMI Case History", *Compliance Engineering*, pp. 11-17, 111-112 (Fall, 1991).

13. "VLF Radiation Emission Levels Vary Widely Among Popular PC Monitors", *InfoWorld*, pp. 74-78 (November 12, 1990).

14. *Biological Effects of Power Frequency Electric and Magnetic Fields: Background Paper*, Office of Technology Assessment, Congress of the United States, p. 3 (May 1989).

15. Art Wall, "FCC Measurements", *NIST Metrology Short Course* (June 19-22, 1990).

16. Edwin L. Bronaugh and William S. Lambdin, *Electromagnetic Interference Test Methodology and Procedures* (Gainesville, Virginia: Interference Control Technologies, Inc.; 1988).

17. R.G. Struzak, "Terrestrial Electromagnetic Environment", in *Electromagnetic Compatibility in Radio Engineering*, ed. by W. Rotkiewicz (Gainesville, Virginia: Elsevier; 1982).

18. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 22-1 (January 1991).

19. *Guidance to Certification of Aircraft Electrical/Electronic Systems for Operation in the High Intensity Radiated Fields (HIRF) Environment*, Society of Automotive Engineers, SAE AE4R Committee Report (July 1991).

20. Edward H. Phillips, "GAMA Says Criteria for Testing HIRF's Effects Too Complex, Costly", *Aviation Week & Space Technology*, pp. 53-54 (March 11, 1991).

21. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, pp. 28-1-3 (January 1991).

22. Winn L. Rosch, "What You Should Know About the FCC Emission Standards", *PC Magazine*, pp. 351-357 (February 26, 1991).

23. Isador Straus, "FCC Gives Nod to ANSI Test Methods", *Compliance Engineering*, pp. 13-22 (Spring, 1991).

24. Isador Straus, "Electromagnetic Compatibility and the Common Market", *Compliance Engineering*, pp. 13-22 (Summer, 1991).

25. "VLF Radiation Emission Levels Vary Widely Among Popular PC Monitors", *InfoWorld*, pp. 74-78 (November 12, 1990).

26. *IEEE Standard Procedures for the Measurement of Electric and Magnetic Fields from Video Display Terminals (VDTs) from 5 Hz to 400 kHz*, IEEE P1140 Draft-8 (November 24, 1991).

27. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 37-1, 2, 3, and 5 (January 1991).

28. *1991 Electronic Market Data Book*, Electronic Industries Association, p. 100 (1991). The estimate is from Delco Electronics.

29. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 37-11 (January 1991).

30. *SAE J551*,"Electromagnetic Emissions Procedures for Vehicle Components (Except Aircraft)", Society of Automotive Engineers (October 1985).

31. *SAE J1113*, "Electromagnetic Susceptibility Procedures for Vehicle Components (Except Aircraft)", Society of Automotive Engineers (June 1984).

32. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 46-2 (January 1991).

33. *MDS-201-0004*, "Electromagnetic Compatibility Standard for Medical Devices", Food and Drug Administration (October 1979).

34. William S. Staewen, "The Electromagnetic Compatibility Standard for Medical Devices", *Biomedical Instrumentation & Technology*, pp. 230-251 (May/June 1991).

35. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 46-2 (January 1991).

36. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, pp. 38-12 to 14 (January 1991).

37. Donald R. Shepherd, "Susceptibility: Passing the 200 V/m Barrier", *Compliance Engineering*, pp. 13-17 (Fall, 1990).

38. Stephen A. Halperin, "Anything But Static: The EOS/ESD Association Takes Charge", *Compliance Engineering*, pp. 13-28 (Winter, 1991).

39. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 31-1 and 2 (January 1991).

40. John D. M. Osburn, *Military Electromagnetic Compatibility Standards of the United States* (Gainesville, Virginia: Interference Control Technologies, Inc.; 1988). John D. M. Osburn, *Supporting and Unique Military Electromagnetic Compatibility Standards of the United States* (Gainesville, Virginia: Interference Control Technologies, Inc.; 1988).

41. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce, p. 31-6 (January 1991). "While Motorola had not yet filed with the FCC, it announced an innovative satellite-cellular telephone system call Iridium in June 1990, planned for full service in 1996."

42. B. E. Keiser, *Principles of Electromagnetic Compatibility* (Dedham, Massachusetts: Artech House; 1979).

43. M. I. Montrose, "Overview on Design Techniques for Printed Circuit Board Layout Used in High Technology Products", *Proceedings of the 1991 IEEE International Symposium on Electromagnetic Compatibility*, Cherry Hill, NJ, pp. 61-66 (August 1991).

44. R. M. Showers and E. Darvin, "Replace System Emission Tests with Unit Tests", *Proceedings of the 1991 IEEE International Symposium on Electromagnetic Compatibility*, Cherry Hill, NJ, pp. 163-168 (August 1991).

45. The "NIST EMI Metrology Short Course" was offered June 19-22, 1990, to an audience of about 50 from industry and 10 from government. The course is normally offered about every two or three years.

46. Bert G. Simson, *Conformity Assessment Workshop on Electromagnetic Compatibility*, published as an NIST document with number NISTIR 4611 (June 1991). The workshop was held on April 4, 1991.

47. R. M. Showers, "The Uniform Standards Initiative", *Proceedings of the 1991 IEEE International Symposium on Electromagnetic Compatibility*, Cherry Hill, NJ, pp. 332-336 (August 1991).

48. Three trade journals on Electromagnetic Compatibility are : (1) *Compliance Engineering*, (2) *EMC Test and Design*, and (3) *Interference Technology Engineers' Masters*, *The International Journal of EMC*.

49. Jeffrey K. Eckert, *Commercial EMC Standards of the United States* (Gainesville, Virginia: Interference Control Technologies; 1988).

50. John D. M. Osburn, *Military Electromagnetic Compatibility Standards of the United States* (Gainesville, Virginia: Interference Control Technologies; 1988).

51. C. M. Wintzer, *International Commercial EMC Standards* (Gainesville, Virginia: Interference Control Technologies,Inc.; 1988).

52. Isador Straus, "FCC Gives Nod to ANSI Test Methods", *Compliance Engineering*, pp. 13-22 (Spring, 1991).

53. *IEEE Standard Procedures for the Measurement of Electric and Magnetic Fields from Video Display Terminals (VDTs) from 5 Hz to 400 kHz*, IEEE P1140 Draft-8 (November 24, 1991).

54. B. Rakoski, "MIL-STD-461 Update", *EMC Excellence*, p. 13 (July/August 1991).

55. Isador Straus, "Electromagnetic Compatibility and the Common Market", *Compliance Engineering*, pp. 13-22 (Summer, 1991).

56. Bert G. Simson, *Conformity Assessment Workshop on Electromagnetic Compatibility*, published as an NIST document with number NISTIR 4611 (June 1991). The workshop was held on April 4, 1991.

57. "VLF Radiation Emission Levels Vary Widely Among Popular PC Monitors", *InfoWorld*, pp. 74-78 (November 12, 1990).

58. "Interim Guidelines on Limits of Exposure to 50/60 Hz Electric and Magnetic Fields", International Non-ionizing Radiation Committee of the International Radiation Protection Association, *Health Physics*, Vol. 58, No. 1, pp. 113-122 (January 1990). According to the preface in the guidelines: "The IRPA/INIRC, in cooperation with the Environmental Health Division of the World Health Organization (WHO), has undertaken responsibility for the development of health criteria documents on NIR [non-ionizing radiation]. These form part of the WHO Environmental Health Criteria Programme, which is sponsored by the United Nations Environment Programme (UNEP)."

59. A decibel is a logarithmic measure of a power ratio. The number of decibels is equal to 10 times the logarithm of the power ratio to the base 10. An increase of 3 decibels from one power level to another represents a factor of approximately 2 increase. An increase of 10 decibels from one power level to another represents a factor of 10 increase.

60. A. Lehto, J. Tuovinen, and A. Raisanen, "Reflectivity of Commercial Absorbers at 100 - 200 GHz", *IEEE Antennas and Propagation Symposium*, London, Ontario, pp. 1202-1205 (June 1991).

61. M. T. Ma, M. Kanda, M. L. Crawford, and E. B. Larsen, "A Review of Electromagnetic Compatibility/Interference Measurement Methodologies", *Proceedings of the IEEE*, pp. 388-411 (March 1985).

62. D. Hansen, P. Wilson, D. Koenigstein, and H. Schaer, "A Broadband Alternative EMC Test Chamber Based on TEM Cell Anechoic-Chamber Hybrid Approach", *Proceedings of the 1989 International Symposium on Electromagnetic Compatibility*, Nagoya, Japan, pp. 133-137 (September 1989).

63. M. T. Ma, M. Kanda, M. L. Crawford, and E. B. Larsen, "A Review of Electromagnetic Compatibility/Interference Measurement Methodologies", *Proceedings of the IEEE*, pp. 388-411 (March 1985).

64. William G. Duff, *Fundamentals of Electromagnetic Compatibility*, pp. 10.13-10.24 (Gainesville, Virginia: Interference Control Technologies, Inc.; 1988).

65. John D. M. Osburn, *Military Electromagnetic Compatibility Standards of the United States* (Gainesville, Virginia: Interference Control Technologies, Inc.; 1988).

66. W. Boxleitner, "How to Defeat Electrostatic Discharge", *IEEE Spectrum*, pp. 36-40 (August 1989).

67. P. F. Wilson, A. R. Ondrejka, M. T. Ma, and J. M. Ladbury, "Electromagnetic Fields Radiated from Electrostatic Discharges Theory and Experiment", NBS Technical Note 1314 (February 1988).

68. P. F. Wilson and M. T. Ma, "A Study of Techniques for Measuring the Electromagnetic Shielding Effectiveness of Materials", NBS Technical Note 1095 (May 1986).

69. D. A. Hill, M. L. Crawford, M. Kanda, D. I. Wu, *Aperture Coupling to Shielded Transmission Lines: Theory and Experiment*, NISTIR 3988 (April 1992).

70. 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", NIST Technical Note 1335 (March 1990).

71. S. Tofani, A. Ondrejka, and M. Kanda, "Time-Domain Method for Characterizing the Reflectivity of Absorbing Materials from 30 to 1000 MHz", *IEEE Transactions on Electromagnetic Compatibility*, pp. 234-240 (1991).

72. R. G. FitzGerrell, "Site Attenuation", *IEEE Transactions on Electromagnetic Compatibility*, pp. 38-40 (1986).

73. E. L. Bronaugh and J. D. M. Osburn, "Radiated Emissions Test Performance of the GHz TEM Cell", *Proceedings of the 1991 IEEE International Symposium on Electromagnetic Compatibility*, Cherry Hill, NJ, pp. 1-7 (August 1991).

74. M. T. Ma, M. Kanda, M. L. Crawford, and E. B. Larsen, "A Review of Electromagnetic Compatibility/Interference Measurement Methodologies", *Proceedings of the IEEE*, pp. 388-411 (March 1985).

75. M. Kanda and D. A. Hill, "A Three-Loop Method for Determining the Radiation Characteristics of an Electrically Small Source", *IEEE Transactions on Electromagnetic Compatibility*, pp. 1-3 (February 1992).

76. M. T. Ma, M. Kanda, M. L. Crawford, and E. B. Larsen, "A Review of Electromagnetic Compatibility/Interference Measurement Methodologies", *Proceedings of the IEEE*, pp. 388-411 (March 1985).

77. M. Kanda and J. Randa, "Estimation of Electromagnetic Fields in Complex Environments", *9th International Zurich Symposium and Technical Exhibition on Electromagnetic Compatibility*, pp. 337-342 (March 1991).

APPENDIX 1

INDUSTRY DATA FROM THE BUREAU OF THE CENSUS

March 1992

Appendix 1

INDUSTRY DATA
FROM THE BUREAU OF THE CENSUS

The Bureau of the Census in the U.S. Department of Commerce surveys U.S. industries periodically to develop a wide variety of economic data. U.S. industries are those manufacturing in the U.S., whatever the nationality of their ownership. Included are data on shipments from U.S. manufacturing facilities, imports and exports, employment and payroll, inventories, and capital expenditures. The information gathered is published in one or more of three principal series of documents, as shown in Table 154. Other publications and databases provide considerable additional data on imports and exports.

Table 154
Reporting Documents of the Bureau of the Census
for U.S. Manufactures

Document	Published	Detail
Current Industrial Reports	annually	7 digits
Annual Survey of Manufactures	annually	5 digits
Census of Manufactures	every five years	7 digits

The data developed by the Bureau of the Census are used throughout the U.S. Government and by many non-government organizations, including industry trade associations, as the basis for tracking industrial progress. For example, the International Trade Administration of the U.S. Department of Commerce draws data from the Bureau of the Census for the annual publication, the *U.S. Industrial Outlook*.[1]

STANDARD INDUSTRIAL CLASSIFICATION SYSTEM

The data in the reports of the Bureau of the Census are arranged according to the Standard Industrial Classification (SIC) System developed under the direction of the Office of Management and Budget (OMB) in the Executive Office of the President. The SIC System uses a structured hierarchy of numerical codes for describing products and services provided within the U.S. economy. The SIC System is redefined every five years. The most recent definition is contained in the *Standard Industrial Classification Manual, 1987*.[2]

The numerical codes in the SIC System contain from two to seven digits, depending on the level of aggregation that they represent. Codes with fewer digits represent higher levels of aggregation. Additional digits are added to provide subdivision. As shown in Table 155, the first two digits in each SIC code are taken together and describe the *major group*. This is the most general level of division in the hierarchy. The third digit subdivides each major group into *industry groups*. The fourth digit subdivides each industry group into *industries*. The fifth digit subdivides each industry into *product classes*. Finally, the sixth and seventh digits are taken together and subdivide each product class into *products*. OMB determines the overall structure to the four-digit level (the *industry* level). The Bureau of the Census in the U.S. Department of Commerce expands the OMB definition to the seven-digit level (the *product* level). Some U.S. Government agencies develop their own SIC breakdowns at the five- and seven-digit levels to serve their special needs. Industry users also sometimes develop their own breakdowns at these levels. For these reasons, only the four-digit level is truly standardized.

Table 155
Levels of the Standard Industrial Classification System

	Example	
Level	SIC Code	Code Title
major group	35	industrial and commercial machinery and computer equipment
industry group	357	computer and office equipment
industry	3572	computer storage devices and parts
product class	35721	computer storage devices
product	3572123	rigid magnetic disk drives less than 5-1/4 inch

In the *Standard Industrial Classification Manual*, sets of two-digit codes are grouped together to form *Divisions*, as shown in Table 156. Virtually all of the products of the electronic and electrical-equipment industries fall into *Division D: Manufacturing*, which includes SIC codes 20-39.[3] Most of the products are concentrated, respectively, in SIC 36, *Electronic and Other Electrical Equipment and Components, Except Computer Equipment* and in SIC 35, *Industrial and Commercial Machinery and Computer Equipment*.

NUMERICAL LIST OF MANUFACTURED AND MINERAL PRODUCTS

For all of the manufactured products in Division D, the structure used by the Bureau of the Census for SIC categories at the seven-digit level is depicted in the *Numerical List of Manufactured and Mineral Products*, which was issued most recently in 1987.[4] This publication also contains a crosswalk to earlier versions of the SIC hierarchy so that changes can be tracked. The publication does not contain data; but it does summarize the types of data collected by the Bureau of Census, and the units of measure for each, for all seven-digit codes for manufactures. The publication also provides annotations for each seven-digit SIC code to indicate if data for that code are available in an issue of the *Current Industrial Reports*. Thus the *Numerical List* is a guide to the structure and availability of types of data for manufactures. Mineral products are also covered.

CURRENT INDUSTRIAL REPORTS

The most detailed current data for the electronic and electrical-equipment industries are contained in the *Current Industrial Reports* (CIRs) from the Bureau of the Census. Each issue contains data on product shipments, expressed as both number of items, or quantity, and dollar value. Data on exports and imports are also included, along with a translation between the SIC categories used for the data

Table 156
Divisions of the Standard Industrial Classification System

Division	Major Groups
A. Agriculture, Forestry, and Fishing	01-09
B. Mining	10-14
C. Construction	15-17
D. Manufacturing	20-39
E. Transportation, Communications, Electric, Gas, and Sanitary Services	40-49
F. Wholesale Trade	50-51
G. Retail Trade	52-59
H. Finance, Insurance, and Real Estate	60-67
I. Services	70-89
J. Public Administration	91-97
K. Nonclassifiable Establishments	99

on shipments and the different classification categories used for data on exports and imports. Exports and imports are classified according to an international system, adopted by the U.S. in January 1989, called the Harmonized Commodity Description and Coding System (or Harmonized System for short). More specifically, the U.S. Government classifies exports according to a hierarchy that is based on the Harmonized System and that is described annually by the Bureau of Census in its publication *Harmonized System-Based Schedule B, Statistical Classification of Domestic and Foreign Commodities Exported from the Unites States*. The U.S. Government classifies imports according to a hierarchy that is also based on the Harmonized System and that is described annually by the International Trade Commission in its publication *Harmonized Tariff Schedule of the United States, Annotated (For Statistical Reporting Purposes)*.

Issues in the *Current Industrial Reports* series are prepared annually, but only for selected subsets of categories in the SIC System. The issues are based on data collected from industry in survey instruments specially designed for the purpose. SIC categories are selected for coverage based on a combination factors, including the demand from users of the data and the availability of funding for developing the data. The private sector funds the development of some issues of the *Current Industrial Reports*. Such issues then become available to all users. Over 100 issues are presently available for various sectors of manufactures.

ANNUAL SURVEYS OF MANUFACTURES

The *Annual Survey of Manufactures*, as its name suggests, is also published annually but only for years between those addressed in the *Census of Manufactures* which is published every five years (1982, 1987, etc.) The *Annual Survey of Manufactures* covers the entire manufacturing industry (SIC codes 20 through 39). It is based on its own survey instrument but resolves data only to the five-digit level. Because of its reduced level of detail, it is the most compact overview of U.S. shipments of manufactured products. A single bound document contains all of the shipments data for manufactures.[5]

CENSUS OF MANUFACTURES

The *Census of Manufactures* is also based on its own survey instrument. It resolves data to the seven-digit level. It is published in over 80 volumes, each of which addresses from one to four three-digit SIC categories.

The SIC categories reflected in the above documents of the Bureau of the Census are used for defining the electronics industry in Appendix 2 and the electrical-equipment industry in Appendix 3. Much of the economic data in other chapters of this document comes directly from the above documents, or was compiled entirely or partly from those documents by other organizations.

ENDNOTES

1. *1991 U.S. Industrial Outlook*, International Trade Administration, U.S. Department of Commerce (January 1991).

2. *Standard Industrial Classification Manual, 1987*, Office of Management and Budget, Executive Office of the President (1987). Available from the National Technical Information Service as publication number PB87-100012.

3. The only exception is integrated software sold with electronic products. It is classified with *services* in SIC 73.

4. *1987 Census of Manufactures and Census of Mineral Industries: Numerical List of Manufactured and Mineral Products*, publication number MC87-R-1, Bureau of the Census, U.S. Department of Commerce (February 1989). Available from the National Technical Information Service as publication number PB89-238042.

5. *1989 Annual Survey of Manufactures: Value of Product Shipments*, publication number M89(AS)-2, Bureau of the Census, U.S. Department of Commerce (June 1991).

APPENDIX 2

DEFINITION OF

U.S. ELECTRONICS INDUSTRY

Appendix 2

DEFINITION OF U.S. ELECTRONICS INDUSTRY

according to the Electronic Industries Association
in terms of the
Standard Industrial Classification System of the
Office of Management and Budget

The Electronic Industries Association (EIA) has developed a definition of the products included in the electronics industry. The major categories of products in that definition are shown in Table 157. The categories of (1) aerospace, (2) automatic controls, industrial apparatus, and other instruments, and (3) motor vehicles focus principally on electronic content not elsewhere included in the definition.

Table 157
EIA's Definition of U.S. Electronics Industry:
Major Categories

- communications equipment
- electronic components
- computers and industrial electronics
  - computers and peripheral equipment
  - industrial electronics
  - electromedical equipment
- consumer electronics
- electronic-related products and services
  - aerospace
  - automatic controls, industrial apparatus, and other instruments
  - systems integration and computer services
  - motor vehicles
  - electronic-related office equipment

The definition can be expressed in terms of the Standard Industrial Classification (SIC) System described in Appendix 1 as shown in Table 158. EIA uses this definition to extract data on U.S. shipments and trade for electronic products from data collected by the U.S. Government. These data come primarily from the *Current Industrial Reports* published annually by the Bureau of the Census of the U.S. Department of Commerce. EIA publishes the data it assembles, along with detailed commentary, in its annual data book, the *19xx Electronic Market Data Book*.

For simplicity, the definition shown in Table 158 begins at the three-digit level. All three-digit codes are subdivided to the four-digit level to provide useful detail about what is included. If only a part of the products associated with a given code is included in EIA's definition, then the word *partial* appears next to that code. Further subdivision is then shown in an effort to reach one code, or a set of codes, that describe products that are fully included in the definition. Such codes are shown in **bold** type. In some cases, there are no codes describing products that are fully included, so the word *partial* appears next to even the most detailed code shown. Endnotes provide clarification when this occurs.

The definition in Table 158 was developed by EIA for the 1991 issue of the *Data Book*. Changes that EIA contemplated for the 1992 edition of the *Data Book* are described in the endnotes. For those cases in which additional SIC codes may be added or existing codes may be moved, square brackets [....] in Table 158 mark the receiving locations.

EIA adjusts the data that it uses in several ways. For example, EIA projects the data one year ahead and labels the projections as *estimates*. This process enabled EIA to provide data for 1990 in the 1991 edition of the *Data Book*, even though data for 1989 were the most recent available from the Bureau of the Census at the time of preparation of that edition. In each annual edition of the *Data Book*, EIA updates the data for earlier years based on the most recent data from the Bureau of the Census. In some cases, as noted above, EIA must subdivide the data in individual SIC codes if those codes are not already subdivided in a way useful for describing the electronics industry. For example, some codes describe products, such as aircraft, that are not electronic equipment themselves but that contain electronic equipment not sold separately and thus not captured in separate SIC codes. EIA estimates such electronic content for inclusion, and describes such inclusions accordingly. The endnotes explain where estimates of this type have been made. Other estimates of similar nature are also described in the endnotes. EIA also substitutes for some data of the Bureau of the Census other data that EIA obtains for selected electronic components directly from industry on a very current basis.[1]

Table 158
EIA's Definition of U.S. Electronics Industry in Terms of SIC System

	Standard Industrial Classification Code	
COMMUNICATIONS EQUIPMENT		
Communications Equipment	**366**	
telephone and telegraph apparatus	**3661**	
radio and television broadcasting and communications equipment	**3663**	
communications equipment, not elsewhere classified	**3669**	
[Rolling, Drawing, and Extruding of Nonferrous Metals]	[335]	partial
[nonferrous wiredrawing & insulating]	[3357]	partial
[fiber optic cable]	[**33579**][2]	
Search and Navigation	381	partial
search and navigation equipment	3812	partial
search and detection, navigation and guidance systems and equipment	**38122**	

(Table 158 continued)

	Standard Industrial Classification Code	
Measuring and Controlling Devices	382	partial
measuring and controlling devices, not elsewhere classified	3829	partial
commercial, geophysical, meteorological, and general purpose instruments	38295	partial
meteorological electronics equipment and radio astronomy equipment	**3829551**	
geophysical electronics equipment	**3829559**	

ELECTRONIC COMPONENTS

Electronic Components and Accessaries	367	partial
electron tubes	**3671**	
printed circuit boards	**3672**[3]	
semiconductors and related devices	**3674**	
electronic capacitors	**3675**	
electronic resistors	**3676**	
electronic coils and transformers	**3677**	
electronic connectors	**3678**	
electronic components, not elsewhere classified	3679	partial
filters except microwave and piezoelectric devices	**36791**	
microwave components and devices	**36793**[3]	
transducers, electrical/electronic input or output	**36795**[3]	
switches, mechanical types for electronic circuitry	**36796**	
printed circuit assembly	**36798**[3]	
all other electronic components, not elsewhere classified	36799[4]	partial
antenna systems, except structural towers sold separately	---	
auto antenna	**3679908**[3]	
antenna accesories, sold separately (pedestals, drives, passive reflectors, rotators, radomes, etc.)	**3679909**[3]	
delay lines (distributed constant, lumped constant, magnetostrictive, ultrasonic, etc.)	**3679911**[3]	
oscillators, except instrumentation and crystal types	**3679913**[3]	
rectifiers, electronic, except solid state	**3679914**[3]	
variable frequency oscillators and antenna tuner kits to be assembled by purchaser	**3679915**[3]	
magnetic recording and reproducing heads	---	
audio	**3679916**[3]	
video	**3679917**[3]	
digital	**3679918**[3]	
instrumentation	**3679919**[3]	
static power supply converters for electronic applications, sold separately	---	
regulated	**3679921**[3]	
unregulated	**3679923**[3]	
variable frequency	**3679925**[3]	

(Table 158 continued)

	Standard Industrial Classification Code	
other (a.c., d.c., converters and inverters, klystron, vibrators, etc.)	**3679929**[3]	
electronic cable harnesses and cable assemblies	**3679931**[3]	
cryogenic cooling devices (cryostats, etc.) for infrared devices	**3679933**[3]	
liquid crystal displays (LCD) and other liquid devices	**3679951**[3]	
magnetic bubble memories	**3679953**[3]	
magnetic cores	**3679955**[3]	
electronic parts not elsewhere classified and specialized	---	
electronic hardware	---	
sockets for electronic component insertion	---	
tube sockets	**3679961**[3]	
relay sockets	**3679963**[3]	
discrete semiconductor sockets	**3679965**[3]	
integrated circuit sockets	**3679966**[3]	
other (sockets for crystals and switches, etc.)	**3679969**[3]	
enclosures for electronic equipment	---	
metal enclosures, assembles (welded, etc.)	**3679971**[3]	
metal enclosures, modular assembled (card insertion, rack and panel, etc.)	**3679973**[3]	
other enclosures, not elsewhere classified	**3679979**[3]	
all other electronic parts and specialized electronic hardware, not elsewhere classified	**3679998**[3]	
Electrical Industrial Apparatus	362	partial
relays and industrial controls	3625	partial
relays, general purpose	36251[5]	partial
Rolling, Drawing, and Extruding of Nonferrous Metals	335	partial
nonferrous wiredrawing & insulating	3357	partial
electronic wire and cable	**3357A**	
telephone and telegraph wire and cable	**3357B**	
Household audio and video equipment	365	partial
household audio and video equipment	3651	partial
speaker systems, microphones, home-type electronic kits, and commercial sound equipment, including public address systems	36515	partial
loudspeakers sold separately, including radio, television, and automobile radio speakers	**3651554**[3]	

(Table 158 continued)

COMPUTERS AND INDUSTRIAL ELECTRONICS

Computers and Peripheral Equipment

	Standard Industrial Classification Code	
Computer and Office Equipment	357	partial
electronic computers	**3571**	
computer storage devices	**3572**	
computer terminals	**3575**	
computer peripheral equipment, not elsewhere classified	**3577**	
calculating and accounting equipment	3578	partial
coin and currency handling machines, cash registers, accounting, bookkeeping, and billing machines	35781	partial
automatic bank note dispensers	**3578111**	
funds transfer terminals	**3578154**	
point-of-sale terminals (POS) (retail devices)	**3578155**	
[calculators]	[**35782**][6]	
[other areas]	[------][7]	
Miscellaneous Electrical Equipment & Supplies	369	partial
magnetic and optical recording media	3695	partial
magnetic and optical recording media (unrecorded)	36950[8]	partial
rigid disks	---	
less than 5-1/4 inch (20.7 cm)	**3695001**	
5-1/4 inch (20.7 cm) but less than 8 inch (31.5 cm)	**3695002**	
8 inch (31.5 cm) and more	**3695003**	
flexible disks	---	
less than 5-1/4 inch (20.7 cm)	**3695005**	
5-1/4 inch (20.7 cm) but less than 8 inch (31.5 cm)	**3695006**	
8 inch (31.5 cm) and more	**3695007**	
optical disks	**3695009**	
computer tape (reel tape)	**3695011**	
computer cassette and cartridge	**3695013**	
other magnetic recording tape, including instrumentation tape and high density recording (HDDR)	**3695027**	
drum	**3695029**	
other magnetic recording media	**3695031**	
parts for magnetic and optical recording media	**3695032**	

Industrial Electronics

Electrical Industrial Apparatus	362	partial
relays and industrial controls	3625	partial
specific purpose industrial controls	**36252**	

(Table 158 continued)

	Standard Industrial Classification Code	
general purpose industrial controls	**36253**	
motor controller accessories and parts for industrial controls	**36254**	
Miscellaneous Electrical Equipment & Supplies	369	partial
[magnetic and optical recording media]	[3695]	partial
[magnetic and optical recording media (unrecorded)]	[36950][8]	partial
[audio range tape]	[---]	
[other, including pancakes]	[**3695019**][9]	
[video tape]	[---]	
[cassette, 3/4 inch (19 mm) and 2 inch (51 mm)]	[**3695023**][10]	
[other video tape, including pancakes]	[**3695025**][11]	
electrical equipment & supplies, not elsewhere classified	3699	partial
electronic teaching machines, teaching aids, trainers, and simulators	**36991**	
laser systems and equipment, except communications	**36992**	
ultrasonic equipment (except medical)	**36995**	
other electronic systems and equipment, not elsewhere classified	**36997**	
other electrical products, not elsewhere classified	36998	partial
automatic garage door openers, electronic	**3699804**	
Measuring and Controlling Devices	382	partial
process control instruments	**3823**	
instruments to measure electricity	3825	partial
integrating instruments, electrical	**38251**	
test equipment for testing electrical, radio and communication circuits, and motors	**38252**	
measuring & controlling devices, not elsewhere classified	3829	partial
physical properties and kinematic testing equipment	**38292**	
nuclear radiation detection and monitoring instruments	**38294**	
Miscellaneous Fabricated Metal Products	349	partial
industrial valves	3491	partial
automatic regulating and control valves	34918	partial
all other actuation, including electric-actuated and electro-hydraulic actuated	---	
[collective]	**34918 11-21**[12]	
valve actuators	**3491823**	
regulator valves	---	
[collective]	**34918 25-49**[13]	
General Industrial Machinery	356	partial
general industrial machinery, not elsewhere classified	3569	partial
industrial robots, accesories, subassemblies, components, and parts	**35697**	

(Table 158 continued)

	Standard Industrial Classification Code	
Miscellaneous Manufactures	399	partial
manufacturing industries, not elsewhere classified	3999	partial
coin operated amusement machines	39992	partial
electronic games, arcade and amusement center type	**3999222**	
[other areas]	[------][14]	

Electromedical Equipment

Medical Instruments and Supplies	384	partial
X-ray apparatus and tubes	**3844**	
electromedical equipment	**3845**	

CONSUMER ELECTRONICS

Household Audio and Video Equipment	365	partial
household audio and video equipment	3651	partial
household and automobile radios, and radio phonograph combinations	**36511**	
television receivers, including television combinations	**36512**	
recorders, phonographs, and radio and television chassis	**36514**	
speaker systems, microphones, home-type electronic kits, and commercial sound equipment, including public address systems	36515	partial
bookshelf loudspeaker systems	**3651556**	
floor standing loudspeakers systems	**3651557**	
other loudspeaker systems	**3651568**	
Miscellaneous Electrical Equipment & Supplies	369	partial
magnetic and optical recording media	3695	partial
magnetic and optical recording media (unrecorded)	36950[8]	partial
audio range tape	---	
reels	**3695015**[9]	
cassettes	**3695017**[9]	
video tape	---	
cassette, 8 mm and 1/2 inch (12.7 mm)	**3695021**	
cassette, 3/4 inch (19 mm) and 2 inch (51 mm)	**3695023**[10]	
Electronic Components and Accessories	367	partial
electronic components, not elsewhere classified	3679	partial
electronic components and subassemblies, not elsewhere classified	36799[4]	partial
earphones and headsets, except telephone	**3679901**	
phonograph cartridges and pickups	**3679903**	
phonograph needles and styli	**3679905**	

(Table 158 continued)

	Standard Industrial Classification Code	
Dolls, Toys, Games and Sporting and Athletic Goods	394	partial
games, toys, and children's vehicles, except dolls and bicycles	3944	partial
electronic games and toys, excluding disks, tapes, and cartridges	39447[15]	partial
home electronic games	---	
for attachment to television receiver	**3944712**	
other	**3944714**	

ELECTRONIC-RELATED PRODUCTS AND SERVICES

Aerospace[16]

Aircraft and Parts	372	partial
aircraft	3721	partial
aircraft engines and engine parts	3724	partial
aircraft parts and auxiliary equipment, not elsewhere classified	3728	partial
Guided Missiles, Space Vehicles, Parts	376	partial
guided missiles and space vehicles	3761	partial
guided missile and space vehicle propulsion units and propulsion unit parts	3764	partial
guided missile and space vehicle parts and auxiliary equipment, not elsewhere classified	3769	partial

Automatic Controls, Industrial Apparatus, and Other Instruments[14]

Electrical Industrial Apparatus	362	partial
electrical industrial apparatus, not elsewhere classified	3629	partial
Laboratory Apparatus and Analytical, Optical, Measuring, and Controlling Instruments	382	partial
laboratory apparatus and furniture	3821	partial
environmental controls	3822	partial
fluid meters and counting devices	3824	partial
analytical instruments	3826	partial
optical instruments and lenses	3827	partial

Systems Integration and Computer Services[17]

Computer Programming, Data Processing, and Other Computer-Related Services	737	partial
computer integrated systems design	**7373**	
computer maintenance and repair	**7378**	
computer related services, not elsewhere classified	**7379**	

(Table 158 continued)

Motor Vehicles[16]

	Standard Industrial Classification Code	
Motor Vehicles and Equipment	371	partial
motor vehicles and car bodies	3711	partial

Electronic-Related Office Equipment[7]

	Standard Industrial Classification Code	
Computer and Office Equipment	357[18]	partial
calculating and accounting machines	3578[18]	partial
office machines, not elsewhere classified	3579[18]	partial
automatic typing and word processing machines (with memory)	**35792**[18]	

ENDNOTES

1. The selected components are: (1) capacitors, (2) resistors, (3) passive networks, and (4) TV and FM antennas and parts.

2. Optical fiber cable is not included in EIA's definition for the *1991 Electronic Market Data Book* but will be included in the definition for the 1992 edition.

3. This SIC code is included in the category "Other Components" in the *1991 Electronic Market Data Book*.

4. All seven-digit codes under SIC 36799 are included either under *Electronic Components* or under *Consumer Electronics*, except SIC 3679907 (home antennas) for which EIA uses data that it collects. All seven-digit codes that appear under SIC 36799 in *Current Industrial Reports* [MA36Q(89)] are included under *Electronic Components*, except SIC 3679907. The remaining three seven-digit codes under SIC 36799 (01, 03, and 05) are included under *Consumer Electronics* and appear in *Current Industrial Reports* [MA36M(89)].

5. Relays from almost all seven-digit SIC codes within 36251 were included in EIA's definition for the *1991 Electronic Market Data Book*. Those excluded were SIC 3625101 [industrial control relays (all voltages), not elsewhere classified] and three SIC categories for which the Bureau of Census did not report data in order to avoid disclosing figures for individual companies, due to the very small number of organizations reporting: SIC 3625131 [round and square can multipole airframe relays and contactors (both sealed and not sealed, all sizes)], SIC 3625169 [stepping switches, stepping and impulse relays], and SIC 3625175 [switchgear and protective relays]. The extent of inclusion of relays in the EIA definition for the 1992 edition will be determined later.

6. For calculators, no data has yet been reported by the Bureau of the Census. When it is reported, this SIC code will be included in the EIA definition.

7. The SIC codes under *Electronic-Related Office Equipment* will move to *Computers and Peripheral Equipment* in EIA's definition for the *1992 Electronic Market Data Book*.

8. Some seven-digit SIC codes under 36950 are changing locations from EIA's definition for the *1991 Electronic Market Data Book* to the definition for the 1992 edition. Others are being added in the 1992 edition. See the endnotes for individual seven-digit codes for the details.

9. SIC 3659019 is not included in EIA's definition in the *1991 Electronic Market Data Book* but will be included in the definition for the 1992 edition. It will appear under *Industrial Electronics* if it can be broken out separately from 3695015 and 3695017 with which it is jointly reported by the Bureau of the Census. Otherwise it will appear under *Consumer Electronics*.

10. SIC 3695023 is included in EIA's definition for the *1991 Electronic Market Data Book* under *Consumer Electronics*. It will be moved to *Industrial Electronics* in EIA's definition for the *1992 Electronic Market Data Book*.

11. SIC 3659025 is not included in EIA's definition in the *1991 Electronic Market Data Book* but will be included in the definition for the 1992 edition.

12. SIC 34918 11-21 includes all codes in that range that appear in "Selected Instruments and Related Products", *Current Industrial Reports*, No. MA38B(90)-1, Bureau of the Census, U.S. Department of Commerce (December 1991).

13. SIC 34918 25-49 includes all codes in that range that appear in "Selected Instruments and Related Products", *Current Industrial Reports*, No. MA38B(90)-1, Bureau of the Census, U.S. Department of Commerce (December 1991).

14. All SIC codes under *Automatic Controls, Industrial Apparatus, and Other Instruments* will be moved to *Industrial Electronics* in EIA's definition for the *1992 Electronic Market Data Book*. For all four-digit SIC codes affected, a percentage representation of U.S. products is included in EIA's definition for the *1991 Electronic Market Data Book*, according to the following three criteria: (1) Products containing 15 percent or more of electronic content are included at full value. (2) For products containing less than 15 percent but at least 5 percent of electronic content, just the electronic content is included. (3) For products containing less than 5 percent of electronic content, nothing is included.

15. SIC 3944712 and 3944714 are included in EIA's definition for the *1991 Electronic Market Data Book*. Their inclusion in EIA's definition for the 1992 edition is under review.

16. For the SIC codes under *Aerospace* and under *Motor Vehicles*, the EIA notes: "...we estimate the portion of the finished value of automobiles and aircraft that is attributable to the electronics intrinsic in the finished good. By this we mean the electronics in the finished product, which were constructed by the manufacturer of the finished good, solely for use in the finished good."

17. This section focuses on revenues from systems integration software services. Of the three types of software -- packaged, custom, and integral software -- only integral software is included in EIA's definition since it alone cannot be separated from the electronic hardware.

18. The electronic content of categories within SIC 357, and not included elsewhere in the definition, is included here. SIC 35792 is included in full. Most of the rest came from SIC 3578 which is included in part. In the 1992 edition, the same inclusions will be made but the data will appear under *Computers and Peripherals* under the heading *Other Electronic Office-Related Equipment*.

APPENDIX 3

DEFINITION OF

U.S. ELECTRICAL-EQUIPMENT

INDUSTRY

March 1992

Appendix 3

DEFINITION OF U.S. ELECTRICAL-EQUIPMENT INDUSTRY

in terms of the
Standard Industrial Classification System of the
Office of Management and Budget

SCOPE AND STRUCTURE

A definition of the U.S. electrical-equipment industry has been developed for use in this document. The major categories of products included in the definition, and excluded from it, are shown in Table 159.

Table 159
Definition of U.S. Electrical-Equipment Industry:
Major Categories

within definition	electrical supply equipment	
		generation
		transfer
		manipulation
		storage
	electrical conversion equipment	
		motion
		light
		electrolytic action
		heat
outside of definition	electrical applications equipment	
	electronic equipment	

The first category addresses products that are used to supply electricity. They generate, transfer, manipulate, and store electricity. Manipulation, as used here, means controlling electricity and converting it into other electrical forms (alternating current to direct current, for example). The second category of products is the bridge between those that supply electricity and those that apply it to practical purposes. The products in this second category convert electrical energy into the four principal forms of non-electrical energy used in practical electrical applications: motion (motors), light (lamps), electrolytic action (carbon and graphite electrodes), and heat (resistive devices). Electrolytic action is important in the industrial sector of the economy, but not in the commercial and residential sectors. The other forms are important in all three sectors. Essentially, the second category contains the *electrical components* that other products employ to apply electricity to practical purposes. This definition of the electrical-equipment industry ends with these components.

EXCLUSIONS

The definition excludes the products that employ the components for practical applications. Such products are highly diverse and include, for example, household appliances, transportation equipment, and manufacturing equipment. Products of this type are excluded from the definition for two principal reasons: (1) Most of them are as much the products of other industries as they are of the electrical-equipment industry; and (2) they are difficult to bound because electricity is used so widely.

Also excluded from the definition are electronic products. For the most part, they are the products that apply electricity in electrical form rather than as motion, light, electrolytic action, or heat. Electronic products can be thought of as addressing principally information rather than energy in its various forms. Electronic products are covered separately in the definition of the U.S. electronics industry in Appendix 2.

The definition of the U.S. electrical-equipment industry can be described in terms of the Standard Industrial Classification (SIC) System, as shown in Table 160. The definition has been designed to be distinct from the definition of the electronics industry in Appendix 2. That is, products classified under one of the two industries are not classified under the other. The structure of the SIC System is discussed in Appendix 1.

The definition in Table 160 begins at the three-digit level for simplicity. All three-digit codes are subdivided to the four-digit level to provide useful detail about what is included. If only a part of the products associated with a given code is included in the definition, then the word *partial* appears next to that code. Further subdivision is then shown in an effort to reach one code, or a set of codes, that describe products that are fully included in the definition. Fully included codes are shown in **bold** type. In some cases, there are no codes describing products that are fully included, so the word *partial* appears next to even the most detailed code (largest number of digits) shown. Endnotes provide additional details where relevant. Within major product categories, as defined by subordinate bold headings such as "Generation", the products are listed in order by ascending SIC code.

Representative electrical applications products that are excluded from the definition are shown in Table 161 beginning on page 445. The products are listed in order by ascending SIC code number to show their industry affiliation.

RELATIONSHIP TO NEMA PRODUCTS

The product content of this definition of the U.S. electrical-equipment industry was influenced by the list of SIC categories used by the National Electrical Manufacturers Association (NEMA) to describe the products that it addresses.[1] However, the definition provided here differs from NEMA's list in several ways. For example, NEMA addresses communications equipment and some medical equipment (such as magnetic resonance imaging equipment and X-ray imaging equipment) that can also be classified as electronic equipment. To prevent double counting, such equipment has been included only in the definition of the electronics industry in Appendix 2. Also, NEMA addresses some products, such as arc welding apparatus and electric lamp fixtures (as opposed to the lamps [bulbs] themselves), which are considered here to be electrical applications equipment and thus to be outside of the definition. Further, some products in the definition provided here do not appear explicitly in NEMA's list of included SIC categories, such as a variety of electrical heating elements.

However, these elements may be captured implicitly by other SIC categories or by NEMA’s own product data which extend beyond formal SIC categories.

LIMITATIONS

It is important to note that the definition in Table 160 does not capture every product that should qualify for inclusion in the U.S. electrical-equipment industry according to the characterization in Table 159. The reason for this is that some qualifying products are not broken out separately, or are not locatable at all, in the SIC System and therefore cannot be captured through that system. As an example, consider resistive heating elements. They fall under "heat" in the second category of products in Table 159. Three distinct cases arise for such elements:

(1) Resistive heating elements *are broken out separately* for small household appliances (like hair dryers), so those elements can be captured for those products.

(2) Resistive heating elements *are merged with other components* for some other product categories. Examples include major household appliances (like electric ranges) and industrial furnaces. For such products, the heating elements are included with other supportive components in collective categories for "parts". The fractional content representing resistive heating elements may be difficult to estimate.

(3) Resistive heating elements *are not isolatable* for other product categories such as residential space heating equipment, and thus are not represented explicitly in the definition.

Thus the full size of the electrical-equipment industry, when expressed strictly in terms of SIC codes as it is here, is likely understated.

The titles used in Table 160 are those from documents of the Bureau of the Census, except for parenthetical inserts in square brackets [...] which have been added for explanation.

Table 160
Definition of U.S. Electrical-Equipment Industry in Terms of SIC System

	Standard Industrial Classification Code	
ELECTRICAL SUPPLY EQUIPMENT		
Generation		
Electrical Industrial Apparatus	362	partial
motors and generators	3621[2]	partial
integral horsepower motors and generators, except for land transportation equipment	36212[3]	partial
a.c. (non-commutated) generators	---	
[collective for sizes from 0.75 to over 1000 kva]	**36212 51-59**	

(Table 160 continued)

	Standard Industrial Classification Code	
motors and generators for land transportation (including those used in associated control equipment)	36213[4]	partial
used for gasoline-electric and diesel-electric buses, trucks, locomotives, and rail cars	3621321[5]	partial
all other types of land transportation	3621341[6]	partial
prime mover generator sets, except steam (or hydraulic) turbine and electric motor-driven	**36214**	
fractional [horsepower] motor generator sets and other rotating equipment, including hermetics	36217[7]	partial
other motor generator sets	---	
a.c. and d.c. output rated less than 746 watts	3621721[8]	partial
other rotating equipment	---	
rate generators, resolvers, and combinations (less than 746 watts)	**3621753**	
all other rotating equipment	---	
rated at less than 746 watts	3621798[9]	partial
integral [horsepower] motor generator sets and other rotating equipment, including hermetics	36218[10]	partial
synchronous converters and double current generators, 746 watts and more	3621811[11]	partial
other motor generator sets	---	
a.c. and d.c. output rated at 746 watts or more	---	
a.c. output, based on the rating of the largest a.c. generator	---	
746 watts to less than 150 kw	3621832[12]	partial
150 kw and over	3621833[13]	partial
d.c. output, based on the rating of the largest d.c. generator	---	
746 watts to less than 170 kw	3621834[14]	partial
170 kw and over	3621836[15]	partial
other rotating equipment	---	
all other rotating equipment	---	
rated at 746 watts or more	3621898[16]	partial
parts and supplies for motors, generators, generator sets, and other rotating equipment (regardless of output rating)	36219[17]	partial
Electrical Industrial Apparatus	362	partial
carbon and graphite products	3624	partial
all other carbon and graphite products	36249	partial
brushes, contacts, and brush plates	---	
automotive brushes (starter, generator, and alternator)	3624911[18]	partial

(Table 160 continued)	Standard Industrial Classification Code	
Electrical Equipment & Supplies	369	partial
engine electrical equipment	3694[19]	partial
battery charging alternators, generators, and regulators	**36942**	
complete engine electrical equipment	36947[20]	partial
magnetos, magneto-dynamos, magnetic flywheels	**3694704**	
parts for engine electrical equipment	36949[21]	partial
armatures and field coils for alternators and generators	**3694907**	

Transfer (transmission and distribution)

Paints, Varnishes, Lacquers, Enamels, and Allied Products	285	partial
Paints, Varnishes, Lacquers, Enamels, and Allied Products	2851	partial
product finishes for original equipment manufacturers (OEM), excluding marine coatings	28512	partial
electrical insulating coatings	**2851253**	
Fabricated Rubber Products, Not Elsewhere Classified	306	partial
fabricated rubber products, not elsewhere classified	3069	partial
industrial rubber products, not elsewhere classified	3069C	partial
pressure sensitive tape, rubber-backed (including friction)	**3069C11**	
Miscellaneous Plastic Products	308	partial
plastic products, not elsewhere classified	3089	partial
electrical and electronic fabricated plastics products, excluding foam and reinforced plastics	30892	partial
household and commercial appliances	3089220	partial
other electrical and electronic, including wiring devices and parts	3089290	partial
reinforced and fiberglass plastic products, not elsewhere classified		
electrical and electronic	3089A12	partial
Pottery and Related Products	326	partial
Porcelain Electrical Supplies	3264	partial
porcelain, steatite, and other ceramic electrical and electronic products	32640	partial
[porcelain and steatite products]	**32640 10-41**	
Rolling, Drawing, and Extruding of Nonferrous Metals	335	partial
aluminum rolling and drawing, not elsewhere classified	3355	partial
aluminum and aluminum-base alloy wire and cable, except covered or insulated (including ACSR), produced in aluminum rolling mills	33551	partial
bare wire for electrical transmission	**3355111**	
aluminum cable, steel reinforced (ACSR)	**3355161**	

(Table 160 continued)

	Standard Industrial Classification Code	
nonferrous wiredrawing and insulating	3357	partial
aluminum and aluminum-base alloy wire and cable, except covered or insulated (including ACSR) produced in nonferrous wiredrawing plants	33571	partial
bare wire for electrical transmission	**3357111**	
aluminum cable, steel reinforced (ACSR)	**3357161**	
copper and copper-base alloy wire (including strand and cable), bare and tinned for electrical transmission	**33572**	
apparatus wire and cordage	**33576**	
magnet wire	**33577**	
power wire and cable	**33578**	
control and signal wire and cable	**3357C**	
building wire and cable	**3357D**	
other insulated wire and cable	**3357E**	
Electric Distribution Equipment	361[22]	partial
transformers, except electronic	**3612**	
Electric Lighting and Wiring Equipment	364	partial
current-carrying wiring devices	**3643**	
noncurrent-carrying wiring devices	**3644**	
Electrical Equipment & Supplies	369	partial
engine electrical equipment	3694[19]	partial
ignition harness and cable sets	**36941**	
complete engine electrical equipment	36947[20]	partial
ignition coils (all types)	**3694701**	
distributors (all types)	**3694702**	
other complete ignition equipment, including electronic ignitions	**3694705**	
parts for engine electrical equipment	36949[21]	partial
parts for ignition distributors	---	
distributor heads and rotors	**3694911**	
breaker point sets	**3694912**	
condensers (capacitors)	**3694913**	
other parts for engine electrical and/or electronic equipment	**3694919**	

Manipulation (control and electrical-to-electrical conversion[23])

Electric Distribution Equipment	361[22]	partial
switchgear and switchboard apparatus	**3613**	
Electrical Industrial Apparatus	362	partial
relays and industrial controls	3625	partial

(Table 160 continued)

	Standard Industrial Classification Code	
relays for electronic circuitry, industrial control, overload, and switchgear type	36251[24]	partial
industrial control relays (all voltages), not elsewhere classified	**3625101**	
electrical industrial apparatus, not elsewhere classified	3629[25]	partial
capacitors for industrial use (except for electronic circuitry)	**36291**	
rectifying apparatus	36292[26]	partial
electrical equipment for industrial use (coil windings [electrical], solenoids, surge suppressors, cathodic protection equipment, electrical discharge equipment, and other miscellaneous equipment for industrial use, not elsewhere classified)	**36293**	
Electrical Equipment & Supplies	369	partial
engine electrical equipment	3694[19]	partial
complete engine electrical equipment	36947[20]	partial
other complete electrical and/or electronic equipment for internal combustion engines, including complete engine control	**3694709**	

Storage

Electrical Equipment & Supplies	369	partial
storage batteries	**3691**	
primary batteries, dry and wet	**3692**	

ELECTRICAL CONVERSION EQUIPMENT (to non-electrical forms)

Motion (motors)

Electrical Industrial Apparatus	362	partial
motors and generators	3621[2]	partial
fractional horsepower motors, excluding hermetics	**36211**	
integral horsepower motors and generators, except for land transportation equipment	36212	partial
used in aircraft and spacecraft (excluding generators)	**3621204**	
a.c. (non-commutated)	---	
motors	---	
[collective: single phase, polyphase, and synchronous]	**36212 10-2H, 55**	
d.c. (excluding all arc welding generators, and battery charging generators for internal combustion engines)	---	
motors and generators	---[27]	

(Table 160 continued)

	Standard Industrial Classification Code	
[collective: mechanically and electrically commutated]	**36212 6A-84**	
motors and generators for land transportation (including those used in associated control equipment)	36213[4]	partial
used for trolley cars, trolley coaches, rapid transit cars, trolley locomotives, third-rail locomotives, multiple unit cars for railway service, and mining locomotives	**3621301**	
used for gasoline-electric and diesel-electric buses, trucks, locomotives, and rail cars	3621321[5]	partial
all other types of land transportation	3621341[6]	partial
fractional [horsepower] motor generator sets and other rotating equipment, including hermetics	36217[7]	partial
other motor generator sets	---	
a.c. and d.c. output rated less than 746 watts	3621721[8]	partial
other rotating equipment	---	
synchro-type components, less than 746 watts	**3621757**	
all hermetic motors	---	
5.5 inch stator core diameter, and smaller	**3621797**	
all other rotating equipment	---	
rated at less than 746 watts	3621798[9]	partial
integral [horsepower] motor generator sets and other rotating equipment, including hermetics	36218[10]	partial
synchronous converters and double current generators, 746 watts and more	3621811[11]	partial
other motor generator sets	---	
a.c. and d.c. output rated at 746 watts or more	---	
a.c. output, based on the rating of the largest a.c. generator	___	
746 watts to less than 150 kw	3621832[12]	partial
150 kw and over	3621833[13]	partial
d.c. output, based on the rating of the largest d.c. generator		
746 watts to less than 170 kw	3621834[14]	partial
170 kw and over	3621836[15]	partial
other rotating equipment	---	
all hermetic motors	---	
over 5.5 inch stator core diameter	**3621897**	
all other rotating equipment	---	
rated at 746 watts or more	3621898[16]	partial
parts and supplies for motors, generators, generator sets, and other rotating equipment (regardless of output rating)	36219[17]	partial

(Table 160 continued)

	Standard Industrial Classification Code	
carbon and graphite products	3624	partial
all other carbon and graphite products	36249	partial
brushes, contacts, and brush plates	---	
automotive brushes (starter, generator, and alternator)	3624911[18]	partial
other fractional horsepower brushes and and contacts	**3624913**	
other industrial brushes and contacts	**3624915**	
brush plates	**3624917**	
Electrical Equipment & Supplies	369	partial
engine electrical equipment	3694[19]	partial
cranking motors (starters)	**36943**	
parts for engine electrical equipment	36949[21]	partial
armatures, field coils, and drive-end housings for cranking motors	**3694901**	

Light[28]

Electric Lighting and Wiring Equipment	364	partial
electric lamps	**3641**	

Electrolytic Action

Electrical Industrial Apparatus	362	partial
carbon and graphite products	3624	partial
electrodes	36241	partial
electrodes for electric furnaces and electrolytic cell use, including paste for self-baked electrodes	---	
carbon [electrolytic applications only]	3624152[29]	partial
graphite [electrolytic applications only]	3624156[30]	partial

Heat[31]

General Industrial Machinery and Equipment	356	partial
industrial furnaces and ovens	3567	partial
electrical heating equipment for industrial use, not elsewhere classified, (except soldering irons) and parts and attachments	35675	partial
parts, attachments, and components for electrical industrial furnaces and ovens (including electric heating units sold separately)	---	
for space heaters	3567511[31]	partial
for other electrical	3567512[31]	partial

(Table 160 continued)

	Standard Industrial Classification Code	
Electrical Industrial Apparatus	362	partial
carbon and graphite products	3624	partial
electrodes	36241	partial
electrodes for electric furnaces and electrolytic cell use, including paste for self-baked electrodes	---	
carbon [thermal applications only]	3624152[29]	partial
graphite [thermal applications only]	3624156[30]	partial
Household Appliances	363	partial
household cooking equipment	3631	partial
electric ranges, ovens, and surface cooking units, equipment, and parts	36311	partial
parts and accessories for household electric ranges and ovens, such as burners, rotisseries, oven racks, broiler pans, etc. (sold separately)	3631185[31]	partial
electric housewares and fans	3634	partial
parts and attachments for electric fans and small household electric appliances	36349	partial
parts and attachments for other small electric appliances [other than household electric fans]	---	
electrothermal in operation	**3634913**	
household appliances, not elsewhere classified	3639	partial
household dishwashing machines, food waste, disposers, floor-waxing machines, trash compactors, and parts	36395	partial
parts for other household appliances, not elsewhere classified	---	
for electric water heaters	3639591[31]	partial
Electrical Equipment & Supplies	369	partial
engine electrical equipment	3694[19]	partial
spark plugs	**36944**	

OTHER ELECTRICAL SUPPLY OR CONVERSION EQUIPMENT

Electrical Industrial Apparatus	362	partial
carbon and graphite products	3624	partial
all other carbon and graphite products	36249	partial
all other carbon and graphite products, except refractories for electrical uses, including welding products,..., power tube and rectifier parts,..., illuminating carbons,..., resistance elements, etc.	**3624981**[32]	

Table 161
Representative Electrical Applications Equipment

	Standard Industrial Classification Code	
Construction, Mining, and Materials Handling Machinery and Equipment	353	partial
elevators and moving stairways	**3534**	
Metal Working Machinery and Equipment	354	partial
machine tools, metal cutting types [electrical only]	3541	partial
machine tools, metal forming types [electrical only]	3542	partial
power-driven handtools	3546	partial
power-driven handtools (electric)	**35461**	
welding apparatus	3548	partial
arc welding machines, components and accessories	**35481**	
arc welding electrodes, metal	**35482**	
resistance welders, components, accessories, and electrodes	**35483**	
Special Industrial Machinery, Except Metalworking Machinery	356	partial
pumps and pumping equipment	**3561**	
air and gas compressors	**3563**	
blowers and fans	**3564**	
industrial furnaces and ovens	3567	partial
electric industrial furnaces, ovens, and kilns	**35671**	
high-frequency induction and dielectric heating equipment	**35674**	
electric heating equipment for industrial use, not elsewhere classified	**35675**	
Refrigeration and Service Industry Machinery	358	partial
refrigeration and heating equipment	**3585**	
service industry machinery	3589	partial
commercial cooking and food warming equipment [electrical only]	35891	partial
commercial and industrial vacuum cleaners	**35893**	
Household Appliances	**363**	
household cooking equipment	**3631**	
household refrigerators and freezers	**3632**	
household laundry equipment	**3633**	
electric housewares and fans	**3634**	
household vacuum cleaners	**3635**	
household appliances, not elsewhere classified	**3639**	

(Table 161 continued)

	Standard Industrial Classification Code	
Electric Lighting and Wiring Equipment	364	partial
residential lighting fixtures	**3645**	
commercial lighting fixtures	**3646**	
vehicular lighting equipment	**3647**	
lighting equipment, not elsewhere classified	**3648**	
Electrical Equipment and Supplies	369	partial
electrical equipment and supplies, not elsewhere classified	3699	partial
Apparatus wire and cordage manufactured from purchased insulated wire	**36996**	
Electrical products, not elsewhere classified	36998[33]	partial
electric gongs, chimes, bells, etc.	**3699801**	
Christmas tree lighting outfits, including value of bulbs made or purchased and assembled in outfits	**3699802**	
insect killers	**3699803**	
other electrical garage door openers	**3699805**	
electric fence chargers	**3699806**	
electric outboard motors for boats (complete propulsion unit)	**3699807**	
other electrical products, not elsewhere classified	**3699819**	
electrical equipment and supplies, not elsewhere classified, not specified by kind	**36990**	
Railroad Equipment	374	partial
railroad equipment	3743	partial
locomotives, both new and rebuilt, and parts [electric only]	37431	partial
passenger and freight train cars, new, excluding parts [electric only]	37432	partial
street, subway, trolley, and rapid transit cars, all rebuilt railcars, and parts for all railcars [electric only]	37433	partial
Miscellaneous Transportation Equipment	379	partial
transportation equipment, not elsewhere classified	3799	partial
golf cards and industrial in-plant personnel carriers, self-propelled, and parts [electric only]	37993	partial

ENDNOTES

1. *Facts and Figures 1990*, National Electrical Manufacturers Association, p. 44 (1991).

2. All products under SIC 3621 are included in the definition under either generators or motors.

3. All products under SIC 36212 are included in the definition under either generators or motors.

4. All products under SIC 36213 are included in the definition under either generators or motors.

5. All products under SIC 3621321 are included in the definition under either generators or motors.

6. All products under SIC 3621341 are included in the definition under either generators or motors.

7. All products under SIC 36217 are included in the definition and contain generators and/or motors.

8. All products under SIC 3621721 contain both generators and motors and are included in the definition under those categories.

9. Assume SIC 3621798 contains both generators and motors. All products under SIC 3621798 are included in the definition under those categories.

10. All products under SIC 36218 contain generators and/or motors and are included in the definition under those categories.

11. All products under SIC 3621811 contain both generators and motors and are included in the definition under those categories.

12. All products under SIC 3621832 contain both generators and motors and are included in the definition under those categories.

13. All products under SIC 3621836 contain both generators and motors and are included in the definition under those categories.

14. All products under SIC 3621834 contain both generators and motors and are included in the definition under those categories.

15. All products under SIC 3621836 contain both generators and motors and are included in the definition under those categories.

16. All products under SIC 3621898 contain both generators and motors and are included in the definition under those categories.

17. All products under SIC 36219 address either generators or motors and are included in the definition.

18. All products under SIC 3624911 address either generators or motors and are included in the definition.

19. All products under SIC 3694 are included at one location or another in the definition.

20. All products under SIC 36947 are included at one location or another in the definition.

21. All products under SIC 36949 are included at one location or another in the definition.

22. All of the products under SIC 361 are included in either SIC 3612 or 3613.

23. One class of equipment that converts electricity to other forms of electricity employs motor-generator sets. This equipment is split between motors and generators. To prevent double counting, it is not listed under the "manipulation" category.

24. The balance of the relays under SIC 36251 is included in the definition of the electronics industry.

25. The balance of the products in SIC 3629 is considered *electronic*.

26. The part of SIC 3629 that is considered *electronic* is assumed to come from SIC 36292.

27. In spite of the inclusion of *generators* in this subheading, all of the products under this subheading appear to be motors. "Motors and Generators", *Current Industrial Reports,* Issue MA36H(89)-1, p. 4 (1989).

28. There are some illuminating carbons in SIC 3624981, but that code contains so many other carbon and graphite products that listing it under "Light" even as a "partial" may overstate its content relevant to "Light". Therefore, this SIC code is shown only under "Other Electrical Supply or Conversion Equipment" on page 444.

29. Since the electrodes are used either for electrolytic action or for heating, all products under SIC 3624152 are included between those two product categories.

30. Since the electrodes are used either for electrolytic action or for heating, all products under SIC 3624156 are included between those two product categories.

31. Most heating elements are not separately listed in the SIC system at this time. For example, there are no separate categories for resistive heating elements used in residential or commercial space heating, in major household appliances (electric ranges and water heaters, principally), or in industrial process applications.

32. The carbon and graphite products represented in SIC 3624981 here are contained in a great variety of products that fall in several of the categories above without any clearly dominant category.

33. Within SIC 36998, the only seven-digit code excluded is 3699804 for automatic garage door openers which are considered *electronic*.