

THE  
**COMPACT**  
**DISC**  
**HANDBOOK**

EINE EINFÜHRUNG



POHLMANN

# **THE COMPACT DISC HANDBOOK**

## **The Computer Music and Digital Audio Series**

John Strawn, *Series Editor*

### **Digital Audio Signal Processing**

Edited by John Strawn

### **Composers and the Computer**

Edited by Curtis Roads

### **Digital Audio Engineering**

Edited by John Strawn

### **Computer Applications in Music: A Bibliography**

### **Computer Applications in Music: A Bibliography, Supplement 1**

Deta S. Davis

### **The Compact Disc Handbook [Second Edition]**

Ken C. Pohlmann

### **Computers and Musical Style**

David Cope

### **MIDI: A Comprehensive Introduction**

Joseph Rothstein

William Eldridge, *Volume Editor*

### **Synthesizer Performance and Real-Time Techniques**

Jeff Pressing

Chris Meyer, *Volume Editor*

*This book has been printed digitally and produced in a standard design  
in order to ensure its continuing availability*

**OXFORD**  
UNIVERSITY PRESS

Great Clarendon Street, Oxford ox2 6DP

Oxford University Press is a department of the University of Oxford.  
It furthers the University's objective of excellence in research, scholarship,  
and education by publishing worldwide in

Oxford New York

Athens Auckland Bangkok Bogotá Buenos Aires Cape Town  
Chennai Dar es Salaam Delhi Florence Hong Kong Istanbul Karachi  
Kolkata Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi  
Paris São Paulo Shanghai Singapore Taipei Tokyo Toronto Warsaw  
with associated companies in Berlin Ibadan

Oxford is a registered trade mark of Oxford University Press  
in the UK and in certain other countries

Published in the United States  
by Oxford University Press Inc., New York

© 1992 and 1989 ny A-R Editions, inc.  
Originally published as *The Comapact*  
*Disc: A Handbook of theory and use*

The moral rights of the author have been asserted  
Database right Oxford University Press (maker)

Reprinted 2001

All rights reserved. No part of this publication may be reproduced,  
stored in a retrieval system, or transmitted, in any form or by any means,  
without the prior permission in writing of Oxford University Press,  
or as expressly permitted by law, or under terms agreed with the appropriate  
reprographics rights organization. Enquiries concerning reproduction  
outside the scope of the above should be sent to the Rights Department,  
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover  
and you must impose this same condition on any acquirer

ISBN 0-19-816327-4

*Cover artwork: Sony Recording Media Photo:  
Hing/Norton, Inc., New York, NY*

Everybody, including myself, was astonished to find that it was impossible to distinguish between my own voice, and Mr. Edison's re-creation of it.

—Anna Case, Metropolitan Opera soprano, 1915



# Contents

Preface .....	ix
Preface to First Edition .....	xi
Introduction .....	xiii
<b>1 Introduction to the Compact Disc</b>	
Introduction .....	1
Analog versus Digital .....	1
LP versus CD .....	4
Origins of the Compact Disc .....	8
For Further Reading .....	12
<b>2 Fundamentals of Digital Audio</b>	
Introduction .....	15
Sampling .....	15
Aliasing .....	22
Quantization .....	27
Dither .....	32
Pulse Code Modulation .....	35
A Digitization System .....	38
For Further Reading .....	45
<b>3 The Compact Disc System</b>	
Introduction .....	47
System Overview .....	48
Optical Storage .....	48
Error Correction .....	60
Cross Interleave Reed-Solomon Code (CIRC) .....	67
Data Encoding .....	74
Data Decoding .....	85
Subcode .....	90
For Further Reading .....	99

**4 CD Player Design**

Introduction .....	103
Player Overview .....	103
Pickup Design .....	107
Digital Signal Processing .....	119
Filter Design .....	121
Digital Filters .....	127
Digital-to-Analog Converters .....	132
Output Sample-and-Hold .....	140
A Chip Set .....	142
Low-Bit D/A Conversion .....	147
Serial Transmission Format .....	157
For Further Reading .....	163

**5 Practical Concerns**

Introduction .....	167
Design Considerations .....	167
Specifications .....	178
Buying a Player .....	195
Player Notes .....	200
Disc Notes .....	204
For Further Reading .....	211

**6 Diverse Disc Formats**

Introduction .....	213
CD-ROM .....	213
CD-I .....	221
Photo CD .....	237
CDTV .....	242
DVI .....	243
CD-V .....	246
CD-R .....	249
Erasable Discs .....	255
Mini Disc .....	261
CD + G/M .....	267
CD-3 .....	273
For Further Reading .....	273

**7 Disc Manufacturing**

Introduction .....	277
Pre-Mastering .....	277
Disc Mastering .....	286
Electroforming .....	293
Injection Molding .....	295
Metallization and Spin Coating .....	299
Quality Control .....	302
Alternative Replication Methods .....	310
For Further Reading .....	314
<b>Glossary of Technical Terms .....</b>	317
<b>Index .....</b>	341

# Preface

There are about a billion reasons that prompted me to prepare a second edition of this book. Specifically, about a billion CDs are sold every year, and the number is still climbing. Exceeding the expectations of even its most ardent supporters (like me), the compact disc has become one of the most successful consumer electronics products ever introduced. The eager worldwide marketplace has encouraged rapid development of CD technology and spawned entirely new applications for the dimpled disc. If you read a magazine article or two on CD a few years back, and you think you're up to speed on the disc, you're sadly mistaken.

It is the aim of this second edition to celebrate the compact disc's tenth birthday and document this decade's swift series of events. Specifically, this edition more fully describes the laboratory origins of the CD and elaborates on the electronic and optical principles underlying the format. It explains technology weaknesses such as converter low-level nonlinearity, exposes placebos such as disc stabilizers, and suggests new testing procedures to measure problems or lack of them. It discusses emerging digital signal processing applications in areas such as low-bit conversion, noise shaping, and data reduction. It also introduces new formats such as Photo CD, CD-I, CD-R, and erasable disc formats including the Mini Disc. In addition, several structural changes have been incorporated into this edition, rearranging existing material to facilitate access. Furthermore, in an effort to further clarify the presentation, literally every page has seen changes. More obviously, anyone familiar with the first edition will notice that many new pages and figures have been added. Even the references at the end of each chapter and the glossary have been thoroughly updated. In short, this is no minor facelift.

Finally, I should note that I am increasingly encouraged by the conversations I overhear in record stores these days. Hardly anyone still confuses CDs with LPs, partly because consumers are more knowledgeable

of digital audio technology and partly because LPs have all but vanished. Of course, I would never take schadenfreude delight in analog technology's quick demise, heh, heh.

Ken C. Pohlmann

# Preface to First Edition

A bit of innocent eavesdropping prompted me to write this book. I was browsing through CDs in a record store when I overheard someone explaining the compact disc to his buddy. Knowingly, he told him, “It’s the same as an LP, only it’s silver.”

Although I didn’t say so at the time, that remark bothered me considerably. In an unknowing way, his explanation was adequate. The compact disc is indeed a counterpart and successor to the long-playing record, but it’s also much more. The sophistication of the technology underlying the CD system should command considerably more understanding and respect than is reflected in a simple comparison to a mechanical groove.

Along those lines, I decided that a book might help illuminate the technological beauty of the CD. My goal was to explain things without generating more confusion. There is enough of that surrounding high technology already. The result, of course, is now in your hands. When you have finished reading, I hope your impression will be that the book was very simple, even though you will have learned a great deal about compact disc technology. A good teacher is just like a good nurse with a hypodermic needle—you never feel the pain.

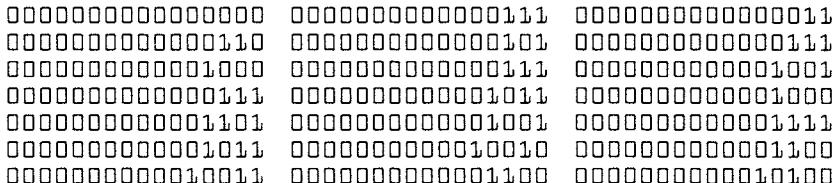
There is at least one philosophical question that I will not try to answer directly: Is digital audio inherently better than analog? This question is similar to asking whether or not there is life after death. Each of us may ascertain the answer, but not until later.

Meanwhile, I hope this book helps to expand your insight into the CD. And the next time you’re in a record store, please try not to embarrass yourself.

Ken C. Pohlmann



# Introduction



00000000000000 000000000000111 000000000000011  
0000000000000110 0000000000000101 0000000000000111  
0000000000001000 0000000000000111 0000000000001001  
0000000000000111 0000000000001011 0000000000001000  
0000000000001101 0000000000001001 0000000000001111  
00000000000001011 00000000000010010 0000000000001100  
00000000000010011 00000000000001100 00000000000010100

First, a word of explanation. From its Dutch and Japanese origins in June 1980, the compact disc system has prospered beyond the wildest dreams of its inventors. The original audio CD was one of the most successful new electronic products ever introduced; everyone was surprised by its rapid acceptance by music lovers. Over the first few years the biggest news concerned shortages, particularly of discs themselves. Then the disc caught the attention of data lovers as well. CD-ROM quickly started its own acceleration into the market of mass storage. As if that wasn't enough, CD-I's future looks similarly promising as an interactive audio-video medium. Of course, while reading data is fun, writing data is even more fun. Thus a standard for write-once CDs was introduced, and erasable CD technology has been developed. Furthermore, CDs with graphics and MIDI information and mini-CDs were also introduced into the marketplace. In short, the compact disc family is a hit and already encompasses more topics than one book could comprehensively cover.

Exactly how important is it to bone up on a new technology such as the compact disc? How significantly will it influence our everyday lives, and for how long? There are a number of factors that determine the lifespan of a technology. Manufacturing cost, product performance, market penetration, user boredom, and innovative competition all move technological evolution forward and heighten expectations. As far as the CD family goes, it is safe to say that the future looks bright indeed. Its market share continues to expand, and new and consequential applications are undoubtedly waiting to be discovered by insightful entrepreneurs. Of course, no technology, except the very primitive kind, lasts forever, and someday

(perhaps relatively soon) the CD will be only a curiosity. Fittingly, the discs themselves should long outlast their usefulness.

Meanwhile, this handbook covers all the fundamentals of the CD and should prove useful to anyone delving into this technology, especially for the first time. Hopefully, it will answer all your questions about audio CDs and pave the way to a greater understanding of the derivative CD formats.

This little opus, composed of seven chapters, starts with the basics and proceeds onwards and upwards. Theoretical topics are treated expeditiously, and mathematics have been avoided almost entirely. The emphasis is on practical, understandable, useful information.

Chapter One is an introduction to the fundamentals of digital audio theory. Digital signals are contrasted with the analog variety, and the pros and cons are discussed. Just for fun, a brief history of the events leading to the CD is chronicled. In addition, the compact disc is weighed against its predecessors. The first chapter, like all the chapters, concludes with a list of references for further reading.

Chapter Two begins our technical discussion. Sampling and quantization, the key analytical methods of digitally interpreting analog signals, such as audio waveforms, are presented. Aliasing, a negative attribute of sampling, and dither, a panacea for problems arising from quantization, are examined. Pulse code modulation, a particularly clever method of encoding digital audio data, is presented. Finally, a practical audio digitization system is used to consolidate all the theory.

Chapter Three examines the technical theory underlying the compact disc system. The encoding process leading to the plastic disc is examined. The disc itself is an impressive piece of handiwork, particularly at the microscopic level, so the intricacies of data pits are presented. The use of error correction is unprecedented in audio storage; the methods used to safeguard the CD's data are discussed. The bit stream from the disc is processed just as heavily, and oppositely, to the input bit stream. The signal processing circuitry composing every CD player is examined. The chapter concludes with a look at the non-audio subcode data, which is stored on every CD and used to control playback.

Chapter Four zeros in on CD player design, starting with the laser pickup as it reads data from the disc at a rate of 4.3218 million bits per second. To assist the pickup in its difficult task, electro-mechanical circuits are used for focusing the beam and keeping it on track. Digital filters must be used to process the waveform before it is presented to the ears. Digital filters and digital signal processing in general are carefully introduced. The digital-to-analog converter's job is to transform data back into an audio waveform. These output stages are summarized with a look at a specific chip set. The chapter concludes by examining low-bit D/A conversion and the AES/EBU and SPDIF standards, serial transmission formats used to convey data from one digital audio device to another.

Chapter Five brings the discussion back to the macroscopic level of the consumer. User notes are presented on a variety of topics. Different player designs are examined and critiqued. Specifications provide a means to distinguish good players from the merely mediocre; measurement techniques are examined. Tips are given to assist in purchasing a CD player. Player care is also important, including the dos and don'ts of preventive maintenance. Your disc collection will probably represent a bigger investment than the player; it is wise to take care of your discs and to be able to evaluate any defects that may be present.

Chapter Six carries the discussion to the diverse versions of the compact disc system. The audio CD was only the first of several CD family members. CD-ROM uses the CD's vast storage potential for non-audio applications, such as data bases and software programs. CD-I merges video with audio on CDs. Imagine a new kind of publishing: combined illustrated, printed, and talking books. DVI provides yet another format for interactivity, and the Photo CD format brings high-quality photographic images to the disc. CD-V merges the CD's digital audio format with that of high quality analog video optical storage. The CD-R format offers users the chance to permanently record their own data. Several fully recordable/erasable optical media, including magneto-optics and phase-change technology, allow users to record and erase data. Although incompatible with the CD, the MD format promises to expand optical disc applications. CD + G/M utilizes subcode storage capacity to provide graphics and MIDI output from audio CDs. CD-3 is a mini-sized CD for applications where shorter playing times are sufficient.

Chapter Seven is an examination of the technology of compact disc manufacturing. The discussion concentrates on the foremost production method, injection molding of polycarbonate plastic. Preparation of master tape, the mastering of the glass disc, the production of pressing molds, injection molding, metallization, spin coating, printing, packaging, and quality control are all discussed. Alternative manufacturing methods may simplify the task of making discs; these new methods are explained.

A glossary containing definitions of key technical terms and abbreviations rounds out the presentation.

In summary, this book attempts to cram as much pleasant and useful information into the smallest possible space, rather like the compact disc itself.

Oh, the bits that begin this introduction are from a compact disc recording of Beethoven's *Ninth Symphony*—as we shall see, a work that played an important role in the development of CD technology. It's a little difficult to catch the melody since the bits represent only the first 200 microseconds of the pianissimo opening of the symphony. It gets a little more interesting when the chorus joins in, about five billion bits later.



## *Chapter One*

# Introduction to the Compact Disc

### **INTRODUCTION**

Let's try a conceptual experiment. From your conceptual seat in your conceptual concert hall, you lean forward in anticipation as the conceptual orchestra begins its conceptual performance of Beethoven's *Tenth Symphony*. (No, I'm not conceptualizing; he really did start a tenth.) Your job is to write down all the information you hear. Ready? Begin!

Whew! After only a few bars, you give up. Even writing down the score in musical notation is overwhelming, much less recording all the timbres, aesthetic considerations, hall acoustics, and so forth.

We have to conclude that music is a surprisingly complex phenomenon; it is filled with information. To store it, we require a system which can deal with incredible amounts of information. It is not surprising that historically the latest and highest technology has always been utilized to make recordings, because only the best technology satisfies our current expectations of what a good recording should sound like. As higher technology was devised and pressed into service, our expectations were redefined. All of which brings us to the topic at hand: digital audio and the compact disc.

### **ANALOG VERSUS DIGITAL**

To understand the nature of analog and digital systems, along with their differences, let's try another conceptual experiment. Suppose that you are stationed in Antarctica. (Since I'm writing this book in Miami in July, this concept has particular appeal to me.) The Audubon Society has charged

you with the important mission of noting the effect, if any, of barometric pressure on the mating habits of penguins.

A barometer measures changes in atmospheric pressure. Attached to the barometer is a recording device, a cylinder with graph paper attached to it, and a pen with ink (and anti-freeze). As the atmospheric pressure changes, the pen traces a line on the slowly rotating cylinder. At the end of every day, the cylinder has completed one rotation, and you change paper. You are thus left with a graph of barometric pressure over time.

What you have is an analog recording. It is analog because it is a continuous representation, an analogy or a model of the actual phenomenon. It's very nice, but its accuracy is limited by numerous factors, including the precision of the cylinder's rotation and the flow of ink. Its usefulness is also limited by the fact that its representation is graphical. For example, to communicate your findings to the Audubon Society headquarters, you would have to fax the original graph, and the copy might not be as accurate as the original.

You figure there might be an alternative, and there is. Instead of relying on the graph paper, you decide to document individual readings themselves. You design a device that reads the pressure from the barometer every minute and prints out the number. At the end of a day, there are 1,440 numbers, representing the changing pressure.

What you have is a digital recording. It is digital because the signal is subjected to measurement at discrete points in time, and the information is stored as discrete numbers. It's very nice because the numbers are inherently more robust than the analog graph, and if your digital device is well designed, the numbers are probably more precise. In addition, it is far easier to communicate your findings; simply reading the numbers over the radio would do the job. Moreover, when the scientists at headquarters write down the numbers, they will have an exact copy of your results.

Suppose, however, that the scientists back at headquarters squawk that the numbers are not as good as a graph and complain that there is no record of pressure between each measurement. You could point out that you aren't as stupid as you look. Atmospheric pressure can change only so fast. For example, suppose the barometric pressure was 30.034 inches at 11:01:00 P.M., and 30.036 inches at 11:02:00 P.M. Although you didn't take a reading at 11:01:30 P.M., you can calculate that the pressure was 30.035, and not, say, 29.022. The atmosphere would not work like that. Thus by knowing how fast things change, you can sample often enough to obtain complete information about what's happening.

On the other hand, an event outside the interest of our barometric experiment, such as the overpressure created by a nearby penguin exploding at precisely 11:01:30 P.M., would not be documented. But that doesn't affect our study. By defining how fast things of interest change, we can ignore those events which happen faster.

At any rate, satisfied that the barometric pressure will be accurately documented, you can get down to the primary scientific task of watching the penguins.

Our conceptual experiment illustrates fairly well the relevant differences between analog and digital representations. In fact, barometric pressure is a close analogy to audio signals, since both are simply pressure changes in air. Of course, changes in barometric pressure would produce a very low frequency (about 0.00001 Hz), but if you speeded it up about 100 million times, it might sound musical, or at least like punk rock.

Back to reality. Let's nail down our comparison between analog and digital systems. The principal distinction lies in the way they represent information. Digital information can exist only in pieces, as discrete values, as numbers. This is vastly different from analog information, where one continuous, infinitely indivisible value is recorded.

Analog and digital systems thus differ considerably. There is no doubt about that. Nevertheless, the basic question has still eluded us: Why digital? At first glance, the use of digital technology for audio purposes seems very cumbersome. After all, we must convert sound into a series of numbers, each of which must accurately describe the sound at that instant in time. We must store these billions of numbers and then convert them back into sound to hear what's going on. That's a lot of work. Moreover, since analog audio technology seemed perfectly adequate for a hundred years, is it really necessary to replace analog with digital?

The answer is this: Sure, digital audio is a lot of hassle, but it's worth it. One justification for digital audio lies in the very nature of its signal. Sound is an analog phenomenon, and so is noise. An analog audio device cannot distinguish between them; hence, the noise of an analog signal is the sum total of all the noise introduced in its path. For example, every time an analog tape is copied, its noise increases. The numbers composing a digital signal will carry an error, introduced when they are first selected, but they are impervious to noise; for example, digital copying does not add noise. A digital number cannot become noisy; it is right or wrong.

Analog reproduction is more frail than digital. An analog system introduces distortion as it attempts to convey the exact analog nature of its information. In contrast, a digital system, at least philosophically, has an easy job. It must be able to distinguish only between 1s and 0s to reproduce the signal. The only theoretical limitations are those dictated by the quantity and accuracy of the numbers. In other words, with digital we can more precisely manipulate and process information and thus achieve a more accurate result.

Along similar lines, digital audio is advantageous because its signal is robust. It is a cruel world out there, and under adverse conditions a digital signal suffers less degradation than an analog signal. Moreover, a digital signal suffers no degradation at all until conditions have deteriorated

beyond a known level. The design performance of a digital system is designed into its circuits. Performance is thus always a known, defined quantity.

Another justification for digital audio is its consistency of performance. Analog devices can work quite well when they are new, but frequent adjustment and maintenance are necessary to ensure consistent performance. Because digital systems often permit a higher degree of circuit integration, they exhibit greater long-term accuracy, with less performance variation or failure. In fact, when a digital system fails, the problem is often an analog part inside.

Digital circuits are also very efficient because logical functions and operational features can be easily designed and implemented. Moreover, such functions and features can be quickly altered or updated; often the hardware circuit remains unchanged, and a change in programming achieves the new result.

Finally, digital audio's sonic performance is excellent. First-generation digital audio products rivaled the result of a century of analog evolution, and digital's evolution has now surpassed analog's best efforts.

Of course, digital audio isn't perfect—far from it. Its complexity breeds considerable challenges that are not found in analog systems. To use and appreciate such equipment properly, a fair amount of technical understanding is required, as demonstrated in the following pages.

## LP VERSUS CD

Many methods of audio storage have evolved since Edison made the first audio recordings in 1877 on a cylinder covered with tin foil. Early acoustical recordings were made on wax cylinder and shellac discs, and many electrical recordings used 78 rpm and long-playing records. Subsequently, numerous magnetic tape formats were developed. However, all of these audio systems shared identical foundations; they recorded and reproduced an analog signal by using a mechanical or electrical pickup in contact with the medium. Overall, this technology is now a mature one and has virtually exhausted the possibilities available within the limitations of analog master tape, phono cartridges, analog circuitry, motors, and mechanical systems.

The difference between analog and digital audio technologies can be illustrated by a close look at a long playing record and a compact disc, illustrated in figure 1.1.

The LP stores its information as an analog groove. Variations in its side-to-side amplitude and depth represent the original audio signal. If someone hit a big drum, the groove would take a big swing. The frequency of the drum sound determines the frequency of the groove's swing, and the amplitude of the drum sound determines the amplitude of the

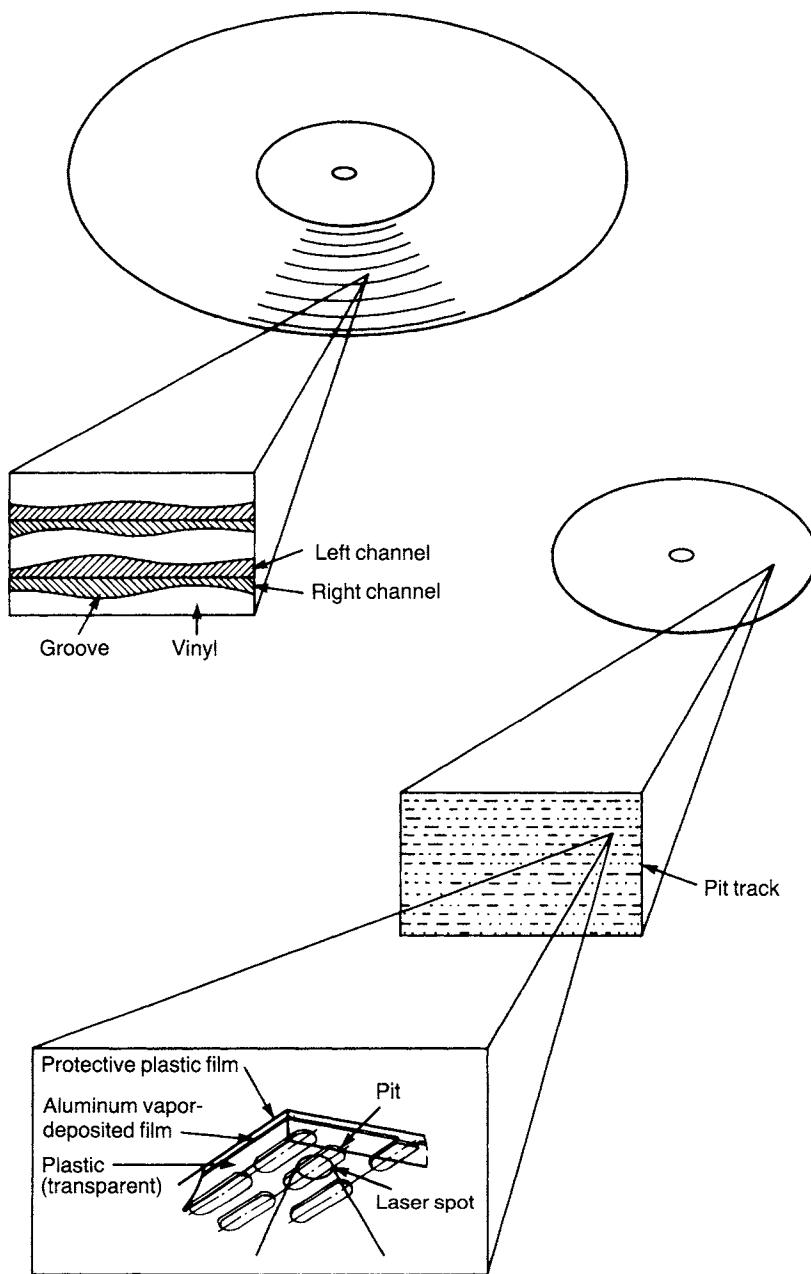
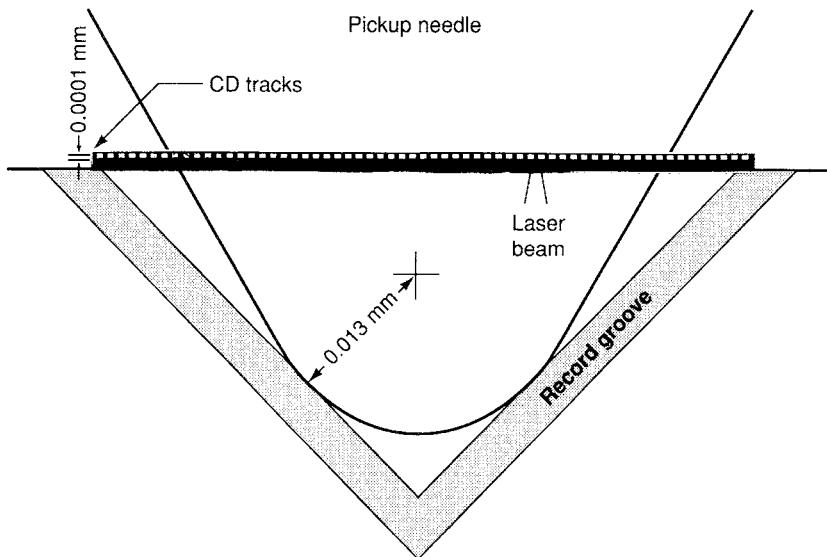


Figure 1.1 LP grooves, storing analog information, contrasted with CD pits, storing digital information.



**Figure 1.2** *The dimensions of compact disc tracks are small compared to those of an LP microgroove.*

swing. In short, the groove is a physical analogy to the original sound wave. Left and right audio channels are stored on either side of the groove walls. To reproduce the information, a stylus runs through the groove and the phonograph cartridge converts the stylus's mechanical movements into an electrical signal which is amplified. Analog tapes follow the same principles, but, instead of being stored as groove undulations, the analog signal is stored in magnetic orientations of ferrite particles.

The CD stores its information digitally. The length of its data pits represents a series of 1s and 0s—numbers (binary bits, actually) which represent the original audio signal. If someone hit a big drum, the numbers would be different than if there was no drum. Generally, the frequency of the signal is reflected in the rate of change in the numbers, and the amplitude is represented by the magnitude of the numbers. Both audio channels are stored along the same pit track. To read the data, we use a laser beam light source. The numbers are read from the disc as flashes of light and are used to reconstruct the original audio signal. Because nothing (except light) touches the pits, they never wear out.

In practice, the compact disc offers numerous advantages over previously existing audio storage media such as the LP record: The LP must be handled with care to preserve its mechanical groove; groove wear is

audible; tracking with a stylus damages the groove; and performance specifications are limited. For example, the signal-to-noise ratio is about 60 dB, channel separation is about 30 dB, and frequency response is not flat.

Phonograph turntables have inherent problems associated with analog devices. The mechanical transport introduces audible artifacts such as wow and flutter because the medium's speed is directly linked to reproduction quality. Electrical circuitry is prone to aging and to temperature drift and instabilities. Considerable phase shift is introduced by analog circuits, and considerable signal equalization is required. The stylus suffers wear and requires periodic replacement to maintain fidelity and prevent accelerated damage to the medium. Turntables are usually operated through mechanical controls that are limited in utility and prone to wear and tear.

The compact disc is designed for better performance in these areas. A CD is more robust because the data surface is embedded inside the disc itself. Because data is read by a focused laser beam, there is no degradation from repeated playings, and the effect of dust and surface damage is minimized. Digital signal processing permits use of error correction and other techniques to improve the reliability of the stored data.

The importance of digital error correction is apparent when we consider the question of scale, which is illustrated in figure 1.2. An LP microgroove is extremely narrow, about the width of a human hair. However, sixty laps of a CD pit track would fit into one LP groove. The audio data rate from that track would be 1.4112 million bits per second with an overall data rate of 4.3218 million bits per second.

The volume of numbers stored on a compact disc necessitates error correction. An analog signal presents basic conceptual problems for any attempt to diagnose and correct errors in the signal. If a groove is distorted, or the record scratched, the turntable's stylus cannot distinguish the defect from the real thing. A digital system easily accommodates and desperately requires error protection. The density of information on a CD is about 100 times greater than that of an LP. With three billion or so pits crammed onto its surface, a CD is vulnerable to defects and scratches. Fortunately, the data may be recorded with error protection codes which permit detection, correction, or concealment of errors. A CD player can thus correct disc errors, whereas a turntable cannot.

A compact disc player offers numerous other improvements over turntables. The transport's speed variation does not affect the quality of audio reproduction; data conversion is independent of variations in disc rotational speed, hence wow and flutter are negligible. Digital circuitry is more immune to aging and temperature problems. A CD player with a digital oversampling filter introduces negligible phase shift. The laser pickup is self-regulating and has great longevity. In addition, the pickup can never damage the medium. Specifications are improved; for example, signal-

to-noise ratio and channel separation are both over 90 dB, and frequency response deviations are within 0.1 dB. The CD also offers longer playing time, smaller size, programmability of tracks, and a subcode for display, control, and optional user information. Integration of circuitry permits portability of players and low manufacturing costs.

Digital applications extend far beyond the task of storing audio data; digital logic, including the use of microprocessors, controls the operation of all CD players. Turntables could also benefit from digital control, but the marriage of stored digital data and digital control offers many advantages. For example, the data on a compact disc not only stores the music but also directly controls the rotating speed of the disc. That would be tough for an analog storage device. Finally, digital control circuits permit flexibility of operation, including programmability and a simpler mechanical interface.

In summary, the introduction of the compact disc revitalized audio technology by changing three fundamental design criteria. First, digital data is used to store and process the audio signal. Second, error correction is employed to make the stored data more robust. Third, the compact disc system uses an optical, non-contact pickup.

Part of the compact disc's strength lies in its noble lineage. Its technology embodies over 100 years of scientific and engineering advances. With that in mind, a quick look at the origins of the compact disc may help to establish a better perspective on this technology.

## ORIGINS OF THE COMPACT DISC

"The exaggeration of sibilants by the new method is abominable, and there is often a harshness which recalls the worst excesses of the past. The recording of massed strings is atrocious." Clearly, not everyone always agrees that advances in audio technology are a good thing. Change is not always welcome, particularly when it occurs with something as intimately and subjectively felt as the perception of musical sound. That was certainly true when Compton Mackenzie, editor of *The Gramophone*, described the new electrical recording process in 1925.

Understandably, change is hard to accept. The introduction of any new technology prompts both praise and condemnation, both of which are sometimes misguided because the context of the technology is misunderstood. The compact disc system is both an evolutionary and revolutionary step in the history of audio technology—nothing more and nothing less. To appreciate the compact disc fully, it may help to consider a very brief chronology of events in its development. Several references at the end of this chapter contain extensive information on the history of audio.

Significant events in the nineteenth century set the stage for digital audio as we know it today. In 1841 Augustin-Louis Cauchy proposed an early sampling theorem. In 1842, Charles Babbage proposed his analytical engine for performing and storing calculations. Construction of this first computer was not accomplished until 1991, when researchers used Babbage's original design to build a working model. In 1854 George Boole published *An Investigation Into the Laws of Thought*, which contained the axioms of switching algebra, the basis of digital circuits. In 1855 Leon Scott de Martinville invented the phonautograph, a machine for recording vibrations using a carbonized paper cylinder and a hog's bristle. In 1876 Alexander Graham Bell publicly demonstrated the telephone. A breakthrough, and the true birth of audio reproduction, occurred in 1877 when Thomas Alva Edison invented the analog phonograph. Ironically, this occurred while he was experimenting with a device for storing digital data—a telegraphic code repeater. Edison's wax cylinder was quickly challenged in 1887 as Emile Berliner invented the audio disc.

The opening years of the twentieth century saw a number of important inventions in the new field of electricity. In 1904 John Ambrose Fleming invented the diode vacuum tube, the first active circuit component. In 1907 Lee DeForest invented the triode amplifier vacuum tube, thereby initiating the modern era of electrical recording and reproduction. The 78 rpm record was introduced in 1915. In 1922 J. R. Carson published a mathematical analysis exploring the nature of time sampling in communications. A critical observation occurred in 1928 when Harry Nyquist published "Certain Topics in Telegraph Transmission Theory," which contains proof of the sampling theorem—the basis of digital audio. In the same year R. V. L. Hartley stated the information capacity of an ideal channel, and 33½ rpm records were demonstrated. In 1937 A. Reeves invented pulse code modulation, and H. Aiken of Harvard proposed the design of an electromechanical computing machine to IBM.

The Second World War stimulated development of computing devices and electrical components. In 1943 the Electronic Numerical Integrator and Computer (ENIAC) was developed at the University of Pennsylvania by the U.S. Army. This dinosaur weighed thirty tons. In 1947 the first magnetic tape recorders were sold in the United States. In 1948 John Bardeen, William Shockley, and Walter Brattain invented the transistor in Bell Laboratories. In the same year Claude E. Shannon published "A Mathematical Theory of Communication," containing work on information theory, including a measure of information and the capacity of a data channel. In 1949 Peter Goldmark developed the microgroove LP record, and 45 rpm records were introduced.

In 1950 Richard W. Hamming published work on error detection and error correction codes, another critical piece in the subsequent development of digital audio. In 1952 D. A. Huffman described a minimum re-

dundancy coding scheme. In 1958 C. H. Townes and A. L. Schawlow invented the laser, and stereo LPs were marketed. Also in 1958, Jack S. Kilby invented the integrated circuit while in the employ of Texas Instruments. In 1959 A. Hocquenghem published error correction codes. In 1960 R. C. Bose and D. K. Ray-Chaudhuri published binary group error correction codes. Early computer music experiments took place at Bell Laboratories and other facilities. I. S. Reed and G. Solomon published a multiple error correction code, later to be incorporated into the compact disc. T. H. Mainan demonstrated the first working model of a laser.

By the 1960s, the essential components required to build a digital audio recorder/reproducer were in place. In 1967 the NHK Technical Research Institute publicly demonstrated a PCM digital audio recorder with 30 kHz sampling and 12-bit companding using a 1-inch helical scan video recorder for the storage medium. In 1969 Sony Corporation demonstrated a PCM digital audio recorder with 47.25 kHz sampling and 13-bit quantization, using 2-inch video tape for storage.

In 1969, Dutch physicist Klaas Compaan had a brilliant idea that, after a dozen years of research and development on the part of two diverse manufacturing conglomerates, would result in one of the most successful consumer electronics products of all time—the compact disc. He learned that engineers at RCA had devised an inexpensive way to manufacture holograms by using a master stamper with microscopic formations to press copies. He believed that the technique could be used to produce discs holding video images. He related the idea to his colleague Piet Kramer in 1970, and, together in the research laboratories of Philips in Eindhoven, they completed a prototype glass disc with a series of 1-millimeter square black and white images that could be projected on a screen. They decided it would be more efficient to record a video signal, rather than images themselves, and to use a track of dimples to encode the analog FM signal. Moreover, they found that a laser source was needed to recover the signal. In July 1972 they publicly demonstrated a color prototype.

Meanwhile, in 1971 large scale integrated (LSI) circuits such as the microprocessor were introduced by Intel and other manufacturing companies. The first digital audio device, the digital delay line, was introduced into recording studios. In 1973 Nippon Columbia and the BBC began using digital recorders for master recording.

Philips engineers began contemplating an audio application for their new video disc system. Forty-eight hours of music could be recorded on a 12-inch disc, but Lou Ottens, a Philips director who had helped develop the compact cassette in the 1960s, knew that small size was key and named the new product the compact disc to drive home his point. But further experiments revealed that the FM signal simply wasn't good enough. It was decided to pursue a digital code; within months a prototype disc with a sampling frequency of 44,000 samples per second was playing through

a 14-bit digital-to-analog converter, delivering monaural audio with a signal-to-noise ratio of 80 dB, a major improvement over the LP record. There was a problem, however. The digital code was particularly sensitive to errors caused by dust or scratches on the disc. Further work was needed.

Meanwhile in 1976 Sony experimented with a large-diameter optical audio disc. In 1977 Mitsubishi, Sony, and Hitachi demonstrated digital audio disc prototypes at the Tokyo Audio Fair. Also in 1977, JVC developed a digital audio processor. In 1978 Philips introduced its video disc player, using optical readout. Also in 1978, Sony marketed the PCM-1600 and PCM-1 digital audio processors, as well as stationary head digital tape recorders for professional use. The Digital Audio Disc Convention was held in Tokyo; thirty-five manufacturers attended.

In 1978 Philips management decided that the compact disc project merited the establishment of a worldwide standard, much like the LP. However, this would require a major corporate effort as well as collaboration between manufacturers. Philips's subsidiary, Polygram, was assigned the task of developing the disc material; they determined that a polycarbonate material was most suitable. It was decided that the data track would start at the disc's center and run outward. Disc diameter was set at 115 millimeters. After considerable development, a gallium arsenide diode laser was selected for the pickup. In March 1979, a prototype system was successfully demonstrated in Europe, then in Japan. Matsushita turned down an offer to collaborate, but Sony accepted.

One after another, design decisions were made. Philips wanted a sampling frequency of 44,050 samples per second, whereas Sony wanted 50,000; they settled on 44,100. Philips dropped earlier work on convolutional error correction codes and accepted the Sony recommendation of a cross-interleaved Reed-Solomon code. Sony also insisted on encoding a 16-bit audio signal, whereas Philips felt 14 bits were sufficient; after lengthy critical listening tests, a 16-bit standard was adopted. To accommodate the extra data, disc diameter was increased to 120 millimeters. Maximum playing time was set at 74 minutes. According to one version of history, the 74-minute time was selected because conductor Herbert von Karajan, a Polygram artist, demanded a 74-minute capacity so his favorite piece, Beethoven's *Ninth Symphony*, would fit on a single disc. The other version claims the wife of Sony chairman Akio Morita insisted that Beethoven's *Ninth Symphony*, her favorite piece, fit on a disc.

In June 1980 the finalized Philips/Sony Compact Disc standard was formally proposed. Matsushita adopted the Compact Disc standard in 1981, as did the Digital Audio Disc Committee. In 1981 Sharp Corporation succeeded in mass producing the semiconductor laser source. The Philips/Sony engineering collaboration ended in 1981, and each company retired to prepare its products. They agreed that if either was unable to design a commercial product, the other would donate its design. The agreement

was not needed. Philips required eighteen months to reduce their prototypes to a first generation integrated chip set which incorporated digital oversampling filters and 14-bit converters with noise shaping. Sony did it in seventeen months but used 16-bit converters and analog brickwall filters.

The compact disc was introduced in Europe and Japan in the fall of 1982 and in the United States in the spring of 1983. The Compact Disc Group was formed to aid in marketing efforts. CD-ROM prototypes were demonstrated. In 1983, 30,000 players and 800,000 discs were sold in the United States. In 1984, second generation CD players were introduced, along with automobile CD players. The first CD pressing plant in this country was opened in Terre Haute. Portable CD players were introduced, followed in 1985 by third generation CD players and CD-ROM computer drives. The CD-I concept was introduced in 1986, the same year that over three million players and over fifty-three million discs were sold in the United States. In 1987, both CD-V and CD-3 discs were introduced. In 1988, the CD-R write-once standard was introduced, and the development of erasable/recordable CDs was announced. In 1990, eight years after its introduction, the CD had achieved a household penetration of 28 percent, with annual sales of 9.2 million players and 288 million discs sold in the United States with worldwide disc sales approaching 1 billion. The CD-I format was launched in 1991.

And there you have it—a very brief look at how we got to where we are now. It's certainly not a complete chronology, but it includes many of the technological prerequisites to the compact disc and to digital audio in general. It serves to demonstrate that any advance in technology is an effect with many causes. The compact disc exists today because of a unique sequence of many technological events. Omit any of these breakthroughs, and the CD as we know it today would have been a technological impossibility.

Of course, the compact disc does not signal the end of development in audio technology. In much the same way that other breakthroughs, such as Edison's wax cylinder, engendered startling consequences, the CD will undoubtedly have the same effect, if not an even more radical one.

## FOR FURTHER READING

Anazawa, T., et al. "A Historical Overview of the Development of PCM/Digital Recording Technology at Denon." In *Proceedings of the AES Seventh International Conference—Audio in Digital Times*. K. C. Pohlmann, chair. New York: Audio Engineering Society, 1990.

Blesser, B., B. Locanthy, and T. G. Stockham, Jr., eds. *Collected Papers from the AES Premiere Conference—Digital Audio*. New York: Audio Engineering Society, 1983.

Deloraine, E. M., and A. H. Reeves. “The 25th Anniversary of Pulse Code Modulation.” *IEEE Spectrum* 2(5):56–63, May, 1965.

Guterl, F. “Compact Disc,” *IEEE Spectrum*, 25(11):102–108, 1988.

McGill, J. F. “An Introduction to Digital Recording and Reproduction.” In John Strawn, ed. *Digital Audio Engineering: An Anthology*. Madison, Wisc.: A-R Editions, 1985. (Originally published in Los Altos, Calif.: William Kaufmann.)

Nijzen, C. G. “And the Music Went Round and Round on Rolls, Disks, or Reels.” Part 1, *Journal of the Audio Engineering Society* 32(3):162–78, March, 1984. Part 2, *Journal of the Audio Engineering Society* 32(4):262–73, April, 1984.

Read, O., and W. L. Welch. *From Tin Foil to Stereo: Evolution of the Phonograph*. Indianapolis: Sams, 1976.

Shannon, C. E. “A Mathematical Theory of Communication.” *Bell System Technical Journal* 27:379–423, 623–56.



## *Chapter Two*

# Fundamentals of Digital Audio

### **INTRODUCTION**

The essence of digital audio lies in its numerical basis. A digital audio system converts an analog waveform into numbers, stores, transmits, or processes those numbers, and then converts them back to an analog waveform. These steps may be divided into the digital audio recording of a music performance, the storage of that performance on a compact disc, and the reproduction of the performance with the CD player.

The digital audio chain, from the recording studio or concert hall to your living room, is predicated on several basic phenomena. First, the analog waveform must be discrete-time sampled at a precise, unvarying rate. This divides the waveform into discrete intervals. With sampling, the signal must be filtered to prevent aliasing, a form of distortion. Next, each sample must be converted to a number; this is the process of quantization. To minimize an inevitable error associated with quantization, a dither signal must be introduced. In many digital audio recorders/reproducers these steps take place in a linear pulse code modulation system. This chapter examines all of these fundamental elements.

### **SAMPLING**

Analog is continuous. Digital is discrete. To transform the original analog audio waveform into a digital audio signal, we examine the waveform at specific points in time. We must choose points on the original analog signal. This decision process is called discrete-time sampling, or simply sampling.

Since we're speaking of time, let's try a clock analogy to illustrate how a digital music system uses sampling. Time seems to flow continuously, and the hands of an analog clock sweep across the circle of hours, covering each part of time. A digital readout clock also tells time, but with a discretely valued display and at discrete points in time. In other words, it displays a sampled time. It's the same with audio. Audio varies continuously in time and may be recorded and reproduced either in continuous analog form or time-sampled digital form. Just as both clocks tell time, both types of recordings play the same music. In summary, sampling means dividing a signal into evenly spaced, discrete points in time.

However, there is a price to pay, in the form of several tricky decisions. The first question we are faced with involves how many measurements we want. In other words, how often do we record a data point? How fast do we sample? And if a digital system samples discretely, what happens to information between samples? To answer these questions, let's go back in time.

### ***Harry's Theorem***

It is February 1928. Shiny new Model As bounce noisily along barely paved roads. *The Jazz Singer*, the first talking motion picture, is opening in theatres across the country. The stock market is spiraling to new heights; everyone's future prosperity is assured. The winter convention of the American Institute of Electrical Engineers is being held in New York City. An engineer from American Telephone and Telegraph Company delivers a paper entitled "Certain Topics in Telegraph Transmission Theory." The engineer apologizes, saying that it is necessary "to include a certain amount of material which is already well known to telegraphic engineers." But at least one of his topics isn't so familiar. He shows that a waveform may be sampled discretely in time; furthermore, the waveform may be completely reconstructed if it is sampled correctly.

The engineer, named Harry, has just completed his paper. We catch him as he steps from the podium to ask him a few questions.

"Harry, the mathematics is great, but can you state this theorem in plain English?"

"That's easy. To understand it, let's consider how a movie camera works. A strip of film runs past the shutter so that a series of still images is photographed. When the movie is projected, these still pictures create a moving image. The frame rate is important: to capture a fast moving object, we would need a fast frame rate. Digital audio is the same. Instead of taking still photographs to create a movie, you take still audio samples to create sound. My theorem tells you how fast you should take samples."

"Is that all there is to it?"

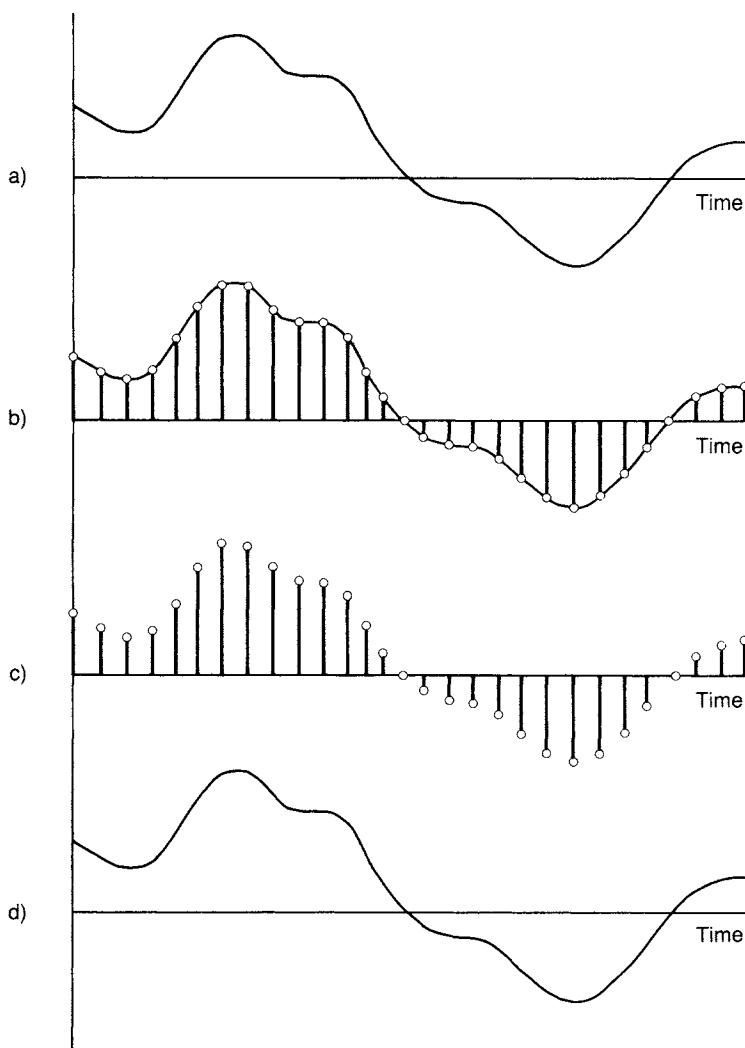


Figure 2.1 With discrete-time sampling, an audio waveform (a) may be sampled (b), and the sample points (c) may be used to reconstruct the waveform (d).

“Pretty much. An audio waveform is a variation in air pressure; it exists continuously in time, over a continuously variable amplitude range. If you put a pencil to paper and start drawing in one direction and never go in the opposite direction or let the pencil line intersect with itself, you’re drawing a representation of an analog audio waveform (figure 2.1a). Now if you go back and periodically tap your pencil along the line (figure 2.1b),

you've sampled the waveform (figure 2.1c). The interesting part is that if you follow my theorem, your sample points can be used to reconstruct the original analog waveform, without loss of information" (figure 2.1d).

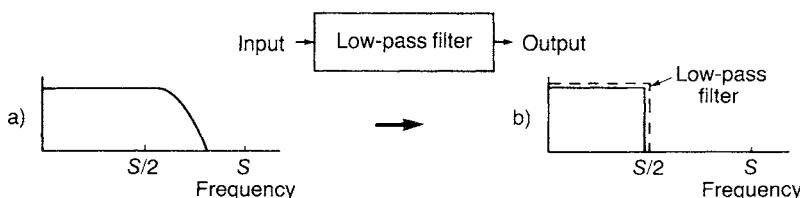
"That's incredible, Harry. Are you sure about this?"

"The math is impeccable. Discrete time sampling is legitimate. Given correct conditions, the theorem says that no information is lost due to sampling in a digitization system. The samples contain the same amount of information as the unsampled signal. The trick is to follow the correct conditions.

"It's just like in our film. Once we determine the speed of the fastest object we wish to photograph, we can set the frame rate accordingly. For sound, the samples correspond to the film frames, the sampling rate corresponds to the frame rate, and the high-frequency signal corresponds to the fast-moving object. Just as we need some relationship between the frame rate and the speed of moving objects, we need to define the relationship between sampling frequency and audio frequencies. That's where my theorem comes in. It states that if we take samples at a frequency at least twice that of the highest audio frequency, complete reconstruction can be accomplished."

"What about the waveform between samples?"

"Not to worry. By definition, the samples have all the information we need to provide complete reconstruction. Of course, the frequency criteria must be strictly observed. If you were making a movie and someone shot an arrow across the set, it might pass by without being recorded on a frame. The same thing is true with sound; a frequency that is too high would not be properly encoded and would even create a kind of distortion called aliasing. To prevent that, we limit the frequency content of the input signal with a low-pass filter. Its job is to ensure that all frequencies greater than one half the sampling frequency are attenuated below the amplitude resolution of the system (figure 2.2). To limit the sampling rate



**Figure 2.2** *The sampling theorem dictates that the frequency content of the audio signal be less than or equal to the half-sampling frequency. The input signal may contain frequencies greater than the half-sampling frequency as shown by the spectrum (a). A low-pass filter removes high frequencies (b) to produce a spectrum of frequencies below the half-sampling frequency.*

required to achieve a given audio bandwidth, we design low-pass filters with a very sharp characteristic—the so-called brickwall filters.”

“They have a cutoff at the half-sampling frequency?”

“Actually, we design them to begin cutting a little before that point to give the filter’s response an opportunity to achieve full attenuation before the half-sampling frequency. Conversely, we place the half-sampling point a little above the highest audio frequency we want in our flat passband. For example, to achieve flat response to 20 kHz, we might sample at 44 kHz; the filter would have 2 kHz, from 20 to 22 kHz, to achieve full attenuation. Whatever the numbers, the idea is clear—the filter bandlimits the input signal, thus ensuring that the theorem’s criteria will be upheld. Thanks to sampling, we can reconstruct the waveform later.”

“What if we want a higher frequency response?”

“No problem—just raise your sampling frequency. But a 20 kHz bandwidth should be okay, since that’s the limit of most people’s hearing.”

“So choice of sampling frequency determines the highest frequency reproducible by the digitization system?”

“Right. And that’s the essence of the theorem.  $S$  samples per second are needed to represent a waveform with a bandwidth of  $S/2$ . ”

“Hey, I just thought of something. If you’re sampling at 40 kHz, the 20 kHz waveform will have only two samples. How can that work?”

“As you stated, when the sampled frequencies become higher, the periods are shorter, and there will be fewer samples per period. At the theoretical limiting case of critical sampling, at an audio frequency of half the sampling frequency, there will be two samples per period. Consider a 20 kHz input sine wave (figure 2.3a). The digitizer would generate two samples, and this would be used at the output to produce a 20 kHz square wave (figure 2.3b).

“Of course, a square wave is quite unlike a sine wave, which brings us to the final piece of system hardware. At the output of every digitization system is another low-pass filter, essentially identical in design to the input

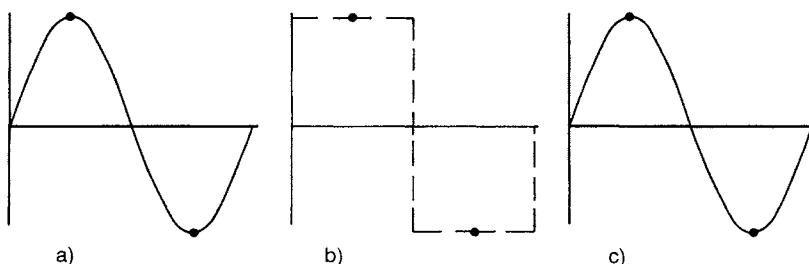


Figure 2.3 With critical sampling, a sine wave input (a) would be reconstructed as a square wave (b), low-pass filtered, and output as a sine wave (c).

low-pass filter. Its job is also to remove all frequencies above the half-sampling frequency. We have to do that again because sampling has generated new frequencies above the audio band, that is, all the high-order harmonics in the square wave. The output filter removes all the unwanted harmonics above the half-sampling frequency, leaving only the original 20 kHz sine wave” (figure 2.3c).

“But what if the input wasn’t a sine wave?”

“It had to be. Our input filter would have removed the harmonics from any 20 kHz waveform, leaving only the fundamental sine wave.”

“We’re losing all the harmonics!”

“Sure. But that’s okay. All complicated waveforms, which are everything other than a sine wave, are basically composed of a fundamental sine wave and a series of harmonic sine waves at multiples of the fundamental. For example, a 1 kHz waveform would have a fundamental of 1 kHz and harmonic sine waves at 2 kHz, 3 kHz, 4 kHz, 5 kHz, etc. However, that information is useful only in your hearing range. You probably can’t hear harmonics above 20 kHz, so there’s no reason for a digital system to ‘hear’ them either. A complex 20 kHz waveform has harmonics at 40 kHz, 60 kHz, 80 kHz, 100 kHz, etc., and most people cannot hear anything past the fundamental. There is no need for a digitization system to attempt to record such high frequencies.”

“I guess your sampling idea holds water.”

“It sure looks like it. Of course, it’s all still theoretical, but sooner or later someone might find an application for the theory—there could be some real possibilities.”

### *The Nyquist Theorem*

The engineer’s name is Harry Nyquist, and his paper contains the sampling theorem that has transformed audio and is used in all digital audio systems. There is no evidence that the other engineers at the convention stood to their feet and cheered after Harry delivered his paper. Indeed, if he had tried to describe to them the music digitization systems we take for granted today, they probably would have laughed. Nevertheless, he established the foundations of digital audio.

There are still some people around who either do not believe Nyquist’s sampling theorem or simply distrust it. For one reason or another, they do not see how little pieces of a continuous bandlimited signal can be used to reconstruct an audio waveform. In all fairness to them, the theorem is somewhat like a brainbuster in the comics section of the Sunday paper, because it seems counterintuitive.

Briefly, let’s reconsider Harry’s analogy. Just as the discrete frames of a movie create moving pictures, so the samples of a digital audio recording create time-varying music; there is little conceptual difference between

the visual and aural systems. In a digital audio system, such as the one shown in figure 2.4, we must limit the highest frequency in the incoming signal to a frequency of one-half the sampling frequency (the half-sampling frequency is also known as the Nyquist frequency). Specifically, the signal is low-pass filtered at 20 kHz or so. When this filtering is accomplished, we can successfully sample the signal such that there is no loss of information due to sampling between the sampled signal at the output and the filtered signal at the input. When the input is filtered, we can restore all the intervening values, without error due to sampling, and thus re-create the original waveform filtered at the upper limit of audibility. An output filter smooths the waveform to return it to its original state.

We merely observe the theorem's dictate that we must sample at a frequency at least twice that of the highest input frequency. In the case of the compact disc, the sampling frequency is 44.1 kHz (a sample every 22.6757 millionths of a second). This permits an audio frequency response from 0 to 22.05 kHz. Thus, each channel of a CD outputs a sample 44,100 times per second. With stereo, that's 88,200 samples per second. And over the course of an hour, that's about 320 million values. Fortunately, such high data rates are well within the scope of mass-produced technology.

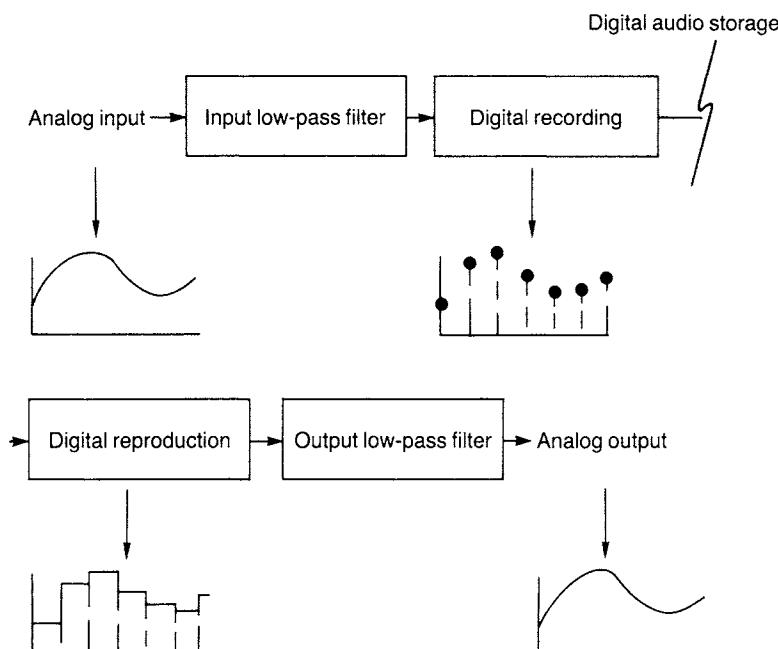


Figure 2.4    *An audio digitization system showing sampling and de-sampling.*

At any rate, the point is clear: A filtered signal may be sampled, stored in discrete values, de-sampled, and reproduced without loss due to sampling.

Why 44.1 kHz? It's just a question of circumstance. Early digital tape recorders used video cassette recorders for storage (some digital PCM recorders still do). A sampling frequency of 44.1 kHz was derived from the video scan frequencies of the television system. Specifically, 50 Hz PAL/SECAM recorders provide the following: (50 lines/second) (294 lines/video field) (3 samples/line) = 44.1 kHz. This sampling frequency is also fortuitous because it allows use of 60 Hz NTSC recorders for CD mastering: (60 lines/second) (245 lines/video field) (3 samples/line) = 44.1 kHz.

However, there are other important considerations pertaining to sampling. We still must discuss exactly why we must have a low-pass filter at the input and the output of a digital music system and what happens if we don't.

## **ALIASING**

One particular challenge to the audio digitization system designer is that of aliasing, a kind of sampling confusion that can take place in the recording side of the chain. A criminal at large can assume an alias to confuse his identity. Similarly, a sampled signal can masquerade as another; this is aliasing. An alias frequency appears as another, new frequency. It is thus the designer's obligation to prevent aliasing from ever occurring. As we will see, an input low-pass filter (sometimes called the anti-aliasing filter) will accomplish this, but only if its cutoff frequency is carefully chosen.

Aliasing is a digital system's improper response to any frequency greater than half the sampling frequency. Nyquist has shown that the highest signal frequency in a sampling system can be no more than half the sampling frequency. As the signal frequency becomes higher and higher, the number of sample points per cycle becomes fewer. When the signal frequency reaches half the sampling frequency, there are only two samples per cycle, which is the absolute minimum needed to record a waveform. A higher frequency would cause the digitization system to alias.

### *Stagecoach Wheels*

What is aliasing? One example occurs in western movies. When the marshall climbs up onto the stagecoach and drives away, the spokes on the stagecoach's wheels begin rotating forward, faster and faster. Then a funny thing happens. The spokes first appear to slow down, and then appear to

rotate backwards, even though the stagecoach is obviously picking up forward speed. As the stagecoach goes even faster, the spokes start moving forward again. The problem lies in the rate of the film's frames. If the camera is taking 24 pictures per second, the sampling theorem dictates that a maximum change of 12 spokes per second can be properly reproduced. Thus a twelve-spoke wheel could reach a maximum speed of one revolution per second for proper reproduction. If the speed increased to 13 spokes per second, the camera would show 11 spokes per second. At 14 spokes per second, a rate of 10 would be reproduced. The reproduced spoke rate is wrong; a form of aliasing is occurring.

Audio waveforms behave similarly. When a waveform is undersampled—that is, when its frequency is greater than the half-sampling rate—aliasing can create new frequencies. Figure 2.5 shows an input waveform, sample points, and a reconstructed waveform. Both waveforms are valid for those samples, but the new waveform is an alias frequency.

A more complete understanding of aliasing may be obtained when sampling is considered as a modulation process. One result of that modulation is the creation of a spectrum, or frequency response, having sidebands, or image frequency responses, centered on multiples of the sampling frequency. These sidebands are a consequence of sampling and result even when the Nyquist theorem is observed.

Figure 2.6 shows an input audio signal (a), the sampling signal (b), and the resulting sampled audio signal (c). The left side of each pair of graphs is the familiar time-domain representation of the waveforms. The right side is the waveform's spectrum, or frequency-domain representation. The input audio signal has been properly bandlimited to the half-sampling frequency of  $S/2$ . The sampled audio signal contains the original spectrum, as well as images of that spectrum, centered at multiples of the sampling frequency. The output filter would remove the high-frequency image spectrum, leaving only the original. The sampling theorem has been observed, and all is well.

### ***Foldover Frequencies***

Aliasing takes place when the sampling theorem is not observed. Fortunately for our understanding, we can draw a simple picture of aliasing. Figure 2.7 shows what happens when the input frequency exceeds the half-sampling frequency; aliasing can be seen as overlap between the signal bands. It is evident that the output filter can remove the higher-frequency image spectra and their aliasing; however, the aliased frequency below  $S/2$  cannot be removed from the audio signal. In this case, the upper frequencies of the audio band become contaminated with the lower frequencies of the side band. The result? Distortion.

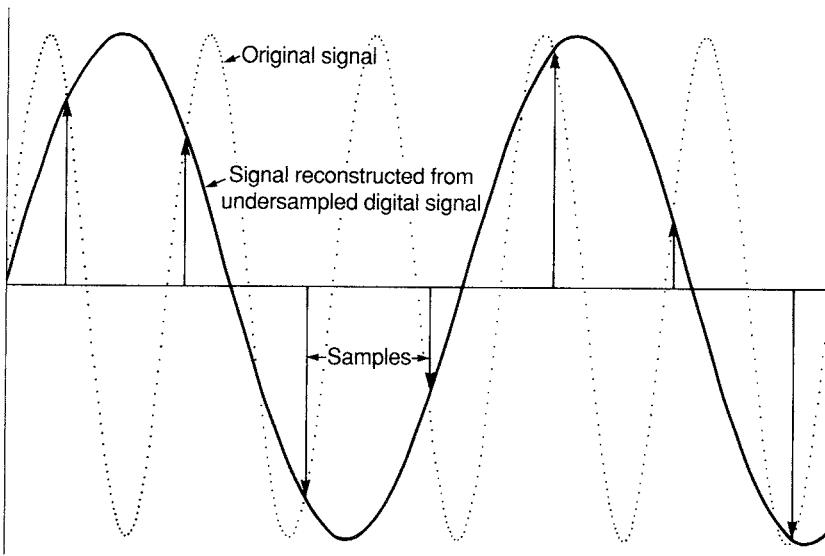


Figure 2.5     *As a result of undersampling, aliasing can create a new frequency.*

As a digital audio system samples these disallowed input frequencies, the spurious new frequencies are “recorded” as well. As the deviate input frequencies ascend, new descending frequencies are created. Specifically, if  $S$  is the sampling frequency and  $F$  is a frequency higher than half the sampling rate, then a new frequency  $A$  is also created at  $A = \pm NS \pm F$  where  $N$  is any integer. In practice, the  $S - F$  component is the most troublesome. For example, if  $S = 44$  kHz, a 23 kHz input signal will alias to produce another frequency at 21 kHz. If we attempt to sample a 24 kHz signal, 20 kHz appears. In other words, a new frequency appears back in the audio band, folded over from the sampling frequency, as the placement of the sidebands predicts. In fact, aliasing is sometimes called fold-over.

To elaborate further upon this sad scenario, in which a false frequency accompanies the actual frequency following the sampler, we must remember that a low-pass filter is used at the output of a digitization system to smooth the staircase function and thus recover the original signal. This output filter (sometimes called the anti-imaging filter) will be designed to cut at half the sampling frequency; thus the errant input frequencies which were above that value will be filtered out, and we would be left with only the aliased frequencies in the audio band.

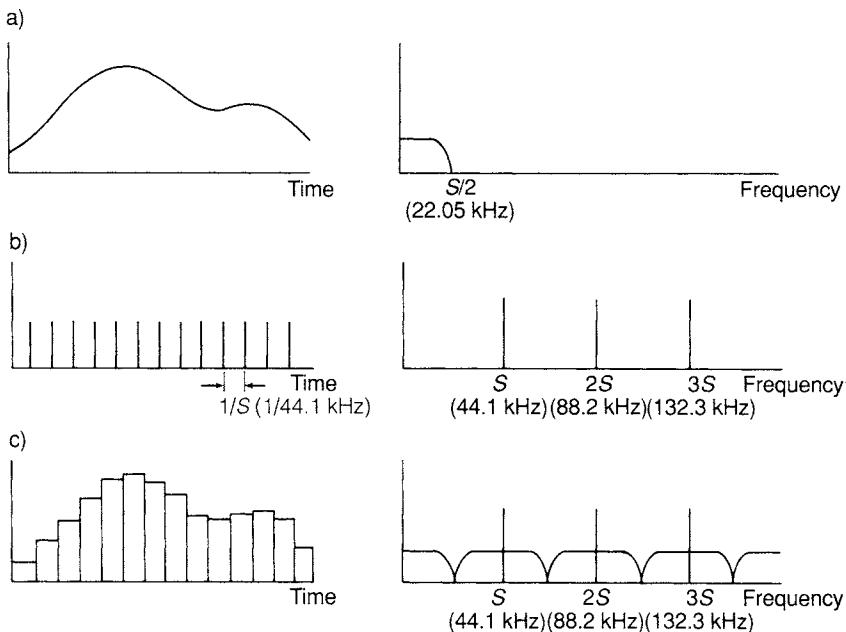


Figure 2.6 When a bandlimited audio signal (a) is sampled by a sampling signal (b), a series of high frequency image spectra (c) are created at multiples of the sampling frequency.

Aliasing can produce some rather strange results in the audio band. If you were sampling at 44 kHz, an input frequency from 0 to 22 kHz would sound fine, but as the frequency ranged from 22 to 44 kHz, we would hear it returning as a frequency descending from 22 kHz to 0 kHz. If we raised the input frequency from 44 to 66 kHz, it would appear again from 0 to 22 kHz. The process would continue infinitely, folding back one sideband after another.

Our sine wave example limited foldover to one fundamental frequency. With complex tones, aliasing frequencies could be generated separately for each harmonic. For example, a complex tone of 5 kHz would normally have harmonics at 10 kHz, 15 kHz, 25 kHz, 30 kHz, etc. With sampling at 44 kHz, however, the 25 kHz fifth harmonic would fold over to 19 kHz, and the sixth harmonic at 30 kHz would pop in at 14 kHz, just shy of the 15 kHz third harmonic. This is illustrated in the spectrum in figure 2.8. A distortion akin to intermodulation distortion would result.

As difficult a problem as aliasing might be in theory, in practice it can be alleviated. In a well-designed digital recording system, aliasing does not occur. The solution is simple. We merely bandlimit the input fre-

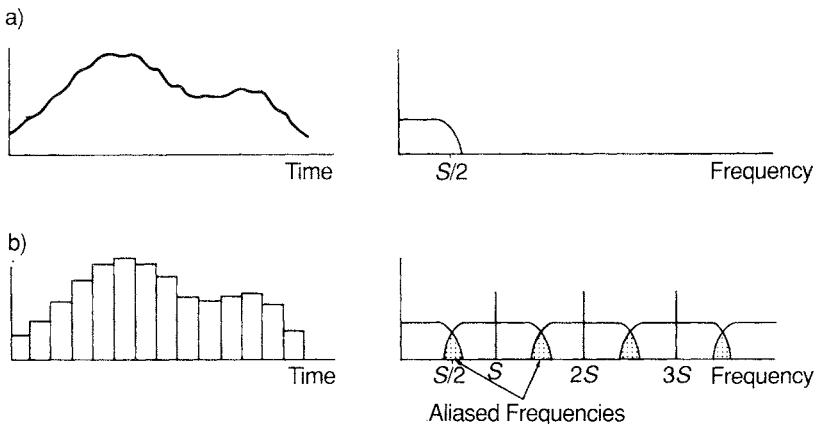


Figure 2.7 When an audio signal is not properly bandlimited (a), the image spectra (b) overlap. This is aliasing.

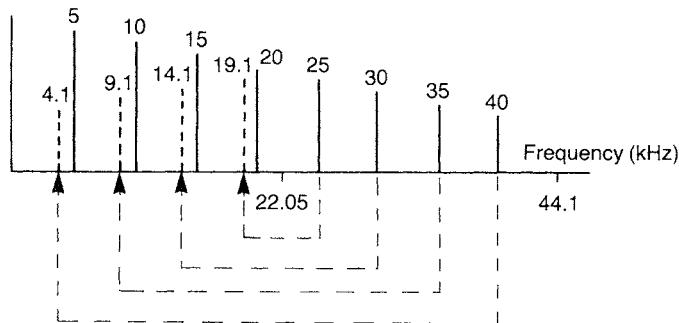


Figure 2.8 The harmonics of a sampled waveform generate alias frequencies.

quencies with a low-pass filter. Its job is to ensure that all frequencies above the half-sampling frequency are attenuated to below the amplitude resolution of the system. The system will not attempt to record these extreme low-level signals. To achieve a maximum audio bandwidth for a given sampling rate, filters with a very sharp characteristic, often called brick-wall filters, are employed. Only certain fundamentals or harmonics are allowed to enter the sampler; thus aliasing cannot occur. In other words, when sampling, one must obey the sampling theorem.

However, time sampling is only one consideration; a digital system must also be able to determine the actual numerical values it will use at sample

time to represent the original waveform's amplitude. That process, amplitude quantization, has its own quirks.

## QUANTIZATION

Whenever there's good news, there is invariably a little bad news, too. At least that's the case with amplitude quantization, the counterpart to time sampling. Whereas sampling is counterintuitively wonderful, quantization is intuitively not wonderful. Hence the need for the short, apologetic introduction that follows.

In essence, the objective in recording and reproducing music can be simply formulated: We want to form a representation of the music. Obviously, the closer our representation is to the original, the better. Unfortunately, reality is stubborn in its ability to defy re-creation, and we are left with a challenging endeavor as we attempt to create an approximation of the original event. The essential problem lies in the complexity of even the simplest acoustic waveform and in the problem of measuring the waveform's infinitely varying amplitude.

As we've seen, sampling is a method of periodically taking a measurement. An audio measurement is meaningful only if both the time and value of the measurement are stored. Sampling represents the time of the measurement; quantization represents the value of the measurement or, in the case of audio, the amplitude of the waveform. Both sampling and quantization are parameters which determine the limitations of an audio digitization system.

While sampling is perfect within its prescribed context, quantization is imperfect. The problem lies in measuring the waveform's amplitude. In the same way that length, an analog quantity, can be measured (e.g., in inches), an audio waveform can be measured (e.g., in volts). In the same way that a measurement of length is always prone to error, quantizing is prone to error.

### *Digital Error*

Let's take a quick ride in my Porsche 911 to demonstrate that, yes, even digital systems have error, and let us differentiate this error from the one inherent in an analog system. Suppose that I've connected three speedometers, one analog and two digital types, like the ones in figure 2.9. Suppose also that I read the first two speedometers just as a police officer pulls up behind me, red lights flashing. Given a good meter face and a sharp eye, I read the analog needle as approximately 75 MPH. My digital speedometer reads 77.3 MPH. The trooper writes me a ticket.

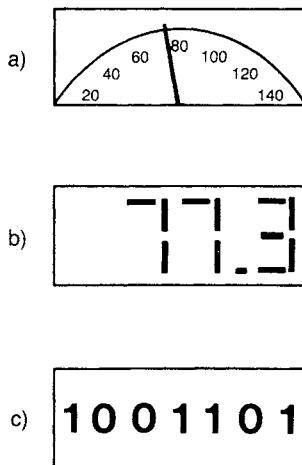


Figure 2.9     *Three types of speedometers: analog (a), digital with decimal read-out (b), and digital with binary readout (c).*

Now, any type of speedometer is always in error. The analog speedometer errs because of the ballistics of the mechanism. That is, it takes the needle a certain amount of time to settle to the correct value. With the digital speedometer, precision is limited by the number of digits displayed. The more digits, the greater our precision, but the last digit will always round off relative to the actual value. For example, 77.3 MPH may have been displayed for an actual 77.34 MPH speed. Under the best conditions the last digit would be completely accurate; for example, 77 MPH would be shown as 77.0 MPH. Under the worst conditions, the rounding off will be one half increment away; for example, 77.55 MPH will be rounded off to 77.5 MPH or 77.6 MPH. No matter how many digits you have, unless you have an infinite number of them, you will never be able to display the analog speed of  $77\frac{1}{3}$  MPH exactly. Nevertheless, a digital readout is an inherently more precise way to measure an analog event.

To solve my speeding problem, I might use a microprocessor to limit my Porsche's top speed. (This is a strictly hypothetical argument, mind you.) Because electronic circuits mainly use binary numbers, I should also express my speed that way, necessitating a third speedometer, a binary type.

### ***Binary Numbers***

Binary numbers are really no different from any others. All number systems employ a radix, also called a base, which defines the number of sym-

bols the system uses. Most of us ten-fingered humans use the decimal system with, noncoincidentally, ten symbols, 0 through 9. From these basic symbols, a method called positional notation allows us to use them to express even very large numbers, such as the amount of the national debt.

In positional notation, each left-most symbol is a power of ten higher than the previous symbol. Thus the number 7085 represents  $7 \times 10^3 + 0 \times 10^2 + 8 \times 10^1 + 5 \times 10^0$ , or  $7 \times 1000 + 0 \times 100 + 8 \times 10 + 5 \times 1$ . The 7 is the most significant digit, while the 5 is the least significant.

Binary numbers operate similarly, except that the system is based on the power of two: 0 and 1 are used to represent the two (and only two) symbols in the system. Such a system makes perfect sense to an electrical machine: It can equate 1 with ON and 0 with OFF. Thus a voltage, or lack of it, represents the only two conditions it must deal with.

As in the decimal system, we employ positional notation to define numbers. Each left-most digit is a power of two higher than the previous digit. For example, the number 1001101 represents  $1 \times 2^6 + 0 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$  or, in its decimal equivalents,  $1 \times 64 + 0 \times 32 + 0 \times 16 + 1 \times 8 + 1 \times 4 + 0 \times 2 + 1 \times 1$ , or 77. In this way, it can be seen that decimal words are interchangeable with binary words, a principle illustrated in table 2.1. In a binary word, the left-most digit is called the most significant bit (MSB), and the right-most is called the least significant bit (LSB). Eight-bit words are called bytes. Conventional digital audio systems use 16-bit words to record, process, and reproduce audio signals. It is the precision of using numbers combined with the speed and reliability of circuits designed to process binary numbers that make digital systems well suited for working with information such as audio signals.

Decimal	Binary	Decimal	Binary
0	0000	8	1000
1	0001	9	1001
2	0010	10	1010
3	0011	11	1011
4	0100	12	1100
5	0101	13	1101
6	0110	14	1110
7	0111	15	1111

Table 2.1     *Digital circuits prefer binary representation rather than decimal.*

### ***Quantization Error***

Numbers are great, but their discrete nature imposes a limitation. An analog waveform may be represented as a series of samples; the amplitude of each will yield a number which represents the analog value at that instant. By definition, an analog waveform has an infinite number of amplitude values, whereas in a digital representation we can choose only from a finite number of discrete values, or steps. The selected values will be only an approximation of the actual. Specifically, after sampling, the analog staircase signal must be approximated by the numerical value nearest to the analog value.

Quantization is the technique of approximating an analog amplitude to form a discrete number. In terms of the quantizing hardware, the number of allowable steps is determined by the length of the data word in bits. Just as the number of digits in our speedometer determined our resolution, the number of bits in our digitization equipment determines its resolution. Two bits would yield four ( $2^2$ ) possible quantization values: (in binary form, 00, 01, 10, and 11). Eight bits would create  $2^8$  or 256 steps (ranging from 00000000 to 11111111). A 16-bit word would accommodate  $2^{16}$  or 65,536 values (ranging from 0000000000000000 to 1111111111111111).

Word length determines the resolution of our digitizing system and hence provides an important specification to measure the system's performance. Yet there will always be an error associated with quantization because the limited number of amplitude choices contained in the binary word can never completely map an infinite number of analog possibilities. Rarely will the chosen step be exactly at the analog value; usually it won't be quite exact. At worst, the analog level we desire to encode will be one-half step away: That is, there will be an error of one-half the least significant bit of the quantization word. For example, suppose the binary word 0010 corresponds to the analog step value of 2.0 V, and 0011 corresponds to 3.0 V, and the actual analog value at sample time is unfortunately 2.5 V. Since 00010½ isn't available, the system will round up to 0011 or down to 0010; either way, there will be an error of magnitude of one-half step.

Figure 2.10a shows how an input signal would be sampled and the error created when the amplitude at sample time is quantized. Figure 2.10b shows the error at each sample time. Note that the error never exceeds a value of one-half step.

Digital hardware performance may be characterized by a ratio of the total number of steps covered by a quantization scheme to the maximum error. This ratio of maximum expressible amplitude to error determines the signal-to-error (S/E) ratio of the digitization system. For example, in a 16-bit system the ratio of maximum amplitude (65,536 steps) to error (0.5 of a step), would be  $65,536/0.5 = 131,072$ . Every added bit doubles

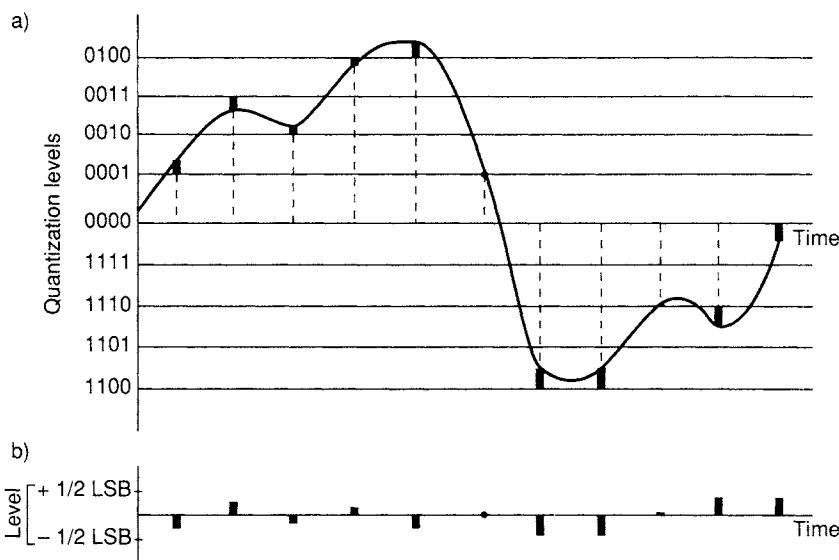


Figure 2.10 When a signal is quantized (a), an error results (b). The magnitude of the error signal can never exceed one-half of a quantization step.

the number of possible values. The more bits the better, because they provide more choices. So the steps are smaller and the magnitude of the error is smaller. In terms of decibels (dB), every added bit yields about a 6 dB increase in the system's S/E ratio because it reduces the quantization error by a factor of two. The S/E relationship can be conveniently expressed in terms of word length as  $S/E \text{ (dB)} = 6.02n + 1.76$  where  $n$  is the number of bits. Using this formula, a 16-bit system yields an S/E ratio of about 98 dB. This equation is discussed more fully in the references listed at the end of this chapter.

Quantization error may be understood as an analog signal whose amplitude cannot exceed one quantization interval (provided anomalies from the real world do not interfere). The difference between the original audio waveform and the regenerated waveform after digital-to-analog (D/A) conversion is due in part to quantization error. The error has a spectrum, or frequency response, which is a function of the original signal. Audibly, quantization error is a form of distortion; fortunately, with 16 bits it exists below the noise level in most analog systems. As we'll see, its effects can be reduced or eliminated in the recording chain with dither, a purposely introduced noise.

Quantization is affected by more than just word length; it is also a question of hardware design. There are many techniques available to ac-

complish quantization, and different strategies determine how the analog signal is matched to the steps. These algorithms influence the relative effects of quantization error. For example, a linear quantizer produces a relatively high error when low-level signals span only a few steps. A nonlinear system could be used to amplify low-level signals to utilize the fullest possible span of steps. However, the resulting error is generally more benign in the linear quantizer and more perceptible in the nonlinear system.

Sampling and quantization are the two fundamental design elements for an audio digitization system. In the case of the compact disc system, a sampling frequency of 44.1 kHz and a word length of 16 bits yields remarkable fidelity, comparable to the best analog systems.

## DITHER

Although a 16-bit system yields a theoretical signal-to-error ratio of 98 dB, it is disturbing to know that as the signal amplitude decreases the relative error increases. When signal level is on the order of one least significant bit—that is, one quantization increment—errors could be audible under low-amplitude signal conditions. So quantization not only loses information, it creates some unexpected problems as well. Let's examine those problems as well as a remedy for quantization error.

### *Granulation Noise*

Digital signals are made up of discrete bits. This is a real blessing; it permits us to employ powerful digital methods for processing and storage. However, in quantization, the discrete nature of digital representation and the ensuing amplitude increments can create nonlinearity. Consider a signal with amplitude on the order of one quantization step, such as the one in figure 2.11. The signal value crosses back and forth across the threshold, resulting in a square wave signal from the quantizer. At very low levels, quantization results in severe distortion. The quantization error floor is sometimes called “granulation noise” because it has been described as having a gritty sound.

To make matters worse, the quantized square wave (like all square waves) is rich in odd harmonics, extending far beyond the half-sampling frequency. Thus aliasing, discussed previously, reappears *after* the anti-aliasing low-pass filter, because of the high-frequency content of the square wave. If the resulting harmonics are very close to a multiple of the sampling frequency, the beat tones drift through the frequency origin to produce a sound called “bird singing” or “birdies.”

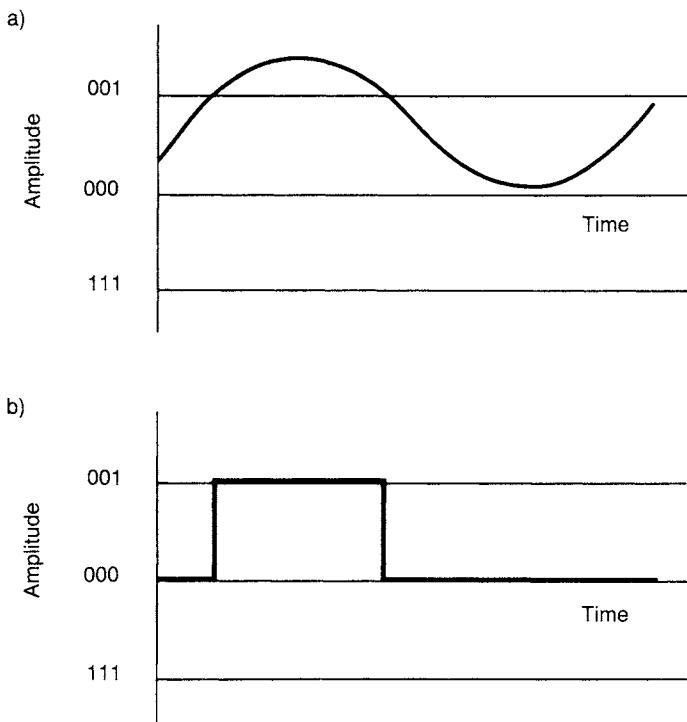


Figure 2.11 Without dither, a low-level signal (a) will be quantized to create a square wave (b).

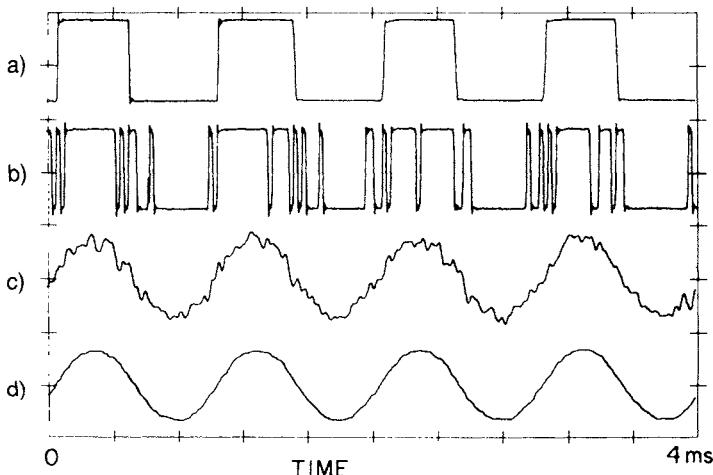
Quantization thus loses low-level information and can lead to aliasing and other effects. These effects occur at the low level of the quantization steps themselves. Thus, the amplitude of the distortion is also very low. Most high-amplitude audio signals mask the error effects. On the other hand, when the signal is low-level, it may not mask the distortion. Yet there is a solution.

#### ***Resolution below the LSB***

Ideally, a digital audio conversion system must suppress any error it introduces. Obviously, the number of bits in the digitization word could be increased with a resultant decrease in error of about 6 dB per bit. This is difficult because of the cost of adding extra bits. An alternative, surprisingly, is to add a small amount of analog noise to the input signal; this noise is referred to as dither. It is important to note that this small amount of noise removes some quantization artifacts from a signal—it doesn't mask the artifacts, it actually removes them.

Let's reconsider our example of a signal with amplitude on the order of a quantization increment, but this time let's add a dither signal; the result is modulation which preserves the information of the original signal. An undithered input sine wave results in the square wave output of figure 2.12a. With dithering, the duty cycle output of figure 2.12b results. When the duty cycle signal is averaged, the sine wave information emerges, as shown in figures 2.12c and d. The averaging illustrates how the ear will actually perceive the signal. It perceives the original signal and some added noise. This is far more desirable than the severely distorting undithered waveform. From a more analytical standpoint, the spectrum of an undithered signal would yield harmonics from the quantized square wave. With dither, only the original fundamental waveform would be preserved with added noise. Thus dither changes the digital nature of the quantization error into noise, and the ear may then resolve signals with levels well below one quantization level.

Those readers who are particularly skeptical of the miraculous quality of dither may want to try the following experiment. Suppose the lines of text on this page represent the amplitude information of an audio waveform, and the spaces between the fingers of your slightly spread hand



**Figure 2.12** An undithered sine wave will produce a quantized square wave (a). With dither, the amplitude of the sine wave is reflected in the quantized signal (b). This is clearly shown when the quantized signal is averaged (c). A longer average results in a smoother waveform (d). (From Vanderkooy and Lipshitz)

represent quantization intervals. Now place your hand over this page with fingers slightly spread; amplitude information represented by lines of text is lost. Now dither the system by moving your hand rapidly up and down the page. Significantly more amplitude information can be retrieved, at the cost of adding noise (your blurred fingers) to the signal. Furthermore, observe that amplitude information smaller than the smallest quantization increment can now be successfully retrieved.

That is an important conclusion: With dither, the resolution of a digitization system is below the least significant bit. By encoding the audio signal with dither, we may recover information, even though it might be smaller than the smallest increment of the quantizer. A dithered system thus behaves more like an analog system. A properly dithered signal will be audible beneath the noise floor (as low as 110 to 120 dB below maximum output level). Moreover, much of the distortion due to quantization is removed from the signal. The performance price? A degradation of about 1.5 dB in the broadband S/E ratio of the digitization system, an increase which is negligible compared to the large S/E ratio inherent in the system. The paper by Vanderkooy and Lipshitz listed at the end of this chapter provides more information on dither.

What kind of noise is used for dither? Many types of dither signals, utilizing noise signals varying in their frequency content and amplitude, have been explored. These include Gaussian, rectangular, and triangular noise. The designer must determine which noise signal is best for a particular application. In some cases, a digital system might not even need a dither generator, because the residual analog noise in the input stages of the system itself can act as dither. Some researchers have proposed digital dither generators, in which a pseudo random noise signal could be first injected to overcome quantization artifacts and then later removed to prevent any degradation in S/E ratio.

Dither thus represents a unique and unlikely concept. The idea of adding analog noise to a digital system seems a little strange, but as we have seen, the outcome is lower distortion and consequently higher fidelity for the digital audio system.

## PULSE CODE MODULATION

Topics such as sampling and quantization form the basis of digital audio. All digital audio systems use these mechanisms to record and reproduce music. However, the methods by which they are utilized can vary tremendously. There are many ways to encode the digital data, with various system architectures to do the job. Pulse code modulation (PCM) is a conventional yet efficient high performance encoding method and forms

the most popular digital audio system architecture. PCM is a modulation process in which an analog signal is expressed in digital form by means of binary coding. Many digital recordings are mastered on PCM digital audio recorders, and at the output of a CD player in many designs, the data returns to its PCM birthright at the digital-to-analog (D/A) converter.

### ***Modulation***

Modulation is the process of encoding source information prior to transmission or storage. In general, a carrier signal forms the basis of the transmission channel, and the source information itself merely affects the carrier's parameters. For example, in amplitude modulation (AM) the carrier's amplitude varies relative to the information's amplitude at a rate relative to the information's frequency. In frequency modulation (FM) the carrier's frequency changes by an amount relative to the information's amplitude and at a rate relative to the information's frequency. A radio receiver is given the task of demodulating, or removing, the carrier and outputting the original information. In AM and FM, both the modulation and the modulating signal are continuous; they are classified as wave parameter modulation schemes.

When the original information is digital, the nature of the modulation can be altered. A parameter of the transmitted pulses, or a code contained in the pulses, carries the information. For example, pulse amplitude modulation (PAM) uses pulse amplitude, pulse width modulation (PWM) uses pulse width, and pulse number modulation (PNM) uses a pulse count to encode the value of the information, as illustrated in figure 2.13.

PAM forms an intermediate waveform in digital audio systems; prior to analog-to-digital (A/D) conversion and following digital-to-analog (D/A) conversion, the audio signal exists as a staircase PAM waveform. PNM is plagued by bandwidth problems; a large number of pulses is required to encode a high resolution signal. For example if 65,536 steps were used to measure the signal's amplitude, then 65,536 pulses would be needed to record a full scale signal level.

### ***PCM***

Pulse code modulation overcomes this bandwidth problem yet retains a numerical coding of information for efficient transmission. In pulse code modulation (PCM) the digital waveform itself encodes the binary digits. Since the binary code itself is retained, the required bandwidth is much less. Its efficiency can be intuitively observed. For example, one can readily read the encoded data directly from the PCM waveform itself. In addition, the coded signal is fully compatible with digital circuits which are usually designed to operate with a binary code. Because of its efficient use

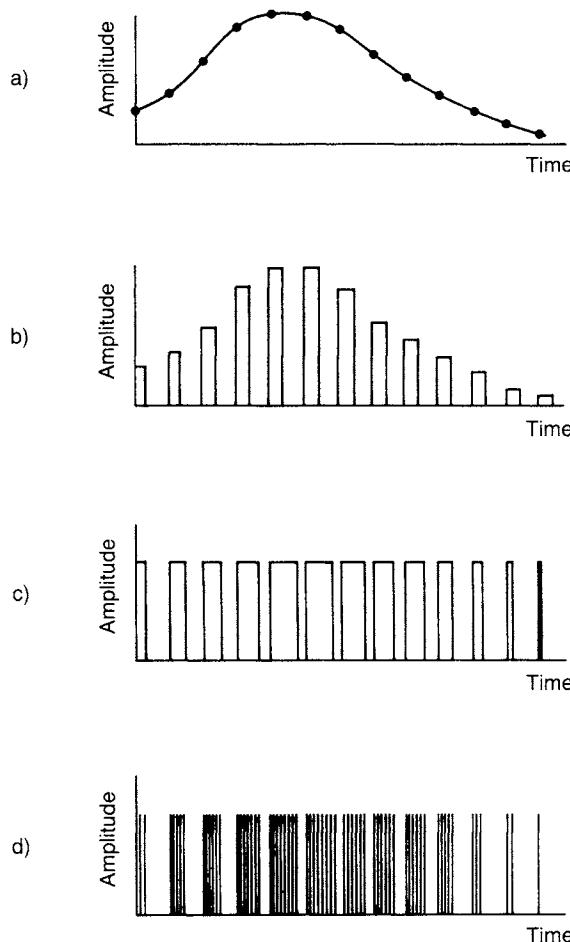


Figure 2.13 An analog signal (a) may be represented through PAM (b), PWM (c), or PNM (d) pulse modulation.

of bandwidth and its compatibility with off-the-shelf circuitry, PCM has proven to be an expedient means of representing audio data for recording or reproduction. When it comes to storage on magnetic tape or optical disc, other forms of modulation may be more effective. There is one consistent drawback to PCM; like many other coding schemes, it requires a much wider bandwidth than the corresponding analog signal.

Pulse code modulation is efficient when confronting the essential problem of an audio digitization system: the conversion from analog to digital

and back again. When the analog audio signal is sampled, analog values form a PAM waveform. Quantization must document the instantaneous amplitude of the stairstep and generate a binary word for storage. Figure 2.14 shows how PCM directly creates the required binary code at the output of the A/D converter. Similarly, the D/A converter directly accepts PCM data and converts it to a PAM waveform. Additionally, PCM data is easily multiplexed—that is, several channels of data may be merged to form one channel.

PCM is an effective modulation scheme. In its absence, at the very least, an audio digitization system would be much more complex and costly.

## A DIGITIZATION SYSTEM

So far, we've examined the various theoretical underpinnings of digital audio, but one does not live by conceptualization alone; one also needs hardware. The time has come to look at a complete digitization system, focusing on a conventional PCM hardware design.

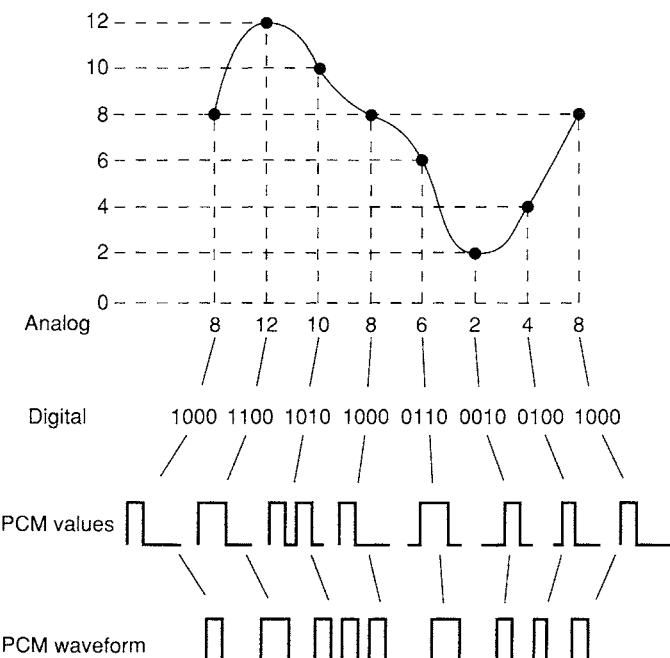


Figure 2.14     *With PCM, amplitude values are recorded directly in the binary waveform.*

### **A PCM System**

An audio digitization system is really nothing more than a kind of transducer which processes the audio for digital storage and then processes it again for reproduction. While that sounds simple, the hardware must be carefully engineered, for success in accomplishing its task (and the resulting quality of the reproduced audio) depends entirely on the quality of the system's design. Whoever said that all digital audio products sound alike never designed one or tried to manufacture it competitively (or listened very closely, either). Let's run through the trials and tribulations of the digitization chain.

A stereo PCM digitization scheme is shown in figure 2.15. The recording section consists of input amplifiers, a dither generator, input (anti-aliasing) low-pass filters, sample-and-hold circuits, analog-to-digital converters, a multiplexer, digital processing and modulation circuits, and a storage medium such as digital tape or optical disc. On the output side are de-modulation and processing circuits, a demultiplexer, digital-to-analog converters, sample-and-hold (aperture) circuits, output (anti-imaging) low-pass filters, and output amplifiers. This hardware collection is the realization of our previous discussions. As we'll see, the output side forms the basis for the design of a compact disc player.

### **Record Section**

A digitization system typically begins with conversion from analog to digital. Of course, in some applications, this is unnecessary since the information begins in digital form. In the case of music, analog-to-digital conversion is usually required because sound usually begins as an acoustic analog waveform.

The input analog amplifiers must maintain the requirement of absolute quality in order not to compromise the fidelity of the ensuing digital system. The dither generator is a controlled noise circuit. The analog signal is low-pass filtered by a very sharp cutoff filter to bandlimit the signal and its entire harmonic content to frequencies below the half-sampling frequency. A number of analog filter designs may be employed for this purpose; they all offer a flat pass-band, sharp cutoff, and attenuated stop-band.

Unfortunately, analog filter designs also create phase shift problems in which the relative timing of different frequencies is varied. For example, low frequencies may arrive at the output of the filter before the high frequencies. Analog filters are also complex circuits, tricky to build, and prone to drift. As with the input amplifier, the input filter must be of high quality; any distortion or noise would be digitized along with the audio signal.

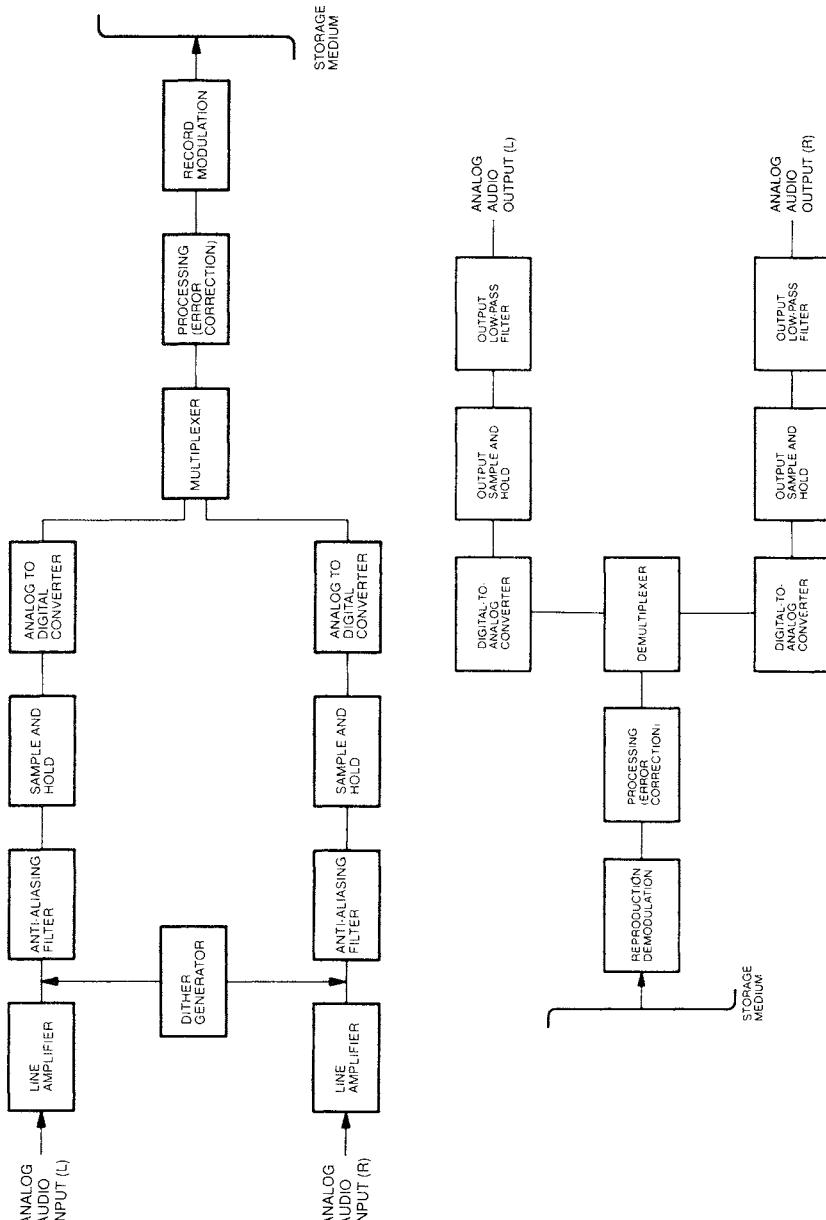


Figure 2.15 *The recording and reproduction sections of a conventional PCM system. (From Pohlmann 1989)*

Increasingly, oversampling input filters are being employed to permit use of lower-order anti-aliasing filters. We'll examine filters more carefully in Chapter Four.

The next circuit in the chain, the input sample-and-hold, samples discrete values of the input signal at a fixed periodic rate (the sampling rate) and holds each analog value while the analog-to-digital conversion takes place. This is required because a varying input might disrupt the analog-to-digital converter. A sample-and-hold circuit is essentially made of a capacitor and switch, such as the one in figure 2.16. The circuit tracks the signal until the sample command causes the switch to open, isolating the capacitor from the signal; the capacitor holds this analog voltage during conversion. The timing of the sample command must be carefully regulated to prevent jitter, the phenomenon of imprecise sample times. The capacitor must be carefully chosen to prevent any loss of voltage, known as droop.

Our signal now appears as a staircase, a hybrid analog signal ready for conversion to digital. The analog-to-digital (A/D) converter is the most

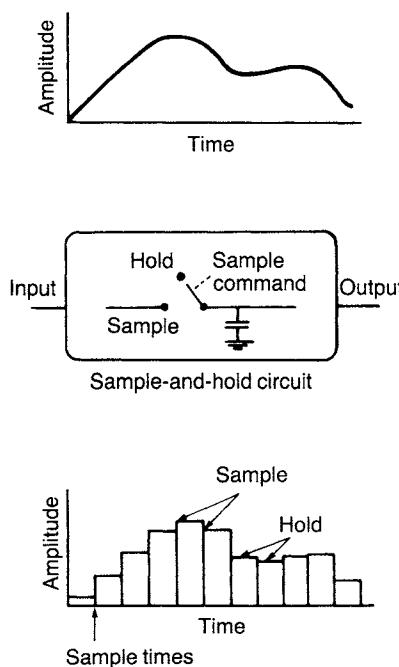


Figure 2.16 A sample-and-hold circuit samples the audio signal to form a staircase waveform.

critical and costly component in a digitization system. Consider that this circuit must resolve the held analog signal into 65,536 steps or more, each with its own binary word, and it must accomplish this task in 10 or 20 microseconds. Fortunately, several circuits are available for this operation. For example, a successive-approximation-register (SAR) A/D converter contains a digital-to-analog converter and a comparator. The SAR tries a digital word, converts it, compares the analog result to the original input, and then corrects its approximation one bit at a time until the proper digital word has been determined and output. Thus a 16-bit SAR A/D converter must make 16 D/A conversions for each A/D conversion. Integrating A/D converters offer another design approach. A capacitor stores the input analog voltage, and then a timer counts as the voltage is discharged. The number of counts in the interval forms the output digital value. Alternatively, low-bit A/D converters have been developed in which the audio signal is represented by only a few bits at a very high sampling rate. Whichever method is used, we accomplish our initial goal of quantizing the analog signal. The signal is now digital.

The output of the A/D converter is raw binary data ready to be processed as the designers see fit. However, a number of operations must occur. First, the A/D output is often parallel data in which all 16 bits of the data word appear at once on 16 lines; yet storage devices permit storage only of serial data, in which the bits appear one after another. Data is therefore converted from parallel to serial. In addition, the data for both left and right audio channels is multiplexed, or combined into one channel. Synchronization words, special non-audio bit patterns, are added to delineate the audio samples.

One of the system's most intriguing operations, error correction encoding, takes place at this point. In analog recording, an error is an error, and that's that. In digital we may provide for error detection and correction. The data stream, therefore, is provided with redundant information, extra data created from the original data to help detect and correct errors. We'll discuss error correction in more detail in Chapter Three. Finally, the data is modulated—that is, it is changed into a signal specifically suited for storage.

The storage medium itself is often a fixed or rotating head tape recorder, commonly found in the professional studio. Many recording engineers prefer fixed head designs for ease of editing and because of their familiarity with them from the analog days; multitrack recorders almost always used fixed heads. Alternatively, rotary head recorders, derived from video technology, are used and are more efficient because of their ability to store large amounts of data on a slowly moving tape. Increasingly, hard disk and CD recorders are entering the studio mainstream. No matter what happens in the studio (you're better off not knowing), the

data is eventually transferred to the compact disc for consumer playback. Both CD players and the professional's recorder follow the reproduction side of our digitization chain.

### ***Reproduction Section***

The initial reproduction circuits take care of many housekeeping functions, such as demodulation and demultiplexing. In this process, the data stream is recovered and is again put into parallel form. Our foresight in placing error detection and correction safeguards in the digital signal prior to storage pays off here. Errors introduced during playback are detected and corrected or concealed. This includes mechanical errors introduced by the recorder or player, such as transport wow and flutter, and media errors caused by media defects, dust, dirt, grape jelly, and fingerprints on the tape or disc.

Whereas analog-to-digital conversion is required only when the signal is originally analog, digital-to-analog conversion is always required, because our ears meaningfully respond only to analog waveforms. The digital-to-analog converter's task is the reverse of the analog-to-digital converter's, yet its operation is inherently simpler. Many CD players use low-bit D/A converters operating at very high oversampling rates. Multi-bit and low-bit D/A converters will be discussed in more detail in Chapter Four. Whichever type is employed, the D/A produces an analog voltage corresponding to each digital sample.

The output values from the D/A are held by a sample-and-hold circuit. The sample-and-hold circuit (as noted before, sometimes called an aperture circuit) essentially consists of a switch that is timed to wait for the D/A conversion. When the D/A output voltage is stable and any glitches have passed, the switch passes the voltage, switching on and off at the original sampling rate. The output of the sample-and-hold is a pulse amplitude staircase signal.

However, the analog signal from the sample-and-hold circuit cannot be output directly. The de-sampled waveform contains an infinite spectrum of images of the original spectrum. Viewed in another way, the sharp edges of the staircase waveform represent ultrasonic frequencies not present in the original waveform. Although the images are above the highest audible frequencies, they must be removed because they could cause intermodulation distortion, among other problems. The staircase must be smoothed, or filtered, to remove the ultrasonic frequencies.

Reconstruction of the audio waveform is completed by the output low-pass (anti-imaging) filter. Specifically, the filter's response to the individual impulses in the PAM staircase add together to re-create the original waveform. The anti-imaging filter can be substantially identical to the anti-

aliasing filter and performs the same cutoff function. The staircase function is smoothed, all the high frequency components of those sharp edges are removed, and the original analog waveform is recovered. In many designs, a digital oversampling filter is used. With oversampling, the number of samples is multiplied, and the images are thus extended in frequency, and a gentle analog filter can be used to remove them. Oversampling and waveform reconstruction will be discussed in more detail in Chapter Four.

The final part of our digitization system is the analog preamplifier, designed, one hopes, by analog designers, not digital specialists. Our PCM digitization and reconstruction system is complete.

### *Alternative Systems*

The architecture of a fixed, linear PCM system comprises the classic audio digitization design. However, in the interests of economy of either the cost or the data rate, the fixed, linear design may be modified. For example, a compander could be used. The compander changes the amplitude dynamics of the signal; its measurement during digitization is thus made nonlinear. Using techniques similar to those in analog noise reduction circuits, the signal may be compressed prior to the PCM A/D and expanded after the D/A to improve the S/N ratio of a lower-bit word system. In essence, low-level signals are measured with small increments, while high-level signals use large increments.

In other applications, such as satellite transmission, data reduction is critical because higher transmitted data rates add up to lower cost. A block floating point design might be the most efficient. A 16-bit A/D converter may be used to convert a memory block of samples. Then a calculation is made to determine the largest value in the block. Based on this value, all the words in the block are digitally attenuated by a calculated scale factor. Usually, the number of bits output is much lower than the number of bits originally in the converted block. The scale factor must accompany the data block for proper rescaling during D/A conversion. Because only one scaling factor is required for an entire block, and the number of bits in the block has been reduced, data reduction is accomplished. Alternatively, digitization systems may use digital signal processing techniques to achieve data reduction. Such systems often employ psychoacoustic modeling of the ear's response to optimize the reduction yet minimize degradation of the audio signal.

No matter which method is employed, a complete digital audio system isn't really all that difficult, at least in block diagram form. However, we haven't nearly exhausted the mysteries and amazing facts underlying the compact disc, as the remaining pages attest.

## FOR FURTHER READING

Blesser, B. "Digitization of Audio." *Journal of the Audio Engineering Society* 26(10):739–71, October, 1978.

Blesser, B., and B. Locanthy. "The Application of Narrow-Band Dither Operating at the Nyquist Frequency in Digital Systems to Provide Improved Signal-to-Noise Ratio over Conventional Dithering." *Journal of the Audio Engineering Society* 35(6):446–54, June, 1987.

Blesser, B., B. Locanthy, and T. G. Stockham, Jr., eds. *Collected Papers from the AES Premiere Conference—Digital Audio*. New York: Audio Engineering Society, 1983.

Hnatek, E. R. *A User's Handbook of D/A and D/A Converters*. New York: Wiley, 1976.

Mathews, M. V. *The Technology of Computer Music*. Cambridge: MIT Press, 1969.

Nakajima H., T. Doi, J. Fukuda, and A. Iga. *Digital Audio Technology*. Blue Ridge Summit, Penn.: Tab, 1983.

Nyquist, H. "Certain Topics in Telegraph Transmission Theory." *Transactions of the American Institute of Electrical Engineers* 47(2):617–44, April, 1928.

Pohlmann, K. C. *Principles of Digital Audio*. Indianapolis: Sams, 1989.

Pohlmann, K. C., ed. *Advanced Digital Audio*. Carmel, Indiana: Sams, 1991.

Vanderkooy, J., and S. Lipshitz. "Resolution below the Least Significant Bit in Digital Audio Systems with Dither." *Journal of the Audio Engineering Society* 32(3):106–13, March, 1984. Correction, *Journal of the Audio Engineering Society* 32(11):889, November, 1984.

Yoshikawa, S. "An Introduction to Digital Audio." In *Proceedings of the AES Third International Conference*, T. Yamamoto, chair. New York: Audio Engineering Society, 1985.



## *Chapter Three*

# The Compact Disc System

### INTRODUCTION

Storing audio information places great demands on a digital medium. A 60-minute musical selection, recorded in stereo with PCM (pulse code modulation) at a sampling rate of 44.1 kHz and with 16-bit quantization, generates over 5 billion bits. Error correction, synchronization, and modulation are required for successful storage. The total capacity required is over 15 billion bits (that's 15 thousand million). In addition, a commercially popular music storage medium requires random access, small size, convenience of use, robustness, low cost, and ease of replication. Clearly, digital audio's requirements are considerable.

The compact disc format was developed to meet these demands. A CD audio disc holds a specification-defined maximum of 74 minutes (variations can increase this to 80 minutes or more) of high fidelity music on a highly robust and economically manufactured disc. A CD player provides access to any part of the audio program within a second or less, in addition to having many user-convenience features. Playback hardware specifications have evolved to the point where they surpass most listeners' ability to reliably distinguish between players. The format is well suited to audio storage demands.

The magnitude of the challenge of digital audio storage and the magnitude of the achievement of the compact disc have resulted in a storage medium of unprecedented capacity, durability, and flexibility. Thus the compact disc is highly suitable for many non-audio applications. As a result, a number of alternative CD formats have been developed. However, all the CD incarnations, which will be discussed in Chapter Six, utilize the same basic CD specifications that are used in audio CDs. To avoid confusion, the audio CD system is sometimes called CD-DA, for Compact Disc Digital Audio, or simply the CD-Audio format.

## SYSTEM OVERVIEW

Music input to a digital recording and reproduction chain passes through the recording side of a PCM system, as described in Chapter Two. The master recording is stored in digital form on magnetic tape or optical disc. After recording, mixing, and editing, the master is ready to be reproduced via the compact disc. The disc manufacturing facility converts the master tape to a master disc and replicates the disc to produce many CD copies. A CD player functions as the reproduction side of the PCM system, outputting the originally recorded music.

When digital technology is used for all the intermediate steps, a sophisticated chain of events transpires—that is, the music is converted to binary code when it enters the recording studio (and stays in binary code throughout the chain) and is converted back to analog form as it leaves the CD player.

Indeed, many CD players have digital outputs so data is preserved in that form until it reaches the power amplifier. In other words, a compact disc contains the same audio information, bit for bit, that was recorded in the studio. To a large degree, you are hearing exactly what was recorded by the engineer. This, of course, is what music recording is all about.

Specifications for the compact disc system are shown in table 3.1. These parameters were jointly developed by Philips and Sony and are defined in the standards document known as the *Red Book*. The CD standard is also contained in the IEC (International Electrotechnical Commission) standard BNN15-83-095 and IEC Publication 908, “Compact Disc Digital Audio System.” Disc and player manufacturers must obtain a compact disc license to utilize these specifications. Disc dimensions are defined in the standard, as are pit dimensions, the physical formations which encode data. Although the disc itself may be made of any transparent material with a refraction index (N) of 1.55, polycarbonate plastic is generally used.

The optical system that reads data from the disc uses a laser beam with a 780 nanometer wavelength. Other optical properties are defined in the CD standard as well. The basic specifications pertinent to player design are contained in the signal format specifications. Sampling frequency, quantization word length, data rate, error correction code, and modulation scheme are all defined in the CD standard (and discussed in detail in the pages to follow). Player manufacturers have considerable freedom in the actual design of their players; this accounts for the wide variety in player features and quality. Similarly, disc manufacturers enjoy a certain degree of latitude in tolerance.

## OPTICAL STORAGE

The physical characteristics of the compact disc are shown in figure 3.1. Disc diameter is 120 millimeters, hole diameter is 15 millimeters, and thick-

**DISC**

Playing time:	74 minutes, 33 seconds maximum
Rotation:	Counter-clockwise when viewed from readout surface
Rotational speed:	1.2–1.4 m/sec.
Track pitch:	1.6 $\mu\text{m}$
Diameter:	120 mm
Thickness:	1.2 mm
Center hole diameter:	15 mm
Recording area:	46 mm–117 mm
Signal area:	50 mm–116 mm
Material:	Any transparent material with 1.55 refraction index, such as polycarbonate
Minimum pit length:	0.833 $\mu\text{m}$ (1.2 m/sec.) to 0.972 $\mu\text{m}$ (1.4 m/sec.)
Maximum pit length:	3.05 $\mu\text{m}$ (1.2 m/sec.) to 3.56 $\mu\text{m}$ (1.4 m/sec.)
Pit depth:	Approx. 0.11 $\mu\text{m}$
Pit width:	Approx. 0.5 $\mu\text{m}$

**OPTICAL SYSTEM**

Standard wavelength:	$\lambda = 780 \text{ nm} (7,800\text{\AA})$
Focal depth:	$\pm 2 \mu\text{m}$
$(\lambda/NA \leq 1.75 \mu\text{m}, NA: \text{Numerical Aperture})$	

**SIGNAL FORMAT**

Number of channels:	2 channels (4-channel recording possible)
Quantization:	16-bit linear quantization
Quantizing timing:	Concurrent for all channels
Sampling frequency:	44.1 kHz
Channel bit rate:	4.3218 Mb/sec.
Data bit rate:	2.0338 Mb/sec.
Data-to-channel bit ratio:	8:17
Error correction code:	CIRC (with 25% redundancy)
Modulation system:	EFM

Table 3.1     *Specification table for the compact disc system.*

ness is 1.2 millimeters. The innermost part of the disc does not hold data; it provides a clamping area for the player to hold the disc firmly to the spindle motor shaft. Data is recorded on an area 35.5 millimeters wide.

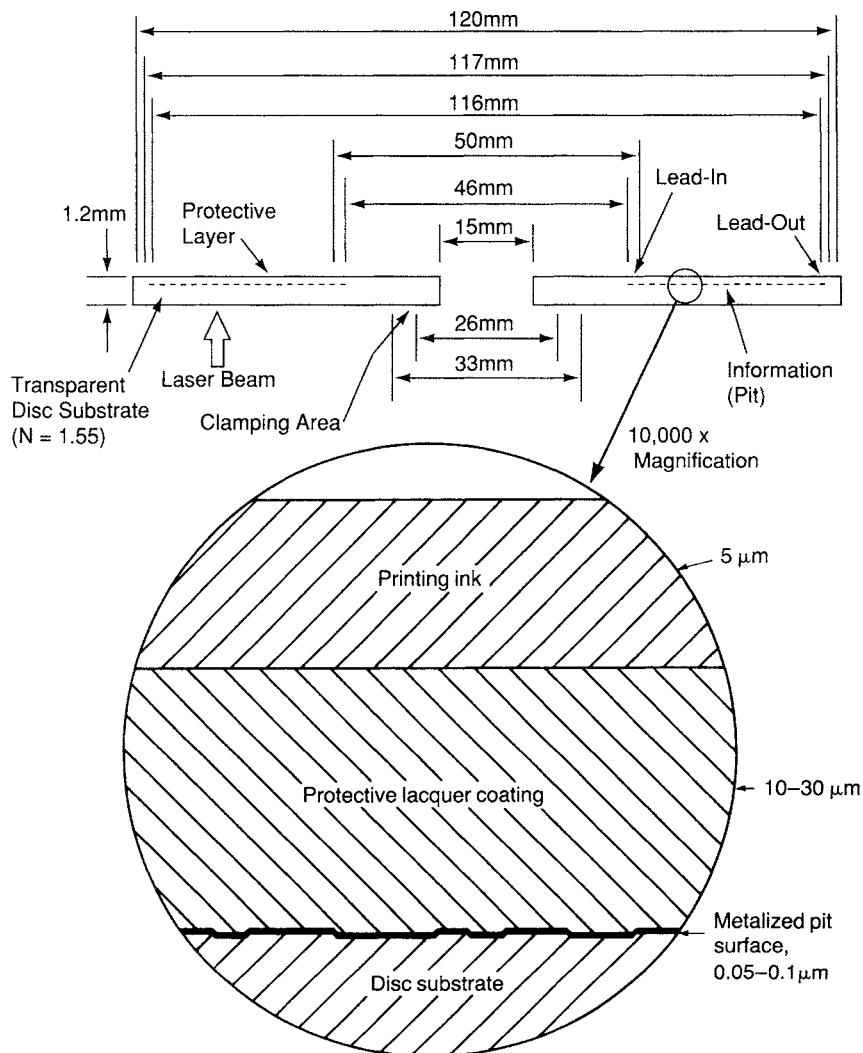


Figure 3.1 *Physical specifications and construction of a compact disc.*

A lead-in area rings the innermost data area, and a lead-out area rings the outermost area. The lead-in and lead-out areas contain non-audio data used to control the player. On many discs, a change in appearance in the reflective data surface marks the end of the musical information.

A transparent plastic substrate forms most of a disc's 1.2 millimeter thickness. Data is physically contained in pits which are impressed along its top surface and are covered with a very thin (50 to 100 nanometers) metal (e.g., aluminum, silver, or gold) layer. Another thin (10 to 30 micrometers) plastic layer protects the metallized pit surface, on top of which the identifying label (5 micrometers) is printed. A laser beam is used to read the data. It is applied from below and passes through the transparent substrate and back again. The beam is focused on the metallized data surface embedded inside the disc. Since data on a disc is read by a light beam, playing a CD causes no more wear to the recording than your reading causes to the words printed on this page.

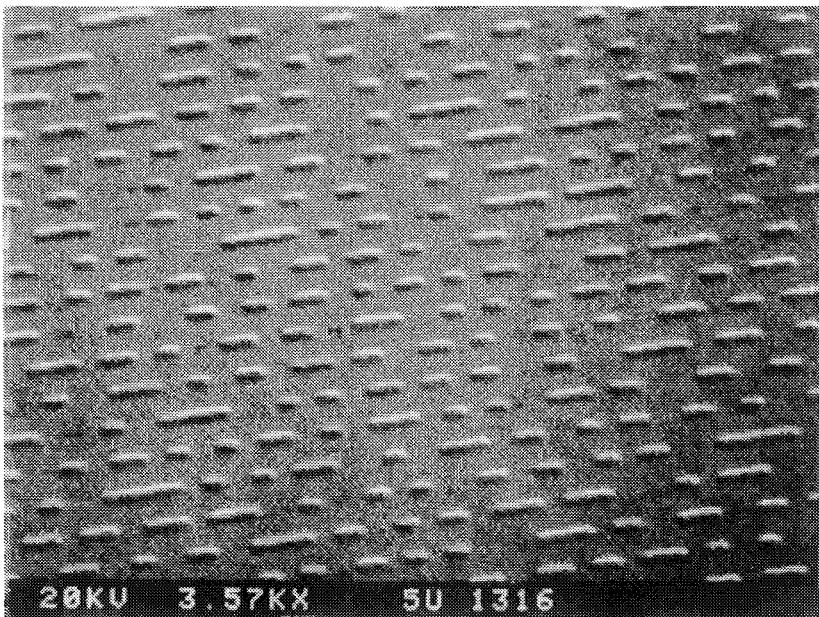
### ***The Pit Track***

The physical method of accomplishing data storage in pits on a flat surface is not directly visible to the naked eye. You would need a scanning electron microscope (such as the one used to take figure 3.2) to get a good look at the pits. The pits are amazingly small, much smaller than we bulky humans can appreciate. For example, a pit is about 0.5 micrometers wide, or about 700 times smaller than a pin prick.

A disc contains a track of pits arranged in a continuous spiral running from the inner circumference to the outer. It is advantageous to start the music on the inside because the outer diameter of a disc, in some manufacturing processes, is more generally prone to manufacturing defects. CDs with shorter playing time thus provide a greater manufacturing yield. In addition, by starting from the inside, adoption of smaller diameter discs (such as 8-centimeter CD-3 discs) or larger diameter discs (such as 20- and 30-centimeter CD-Video discs) is facilitated.

Figure 3.3 diagrams the track pitch, the distance between successive tracks, which is 1.6 micrometers (that's about 600 tracks to a millimeter). There are 22,188 revolutions across the disc's signal surface of 35.5 millimeters. A pit track might contain 3 billion pits. Unspiraled, the track would stretch about 3½ miles. The construction of the CD is so small that its track pitch acts as a diffraction grating, producing a rainbow of colors. In fact, CD pits are among the smallest of all manufactured formations.

A simple hands-on experiment suggested by Nöldeke can be used to demonstrate diffraction caused by the pit track and to measure a CD's track pitch. Anyone who's ever wondered what hallucinogenic drugs are like will definitely want to try this experiment. This effect alone will

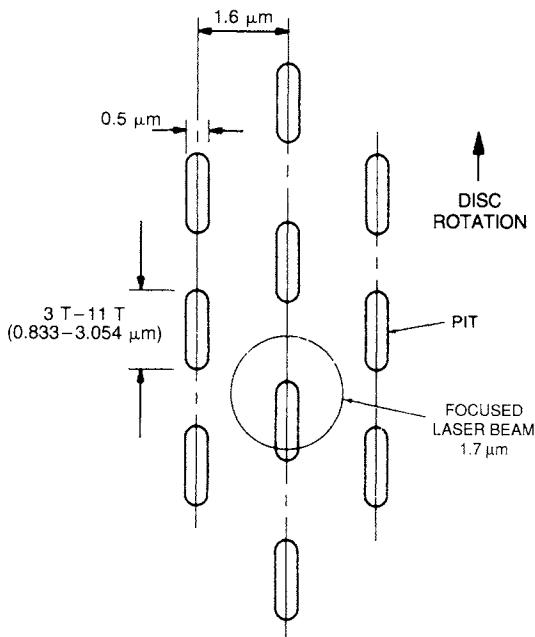


---

Figure 3.2    *Compact disc pit surface. The horizontal line indicates a scale of 5 micrometers.* (Photo: University of Miami)

compensate you for the price of the book and will make you a hit at parties. Fetch a CD from your collection, a ruler from your desk, and seat yourself with a light bulb (60 watts or so) about 2 meters (6 feet) behind you. Hold the CD about a foot from your face (reflective side facing you) and angle it so the light bulb's reflection disappears in the disc center hole. Now slowly move the disc toward your face, keeping the bulb's reflection centered in the hole. The bright psychedelic display is a result of the diffraction created by the pit track.

Now move the disc away from you until the violet ring is located on the edge of the CD. If you measure this, you may find it to be about 20 centimeters from your eye. Now we can calculate the track pitch using the equation  $d = \lambda(1 + a^2/r^2)^{1/2}$  where  $d$  is the track pitch,  $\lambda$  is the wavelength of the violet light (450 nanometers),  $a$  is the distance away (20 centimeters), and  $r$  is the radius of the violet ring (5.5 centimeters). Calculating, we obtain a track pitch value of 1.7 micrometers—only a ten-millionth of a meter away from the actual pitch (on most CDs) of 1.6 micrometers!




---

Figure 3.3    *Track pitch is 1.6 micrometers. Data is read through an intensity-modulated laser beam.*

Further examination of a pit track would reveal that the linear dimensions of the track are the same at the beginning of its spiral as at the end. This means that a CD must rotate with CLV (constant linear velocity), a condition in which a uniform relative velocity is maintained between the disc and the pickup.

To accomplish this, the rotation speed of a compact disc varies depending on the position of the pickup. Because each outer track revolution contains more pits than each inner track revolution, the disc must be slowed down as it plays in order to maintain a constant rate of data. The disc rotates at a speed of 500 rpm when the pickup is reading the inner circumference, and as the pickup moves outward, the rotational speed gradually decreases to 200 rpm. Thus a constant linear velocity is maintained. In other words, all of the pits are read at the same speed, regardless of the circumference of that part of the spiral. The track is a CLV servo system; the player constantly reads frame synchronization words from the data and adjusts the disc speed to keep the data rate constant.

The compact disc's CLV system differs significantly from LPs. A turntable's motor rotates an LP at a CAV (constant angular velocity) of 33½ rpm. As a result, outer grooves have a greater apparent velocity than inner grooves. One ramification of this is that the high-frequency response of inner grooves is markedly inferior to that of outer grooves. If the CD used CAV, pits on the outside diameter of the disc would have to be longer than those on the inner diameter. This would decrease data density and decrease playing time.

Although the CLV of any particular compact disc is fixed, the CLVs used on different discs can range from 1.2 to 1.4 meters/second. In general, discs with playing times of less than 60 minutes are recorded at 1.4 meters/second, while discs with longer playing times use a slower velocity, to a minimum of 1.2 meters/second. The CD player is indifferent to the actual CLV; it automatically regulates the disc rotational speed to maintain a constant bit rate of 4.3218 MHz.

### ***Data Readout***

The fact that the laser beam travels through the disc substrate provides one of the most significant assets of the CD system. When light passes from one medium to another with a different index of refraction, its wavelength changes and it bends. For example, drop a pencil in half a glass of water; the pencil appears to be broken because the air/water boundary has bent the light. The velocity of light decreases when it passes from air to water.

Figure 3.4 shows that a pickup's laser beam responds similarly. The plastic substrate has a refractive index of 1.55 (as opposed to 1.0 for air); the velocity of light slows from  $3 \times 10^5$  kilometers/second to  $1.9 \times 10^5$  kilometers/second. When the velocity of light slows, the beam is bent, and focusing occurs. Because of the refractive index, the thickness of the disc, and the numerical aperture (0.45) of the laser's lens, the size of the laser beam on the disc surface is approximately 800 micrometers but is focused to approximately 1.0 micrometers at the pit surface. The laser beam is thus focused to a point about two times larger than the pit width. Moreover, the effects of any dust or scratches on the substrate's outer surface are minimized because their size (and importance) at the data surface are effectively reduced along with the laser beam. Specifically, any obstruction less than 0.5 millimeters becomes insignificant and causes no error in the readout.

The entire pit surface, as noted, is metalized. The reflective flat surface, called land, causes almost 90 percent of the laser light to be reflected back into the pickup. When viewed from the laser's perspective (underneath), the pits appear as bumps. The height of each bump is between 0.11 and 0.13 micrometers. This dimension is slightly smaller than a care-

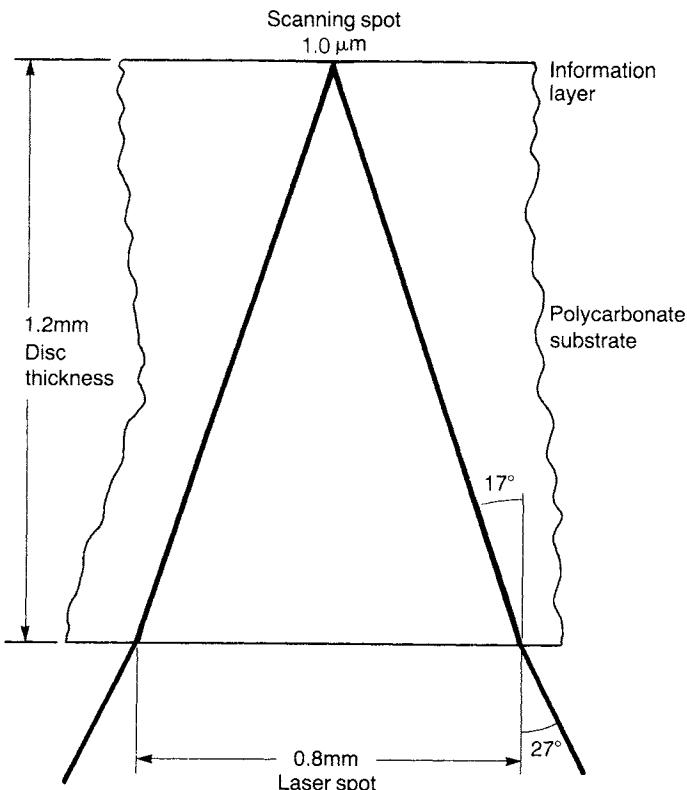
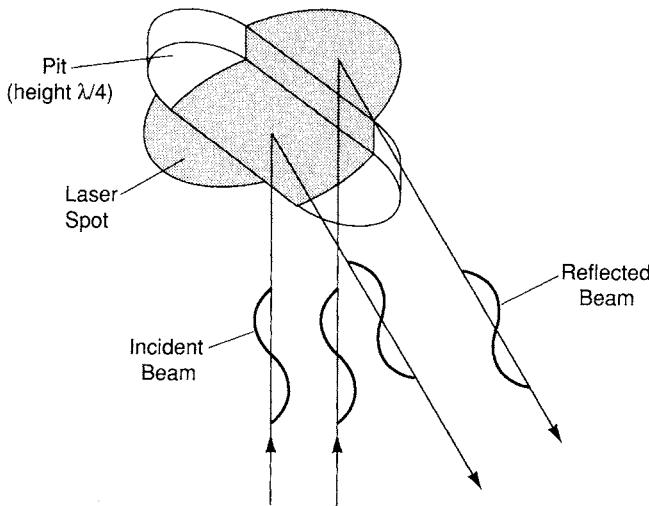


Figure 3.4     *The refractive index of the disc substrate contributes to the focusing of the laser beam.*

fully chosen dimension: the laser beam's wavelength in air is 780 nanometers (some players use 790 nanometers). Inside the polycarbonate substrate, with a refractive index of 1.55, the laser's wavelength is about 500 nanometers. The height of the bumps is thus approximately one-quarter of the laser's wavelength in the substrate.

Assume the height is exactly one-quarter the laser's wavelength. Light striking land thus travels at a distance one-half the wavelength (one-quarter plus one-quarter) further than light striking a bump. This creates a phase difference of one-half wavelength between the part of the beam reflected from the bump and the part reflected from the surrounding land, as shown in figure 3.5. The phase difference causes the two parts of the beam to destructively interfere with and cancel each other, forming a diffraction pattern. In short, a bump disperses light, reducing the intensity of the reflected light.




---

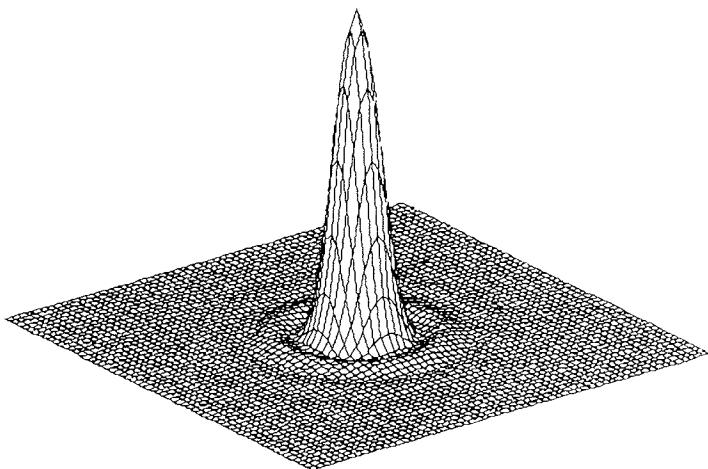
**Figure 3.5** *A bump (pit) height causes a one-half wavelength path difference relative to surrounding land.*

In theory, when the beam strikes an area between pits, virtually all of its light is reflected, and when it strikes a pit, virtually all of its light is cancelled, so that virtually none is reflected. In practice, the laser spot is larger than is required for complete cancellation between pit and land reflections, and pits are made slightly shallower than the theoretical figure of one-quarter wavelength; this yields a better tracking signal, among other things. About 25 percent of the power of the incident light is reflected from a long bump. In any case, the presence of pits and land is thus read by the laser beam; specifically, the disc surface modulates the intensity of the light beam. Thus the data physically encoded on the disc can be recovered by the laser and later converted to an electrical signal.

### ***A Closer Look at Optics***

Many of us studied optics in high school physics class, then promptly forgot it all after the final exam. Let's review some optics and see how the CD functions as an optical device. On the other hand, if memories of high school physics are simply too painful, feel free to skip to the next section.

Although the compact disc pit surface and laser readout may be explained in fairly simple terms, they are in fact a very sophisticated optical system that requires knowledge of optics to understand fully. The numerical aperture of the lens, wavelength of the laser light, thickness and



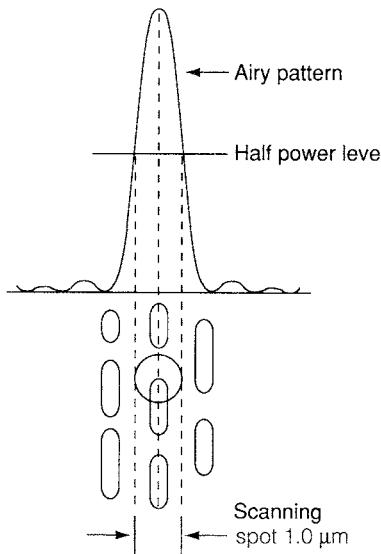
---

Figure 3.6    *The laser beam focused on the CD data surface forms an intensity function known as an Airy pattern. (Courtesy Robert Dunn)*

refractive index of the disc, and size and height of the pits all work together to allow data to be read from the disc.

Diffraction occurs when a single slit is placed in a barrier. From the slit, a new wavefront will radiate as though it came from a point source. A diffraction grating contains a series of slits. Because of interference, a wavefront will leave the grating only in directions where light from all the slits is in phase. This happens straight ahead and at certain other angles. The angle of the first oblique wave is a function of the wavelength of the light and the spacing of the slits. If the spacing of the slits is reduced, the oblique ray will leave at a greater angle to the center line. Fine detail can be resolved only if the diffraction wavefront is collected by the lens. The resolving power of the lens is determined by its numerical aperture (NA).

Even if the lens is perfectly free of aberration, a laser beam cannot be focused to a point with a circular lens of finite aperture. The laser spot focused on the disc data surface is actually an area where the intensity of light varies as a function called the Airy pattern, as shown in figure 3.6. This is a circular diffraction pattern of maximum and minimum light intensities corresponding to different phase differences. The central spot is the zero-order maximum, and the surrounding rings are higher-order maxima. The size of the Airy pattern is a function of the light wavelength and the numerical aperture of the lens. The laser pickup optics are thus



**Figure 3.7** *The reading laser forms an Airy pattern on the data surface. The half intensity spot covers the pit track.*

diffraction-limited as opposed to tolerance-limited. As dictated by the Maréchal criterion, a higher quality lens would still result in the same spot size. For the same reason, any future high-density CD with smaller pit size would demand a laser pickup operating at a shorter wavelength.

The laser spot diameter of 1.0 micrometers is typically specified as the point of half-intensity on the Airy pattern. The first dark ring of the Airy pattern is at 2.1 micrometers, which falls roughly on the adjacent CD pit track, as shown in figure 3.7. This provides good crosstalk performance. Allowable crosstalk between tracks determines track pitch—nominally set at 1.6 micrometers. Crosstalk must be acceptable even in the worst case of a slightly defocused beam and slight tracking error.

The CD surface is a reflection grating and acts similarly to a diffraction grating. Specifically, the pits on the CD cause diffraction to take place. The smaller the pits in relation to the light wavelength, the greater is the angle at which light leaves. Light diffracted by the grating consists of a single zero-order and multiple first-order beams, as shown in figure 3.8. They will partly overlap each other, resulting in cancellation due to interference.

As noted, the area of light striking a pit is about equal to that striking the surrounding land, and the pit is about one-quarter wavelength in height. Light striking the pit is out of phase with that striking the land. This results in destructive interference between the zero- and first-order beams in the

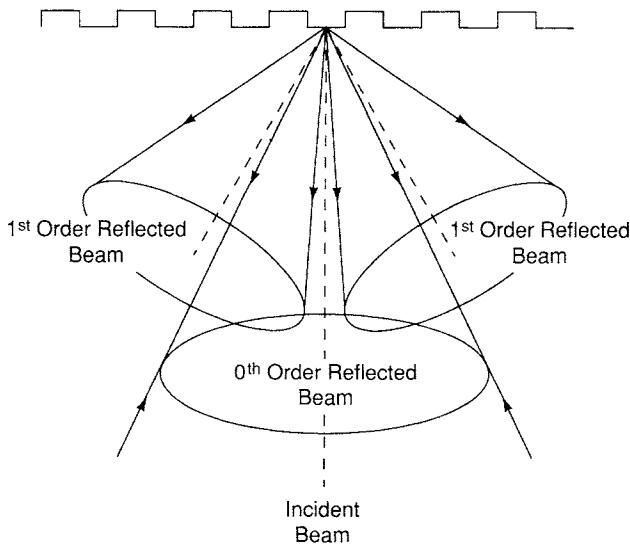


Figure 3.8    *The CD pit surface acts as a reflective grating. Diffraction causes cancellation through interference between reflected beams.*

light ray returning to the lens, and constructive interference along other rays, primarily along a disc radius. For good contrast, the power in the two rays with different phases should be equal. Effectively, a pit reduces the intensity of light returning to the lens. Also as noted, maximum modulation requires a half-wavelength difference and equal intensities of the pit/land reflections. However, pit depth does not require great accuracy. It is more important to achieve equality of pit/land beam areas.

The system is diffraction-limited; that is, it operates at dimensions that are as small as permitted by the wave nature of light at that wavelength. With a spot diameter of approximately 1.0 micrometer, details this size would result in diffracted rays that only just fall within the lens aperture. Finer details yield greater convergence; thus rays diffracted by the pits must fall outside the lens aperture. This is in fact the case with the pits on the data surface since they are narrower than the diameter of the spot. Rays diffracted by the pits must consequently fall outside the lens aperture, decreasing intensity reflected back into the lens. This diffraction-limiting makes it impossible to discern the contours of a pit surface with an optical microscope; a scanning electron microscope is needed.

So much for CD optics, whether visible to the naked eye or not. Not visible at all is the format by which the data is physically encoded on the

disc. For example, contrary to what one might expect, the pits and intervening reflective land on the CD surface do not directly designate 1s and 0s. Rather, each pit *edge*, whether leading or trailing, is a 1, and all areas in between, whether inside or outside a pit, are 0s. This is a much more efficient storage technique than coding the binary bits directly with pits.

Still less obvious are the curiously varying lengths of the pits. The pit and the reflective land lengths vary incrementally, with nine different pit and land lengths. As we shall see, combinations of these varying dimensions physically encode the data.

## **ERROR CORRECTION**

Error correction is one of the great opportunities—and obligations—of digital audio storage media such as the compact disc. It is a great opportunity because it offers something never before possible with analog media—the chance to correct mistakes. When you scratch a long playing record, the grooves are irrevocably damaged, along with the information contained in them. Forever after, there will be a click or pop when the damaged part of the groove passes underneath the phonograph needle.

But when you scratch a CD, the nature of the data on the disc and the player's design offer you a second chance. Specifically, the data on the disc has been specially encoded with an error correction code, and your player uses the code to correct for damaged data. Thus it delivers the original undamaged variety instead, performing its error correction every time the disc is played.

To illustrate a fundamental aspect of error correction, consider the two messages below:

1. Please turn to page 345 in your textbook.
2. Hey gang! Let's party! Hey gang! Let's party!

Assume the first message represents the information in an LP groove and the second is the data in a CD pit track. Now, place one of your fingers vertically across this page, representing a scratch. You'll observe that part of the first message is irrevocably gone, whereas the second message is intact, because it can be reconstructed. Specifically, we have used redundant data to protect the message. By comparing the redundant messages, we can overcome the effect of the error. Note that we haven't prevented the error; we have simply insured against its effect. Clever, eh?

Of course, even the most clever of ideas has its limitations. As you place more fingers over the page, it becomes more difficult to reconstruct the digital message successfully. Thus the severity of the error plays an

important role; our ability to correct errors has limitations. Furthermore, if you place your finger horizontally across the page, you might completely obscure the entire line, destroying both the message and its redundancy. The nature of the error thus plays a part too. Finally, we should note that error correction exacts a cost; in this case, the message coded with error correction requires twice the storage space.

On the other hand, error correction is an obligation. Digital audio on a compact disc might require storage of fifteen billion bits. With such great data density, even the smallest speck of dust would wipe out a considerable number of bits, as shown in figure 3.9. In reality the type of error correction utilized would have to be a great deal more sophisticated than a simple redundancy scheme. Even one bad bit could wreak havoc. For example, if the digital word 0000000000000000 (representing silence) was misread as 1000000000000000 (representing a pretty loud level), a click some 90 dB above the silence would result.

### Detection

Before error correction takes place, the errors must be detected. While this might sound obvious, the problem can be difficult. If presented with a data word, could you tell whether or not an error had occurred? For example, does 1100101000011110 contain an error? Unless you are psychic, there is no way to tell. The message could be sent twice: 1100101000011110 1100101100011110, and close examination would reveal a difference between them. Obviously, both cannot be correct, but which one is the right one? The message could be sent three times: 1100101000011110 1100101100011110 1100101000011110, and you might have a good suspicion of which one was correct, but would you be sure? What if the message were sent three times, and they were all different?

While simple repetition is an inefficient way of going about error detection, a more enlightened variation, data redundancy, is the essence of

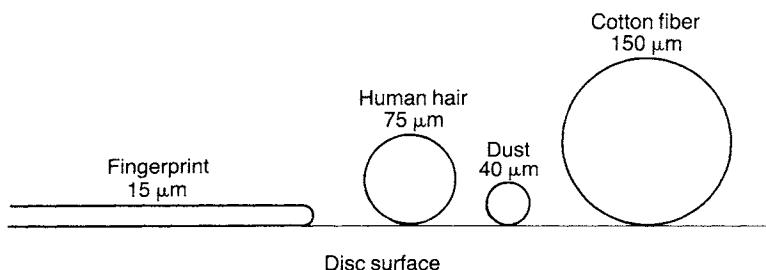


Figure 3.9 Even minuscule obstructions such as dust are large compared to the 0.5 micrometer-wide data pits.

error correction. Data redundancy is extra information derived from the original information; hence it is redundant. In general, the greater the redundancy, the better the error correction. As we have seen, redundancy adds to storage overhead and must be optimized in a practical system; this leads to the development of elaborate error correction codes which make very efficient use of redundancy by coding the information in certain ways.

### ***Correction***

In practice, the redundancy contained in correction codes often takes the form of a parity bit added to every data word. The parity bit is chosen so the number of 1s or 0s in the group (data word plus parity bit) is even or odd. For example, a parity bit may be formed with this rule: If the number of 1s in the data word is even (or zero), the parity bit is made 0; if the number of 1s is odd, the parity bit is made 1. The total number of 1s will always be an even number (or zero); this is called even parity. By the same token, odd parity could be used.

This scheme allows for error detection of received data: Any word with an odd number of 1s must be invalid. Odd parity can detect 1-bit errors in data (actually, it detects all errors involving an odd number of bits). However, if two or any even number of bits are bad, the scheme fails to detect the error, and a bad parity bit would cause good data to be flagged as erroneous. Moreover, the scheme cannot ascertain which bit is bad. Thus it cannot correct errors. Clearly, a more sophisticated scheme is needed.

For correction, in essence, more parity bits are required to solve the problem, and the algorithm used to select the parity bits determines the performance of the error correction code. For example, in some codes the data is divided into blocks, and parity values are added to each block. Consider the example in figure 3.10. In addition to the sixteen data values in 3.10a, nine extra parity values are created and appended to the original data block. They are placed at the end of each row and column and form the sum of that row and column, as shown in 3.10b. Note that a parity value is also included for the parity row and column. If an error occurs in any data (or parity) value, the error can be easily located, and the correct value can be easily calculated, using the other data present.

For example, suppose that the data block in Figure 3.10c was received. As we recalculate each parity value at the receiving end and check it against each transmitted parity value, we observe a disagreement. In fact, there is a disagreement in both a row and a column parity value. The intersection of the row and column points to the error. Furthermore, we can now substitute the correct value, derived from the transmitted parity values. The data in the block is thus substantially more reliable. Of course, instead of twelve values, twenty are now required for the error code. In

a)

10	04	04	02	
05	04	11	02	
01	12	06	03	
00	07	12	06	

b)

10	04	04	02	20
05	04	11	02	22
01	12	06	03	22
00	07	12	06	25

c)

10	04	04	02	20
05	04	11	02	22
01	(18)	06	03	22
00	07	12	06	25

↓

column error

data error

row error

Figure 3.10 A data block (a) may have parity values attached to each row and column (b). In the event of an error (c), parity may be used to correct the error.

practice, a code may take the form shown in figure 3.11. Several parity bits (a) are generated from combinations of input data bits. Transmitted data is parity-tested, and odd parity indicates an error. Using the results of the parity check, a logic table may be used to determine which bit is incorrect. If data 0011 is received as 0111 (c), parity indicates an error in bit 2.

Numerous error correction codes have been devised. For example, Hamming codes derive multiple parity bits from combinations of the data bits. Simple Hamming codes can be constructed which will detect two errors and correct one error. Similarly, Reed-Solomon codes can detect and correct large numbers of errors. Using codes such as Reed-Solomon, a digital audio system can detect and correct errors and supply information that is as good as new. As we will see, the compact disc system uses a Reed-Solomon error correction code as part of its defense against errors.

However, as we have seen, the performance of the error correction system depends on the nature of the error. For example, what if a large

a)

	bit 1	bit 2	bit 3	bit 4	parity 1	parity 2	parity 3
	0	0	1	1	1	1	0

b)

	bit 1, bit 2 bit 3, parity 1	bit 1, bit 2 bit 4, parity 2	bit 1, bit 3 bit 4, parity 3	location of error
even	even	even	even	none
odd	even	even	even	parity 1
even	odd	odd	even	parity 2
even	even	odd	odd	parity 3
even	odd	odd	odd	bit 4
odd	even	odd	odd	bit 3
odd	odd	odd	even	bit 2
odd	odd	odd	odd	bit 1

c)

bit 1 = 0	bit 1 = 0	bit 1 = 0
bit 2 = 1	bit 2 = 1	bit 3 = 1
bit 3 = 1	bit 4 = 1	bit 4 = 1
parity 1 = 1	parity 2 = 1	parity 3 = 0
odd	odd	even

Figure 3.11 An error correction code uses combinations of data bits to form multiple parity bits.

error obliterated both the data and its parity? There would be nothing left to reconstruct the message. It is thus important to understand the nature of the errors and devise protection to fit their nature. Errors can occur in large groups, called burst errors, or in isolated instances, called random errors. The compact disc must guard against both kinds; for example, a badly formed pit could cause a random error, whereas a dust particle could cause a burst error. Clearly, burst errors are the most troubling. A good error correction code uses parity in addition to other processing, such as interleaving.

### ***Interleaving***

Interleaving is employed to guard against the very likely occurrence of burst errors. Interleaving might be thought of as shuffling a deck of cards; data symbols are redistributed in the bit stream prior to recording so that consecutive words are never adjacent on the medium. An error occurring in the medium (for example, caused by a dust particle on the disc) might prevent the successful reading of a number of consecutive values. However, upon de-interleaving the shuffled words are placed back in their original and rightful position in the stream, and the errors are scattered in time. Thus isolated, they are much easier to correct. Interleaving is particularly useful for long burst errors; consecutive errors are scattered by de-interleaving, becoming more like random errors, which are more easily corrected. An interleaving example is shown in figure 3.12. Interleaving appears complicated, but it can be accomplished by simply delaying the data words by differing amounts prior to recording. Delaying the data words again (in a complementary manner) upon playback completes the technique.

Cross interleaving carries the idea one step further. Data is interleaved numerous times, over both short and long time intervals. This provides correctability for larger errors. As we shall see, the CD system uses the Cross Interleave Reed-Solomon code (CIRC) for error protection. It employs parity checking to correct random errors and cross interleaving to permit parity to correct burst errors.

### ***Concealment***

Although correction of massive errors is possible, it would be impractical to implement. In real-life digital audio systems, some errors overwhelm the error correction scheme. These errors are flagged by the correction circuits and passed on to error concealment circuits.

Without concealment, erroneous data that escaped error correction could result in an audible click. Error concealment systems employ interpolation and muting circuits following the CIRC decoder. Although an important part of error protection, they are unrelated to the CIRC. Specifically, they attempt to handle errors too massive for CIRC correction. Using error flags from the CIRC decoder, the player's concealment circuits determine whether to output the data directly, to interpolate it, or to mute the output.

Interpolation is the technique of using valid data surrounding an error as a basis of forming a replacement for the erroneous data. These values are used to calculate a new value to substitute for the error. Because of the high correlation between music samples, an uncorrected error can be

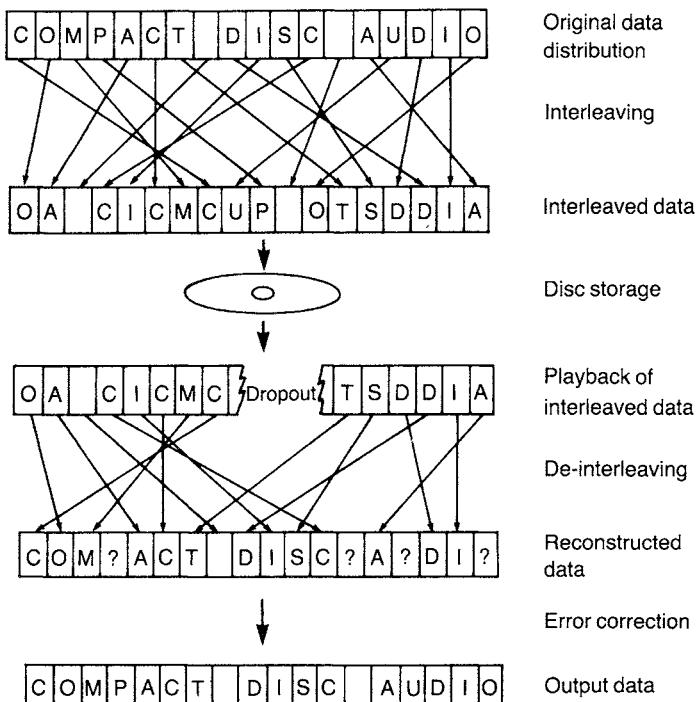


Figure 3.12 When interleaved data is de-interleaved, consecutive errors are scattered, making correction easier.

made virtually inaudible by synthesizing new data from surrounding data. Although clicks are avoided, a momentary increase in distortion is produced.

Numerous interpolation schemes are used, with different performance levels. In its simplest form, zero-order interpolation holds the previous value and repeats it to cover the missing or incorrect word. In first-order interpolation, the erroneous word is replaced with a word derived from the mean value of the previous and subsequent words.

In worst case scenarios, where the error is so massive that even interpolation would fail, we choose to mute the audio signal. The brief silence is preferable to the burst of digital noise, usually heard as a click. By attenuating the signal before and after the mute, even these catastrophic errors are often made inaudible to most listeners.

However, the digital audio signal cannot be muted by switching the bit stream to zero; this could result in an audible click. Muting methods vary

from player to player. For example, the signal may be faded down by multiplying the samples by descending coefficients, usually taken from a half cycle of a cosine waveform. The fade-out must begin prior to the bad data; this is accommodated by feeding the signal through a delay before the muting circuit. The mute signal thus arrives before the bad data. Following the bad data, a fade-in is similarly accomplished. Smooth mutes are surprisingly inaudible.

### CROSS INTERLEAVE REED-SOLOMON CODE (CIRC)

As we have seen, error correction is essential to the success of digital audio. Without it, any digital recording on tape or disc would sound like a badly scratched LP at best or, at worst, be simply unplayable. In fact, without error correction, digital audio would be an impossibility.

The raw error rate from a CD is around  $10^{-5}$  to  $10^{-6}$ , or about one error for every 0.1 to 1.0 million bits. This is impressive storage capability, but considering that a disc will output over 4 million bits per second, the need for error correction is obvious. With error correction, perhaps 200 errors per second will be completely corrected. To achieve such results, the CD employs the cornerstones of error correction: interleaving to distribute errors and parity to correct them. The particular algorithm used for error correction in the compact disc system is, as mentioned earlier, the Cross Interleave Reed-Solomon Code (CIRC). The CIRC circuit uses two correction codes for additional correcting capability and three interleaving stages to encode data before it is placed on a disc. Similarly, CIRC performs error correction while decoding the data during playback.

#### *Reed-Solomon Codes*

The Reed-Solomon code used in CIRC is a highly efficient error correcting code. It is particularly well suited for the compact disc system because its decoding requirements are relatively simple. To illustrate the operation of Reed-Solomon codes, suppose that  $A$ ,  $B$ ,  $C$ , and  $D$  are four data words, and  $P$  and  $Q$  are two parity words. The  $\oplus$  symbol denotes modulo 2 addition (EX OR), and the alphas are weighting factors. The data words are placed into two simultaneous equations:

$$\begin{aligned} A \oplus B \oplus C \oplus D \oplus P \oplus Q &= 0 \\ \alpha^6A \oplus \alpha^5B \oplus \alpha^4C \oplus \alpha^3D \oplus \alpha^2P \oplus \alpha Q &= 0 \end{aligned}$$

Solving for  $P$  and  $Q$ , we have

$$\begin{aligned} P &= \alpha A \oplus \alpha^2 B \oplus \alpha^5 C \oplus \alpha^3 D \\ Q &= \alpha^3 A \oplus \alpha^6 B \oplus \alpha^4 C \oplus \alpha D \end{aligned}$$

<b>Recorded data</b>			
<b>Information words</b>		<b>Parity words</b>	
A	001	P	101
B	101	Q	110
C	011		
D	100		

Table 3.2    *Sample values and calculated parity words.*

where we assume that

$$\begin{aligned}
 \alpha &= 010 \\
 \alpha^2 &= 100 \\
 \alpha^3 &= 011 \\
 \alpha^4 &= 110 \\
 \alpha^5 &= 111 \\
 \alpha^6 &= 101 \\
 1 = \alpha^7 &= 001 \\
 0 &= 000
 \end{aligned}$$

Table 3.2 shows some sample values for  $A$ ,  $B$ ,  $C$ ,  $D$ ,  $P$ , and  $Q$ . To detect errors in the received data, two syndromes, or error patterns, are calculated by these decoding equations:

$$\begin{aligned}
 S_1 &= A' \oplus B' \oplus C' \oplus D' \oplus P' \oplus Q' \\
 &= E_A \oplus E_B \oplus E_C \oplus E_D \oplus E_P \oplus E_Q \\
 S_2 &= \alpha^6 A' \oplus \alpha^5 B' \oplus \alpha^4 C' \oplus \alpha^3 D' \oplus \alpha^2 P' \oplus \alpha Q' \\
 &= \alpha^6 E_A \oplus \alpha^5 E_B \oplus \alpha^4 E_C \oplus \alpha^3 E_D \oplus \alpha^2 E_P \oplus \alpha E_Q
 \end{aligned}$$

where  $A'$  through  $Q'$  are the reproduced words,  $E_A$  through  $E_Q$  are the errors of each word, and the alphas (along with 1 and 0) are the weighting constants. If there is no error, then  $S_1 = S_2 = 0$ . If the word  $A$  is erroneous,  $S_1 = E_A$  and  $S_2 = \alpha^6 S_1$ . In other words, an error results in non-zero syndromes, and the value of the erroneous words can be determined by the difference of the weighting between  $S_1$  and  $S_2$ . In practice, the polynomials used in the CIRC correction code are:

$$\begin{aligned}
 P &= \alpha^6 A \oplus \alpha B \oplus \alpha^2 C \oplus \alpha^5 D \oplus \alpha^3 E \\
 Q &= \alpha^2 A \oplus \alpha^3 B \oplus \alpha^6 C \oplus \alpha^4 D \oplus \alpha E
 \end{aligned}$$

and the resulting syndromes are:

$$\begin{aligned} S_1 &= A' \oplus B' \oplus C' \oplus D' \oplus E' \oplus P' \oplus Q' \\ S_2 &= \alpha^7 A' \oplus \alpha^6 B' \oplus \alpha^5 C' \oplus \alpha^4 D' \oplus \alpha^3 E' \oplus \alpha^2 P' \oplus \alpha Q'. \end{aligned}$$

As noted previously, CIRC employs cross interleaving, the separation of two error correction codes by an interleaving stage. Thus one code can check the accuracy of the other code. Another important aspect of cross interleaving is that error correction is enhanced, but the amount of redundancy is not increased. An example of cross interleaving and the way it facilitates parity checking is shown in figure 3.13. S represents data values (called symbols), P and Q are parity values, and C1 and C2 are correction decoders capable of correcting one and two symbols respectively. If S13 is erroneous, C1Q16 corrects it. If S13 and S14 are erroneous, C1Q16 and C1Q20 correct them. If S13 and S16 are erroneous, C2P12 and C2P16 correct them. If S13, S14, S16, and S17 are erroneous, C2P12 and C2P16 correct them by using C1's error detection pointers. If there are more errors in the P12 and P16 strings, correction becomes impossible. CIRC uses cross interleaving similarly; however, the C1 and C2 decoders are more powerful.

### **CIRC Encoding**

The complete CIRC encoding scheme is shown in figure 3.14. With this encoding algorithm, bits from the audio signal are cross-interleaved, and two encoding stages generate parity symbols, or data values. Error correction encoding begins with the first stage of interleaving, which is designed to assist interpolation. Twenty-four 8-bit symbols are applied to the CIRC encoder. A delay of two symbols is placed between even and odd samples: That is, even samples are delayed by two blocks. Interpolation can be accomplished in the case where two uncorrectable blocks occur. Next, the symbols are scrambled so as to separate even- and odd-numbered data words. This process facilitates concealment.

The C2 encoder accepts the 24-byte parallel word and produces 4 bytes of Q parity. Q parity is designed to correct one erroneous symbol or up to four erasures in one word. An erasure is a word that has been erased by the decoders because detection has ascertained its value is unreliable. The parity symbols are placed in the center of the block to increase the odd/even distance. This enhances interpolation in the case of burst errors.

Cross interleaving follows the C2 encoder. The 28 bytes are delayed by differing periods, which are integer multiples of 4 blocks. As a result of this convolutional interleave, each C2 word is stored in 28 different blocks,

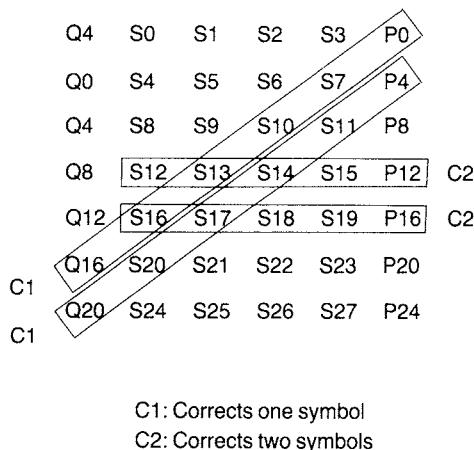


Figure 3.13 Cross interleaving permits more efficient correction of errors by decoders.

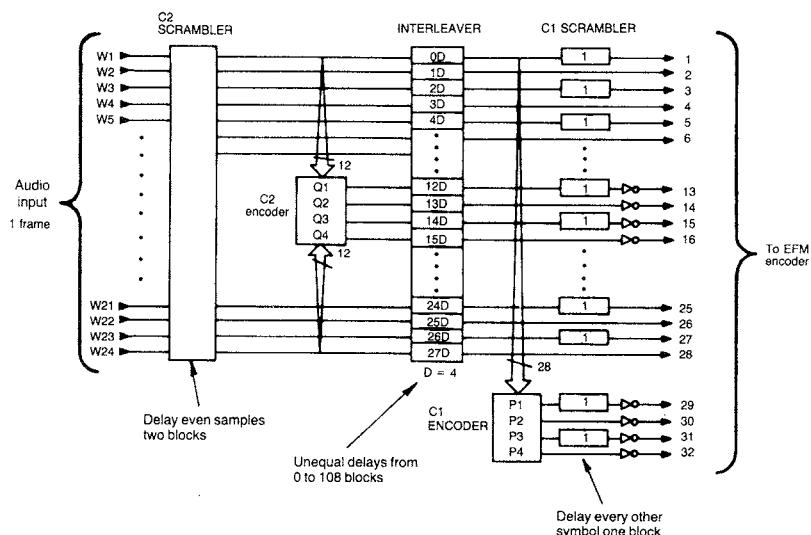


Figure 3.14 CIRC encoding.

distributed over 109 blocks. In mathematical terms, we have crossed a data array in two directions.

The C1 encoder accepts a 28-byte word (from 28 different C2 words) and produces 4 more bytes of P parity. The C1 encoder is used to correct

single symbol errors and to detect and flag double and triple errors for Q correction.

A last interleave stage introduces a fixed delay of one symbol to alternate symbols. This odd/even delay spreads the output words over two data blocks. This prevents random errors from disrupting more than one symbol in one word, even if two adjacent symbols in one block are erroneous. Finally, the P and Q parity symbols are inverted to provide non-zero P and Q symbols with zero data. This assists data readout during areas with muted audio program. Thirty-two 8-bit symbols leave the CIRC encoder.

### ***CIRC Decoding***

Upon playback, following demodulation, data is sent to a CIRC decoder for de-interleaving, error detection, and correction. The CIRC decoding process reverses many of the processing steps accomplished during encoding; the complete CIRC decoding process is shown in figure 3.15. It utilizes parity from two Reed-Solomon decoders and de-interleaving. The first decoder, C1, is designed to correct random errors and to detect burst errors. It puts a flag on all burst errors to alert the second decoder, C2. Given this prior knowledge, and help from de-interleaving, C2 is able to correct burst errors as well as random errors that C1 was unable to correct.

During reproduction, the CIRC decoder accepts one frame of thirty-two 8-bit symbols: twenty-four are audio symbols and eight are parity symbols. Odd-numbered symbols are delayed, and parity symbols are inverted. The delay lines have a delay equal to the duration of one symbol. Thus information of even-numbered symbols of a frame is de-cross-interleaved with that of the odd-numbered symbols on the next frame. This de-interleaving places the even-and odd-numbered audio symbols back into their original order by rearranging their order as read from the disc. Any sequence of errors on the disc is thus distributed among valid data.

In the C1 decoder, errors are detected and corrected by the four P parity symbols; short duration random errors are corrected, and longer burst errors are passed along. Specifically, the C1 decoder can correct a symbol error in every word of thirty-two symbols. If there is more than one erroneous symbol, then all twenty-eight data symbols are marked with an erasure flag and passed on. Valid symbols, those adhering to the encoding rules of C1, are passed along unprocessed.

The delays between the decoders are of unequal length and longer than those at the input to the C1 decoder; this convolutional interleaving enables the C2 decoder to correct longer burst errors. Because the word

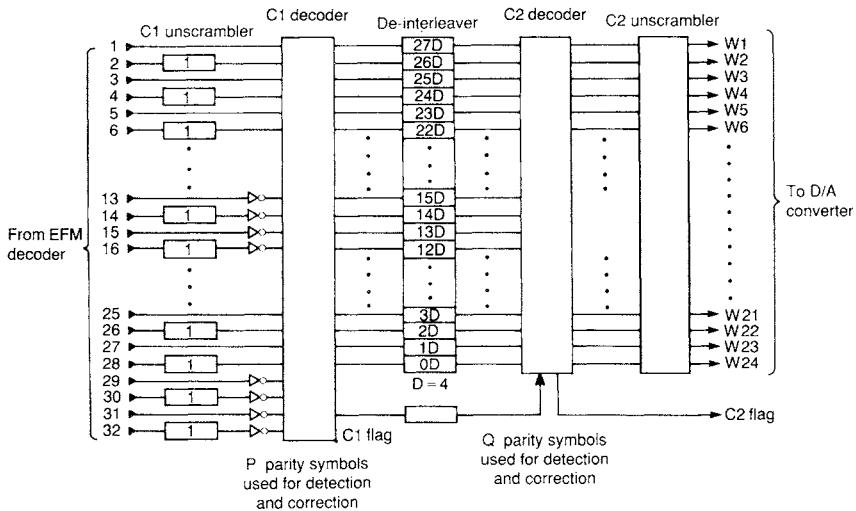


Figure 3.15     *CIRC decoding.*

arriving at C2 contains symbols from C1 decoded at different times, symbols marked with an erasure flag are now distributed among valid symbols. This helps the C2 decoder correct burst errors. Symbols without a flag are assumed error free and passed through unprocessed. In the C2 decoder errors are corrected by the four Q parity symbols. C2 can detect and correct single symbol errors, correcting up to four symbols, if flagged (in some implementations, only two symbols can be corrected). Burst errors and errors which might have occurred in the encoding process itself (rather than in the medium) are corrected. In addition, C2 can correct symbols miscorrected by C1 decoding. When C2 cannot accomplish correction because more than four symbols are flagged, twenty-four data symbols are flagged as uncorrected and passed on for interpolation. Final de-scrambling and delay is performed to assist interpolation.

The use of two correction decoders and cross interleaving helps tackle a particularly difficult error scenario. Interleaving distributes burst errors, perhaps caused by surface contamination, over different words for easier correction. Nevertheless, correction is difficult when a burst error coincides with a random error introduced perhaps by a manufacturing defect. In the CD standard, random errors are defined to be single symbol errors, or no more than  $17T$  (time periods) in length. Any longer errors are burst errors. EFM coding (explained later in this chapter) guarantees that a

random error will never corrupt more than two symbols, and the even/odd interleave guarantees that a two-symbol random error will always appear as a single error in two different C1 words after de-interleaving. Random errors are thus always correctable, and the C2 decoder retains its burst error correction capability.

### ***CIRC Performance***

It should be noted that CD players are not created equal in terms of error correction. Any CD player's error correction ability is limited to the success of the strategy chosen to decode the CIRC code on the disc. For example, for more efficient correction, decoding could be repeated alternatively, in the form C1, C2, C1, C2.

A measure of the overall correction performance is the number of interpolated sample values for a given bit error rate (BER) in a given time. The lower the interpolation rate for a given BER value is, the better the system's random error correction quality will be. Alternatively, BLER (block error rate) measures blocks containing errors. An assessment of system performance must also account for uncorrected errors that pass through the error correction circuitry. These errors may result in an audible click on the output. Table 3.3 summarizes the specifications of the CIRC system.

In theory, CIRC detects and completely corrects burst errors of up to 4000 bad bits, or a period of 1.9 milliseconds, or a physical defect 2.47 millimeters long. Interpolation can conceal errors of up to 13,700 bad bits, corresponding to a physical defect of 8.5 millimeters. Looked at in another way, the raw error rate (before correction) from a compact disc may range from  $10^{-5}$  to  $10^{-6}$ , or one error in every  $10^5$  to  $10^6$  bits. After CIRC, in theory, the error rate is between  $10^{-10}$  and  $10^{-11}$ , or about one uncorrected error in every 10 or 100 billion bits.

However, the quality of the error correcting performance varies from player to player. Depending on the CIRC decoding strategy used, the theoretical probability of an error occurring can vary from quite often to almost never. Three different CIRC decoding strategies are shown in figure 3.16, and their random error correction performance is shown in figure 3.17. Although the simplest strategy, number 1 (in figure 3.16), is satisfactory for most applications, strategy 3 is clearly much better. In the example, other strategies were developed by interchanging the C1 and C2 decoders of strategies 2 and 3. Strategy (2,3) uses the C1 decoder from strategy 2 and the C2 decoder from strategy 3. This illustrates the effect of different error conditions on the success of the correction strategy. For example, strategy 3 is superior to strategy 2 primarily because of its ability

Aspect	Specification
Maximum completely correctable burst length	$\approx 4000$ data bits (i.e. $\approx 2.47$ mm track length on the disc)
Maximum interpolatable burst length in the worst case	$\approx 13,700$ data bits (i.e. $\approx 8.5$ mm track length)
Sample interpolation rate	One sample every 10 hours at $BER = 10^{-4}$ ; 1000 samples per minute at $BER = 10^{-3}$
Undetected error samples (clicks)	Less than one every 750 hours at $BER = 10^{-3}$ ; negligible at $BER = 10^{-4}$

Table 3.3    *Performance specifications of CIRC.*

to correct two independent errors at the C1 decoder. Thus for correcting burst errors, strategy 3's superiority could be drastically reduced.

In short, it is difficult for designers to evaluate error correction performance prior to real use. Figure 3.18 shows an example of error correction performance on three discs. The clean disc represents normal conditions (with a block error rate of  $3.3 \times 10^{-4}$ ). On the disc with the fingerprint, the smudge covers the entire disc. The block error rate is  $5.6 \times 10^{-4}$ . To create the scratches, the disc was rubbed on a wooden table for approximately one minute. This created a block error rate of  $4.5 \times 10^{-3}$ . In practice, for CDs that are in reasonably good shape, you might find one uncorrected error per disc, to be dealt with by error concealment methods.

## DATA ENCODING

Compact disc encoding is the process of placing audio data in a format suitable for storage on the disc. A frame structure provides such a format. The frame furnishes a means to distinguish the data types: audio data and its parity, synchronization word, and subcode. The information contained in a CD frame (prior to modulation) contains a 27-bit synchronization word, 8-bit subcode, 192 data bits, and 64 parity bits. CD frames are assembled when the master disc is encoded. Assembly of the frame involves several processing steps as well as modulation and the addition of merging bits. The complete encoding sequence is shown in figure 3.19.

Encoding begins with the audio data. Six 32-bit PCM audio sampling periods (alternating from 16-bit left and right channels) are grouped in a frame, the left channel preceding the right. Each 32-bit sampling period is divided to yield four 8-bit audio symbols. (The original 16-bit number

## 1) Strategy 1 (Simple Correction Strategy)

C<sub>1</sub> Decoder*if* single- or zero-error syndrome*then* modify at most one symbol

accordingly

*else* assign erasure flags to all symbols  
of the received wordC<sub>2</sub> Decoder*if* single- or zero-error syndrome*then* modify at most one symbol

accordingly

*else if*  $f > 1$ *then* copy C<sub>2</sub> erasure flags  
from C<sub>1</sub> erasure flags*else if*  $f \leq 1$ *then* assign erasure flags to  
all symbols of the  
the received word

## 2) Strategy 2 (Middle Correction Strategy)

C<sub>1</sub> Decoder*if* single- or zero-error syndrome*then* modify at most one symbol

accordingly

*else* assign erasure flags to all symbols  
of the received wordC<sub>2</sub> Decoder*if* single- or zero-error syndrome*then* modify at most one symbol

accordingly

*else if*  $f > 2$ *then* copy C<sub>2</sub> erasure flags  
from C<sub>1</sub> erasure flags  
*else if*  $f = 2$   
*then* try two-erasure decoding  
*else if*  $f < 2$  or *if* two-erasure  
decoding fails*then* assign erasure flags  
to all symbols of the  
received word

## 3) Strategy 3 (Superstrategy)

C<sub>1</sub> Decoder*if* single- or zero-error syndrome*then* modify at most one symbol

accordingly

*else if* double-error syndrome*then* modify two symbols  
accordingly assign erasure  
flags to all symbols of the  
received word*else* assign erasure flags to all  
symbols of the received wordC<sub>2</sub> Decoder*if* single- or zero-error syndrome*then* modify at most one symbol

accordingly

*else if*  $f \leq 4$ *then* *if* double-error syndrome  
and  $L_c = 2$   
*then* modify two symbols  
accordingly*else if* {double-error syndrome  
and  $[(L_c = 1 \text{ and } f \leq 3)$   
or  $(L_c = 0 \text{ and } f \leq 2)]\}$   
or  $(f \leq 2 \text{ and not double-}$   
error syndrome)}*then* assign erasure flags  
to all symbols of the  
received word*else* copy C<sub>2</sub> erasure flags  
from C<sub>1</sub> erasure flags*else* copy C<sub>2</sub> erasure flags from  
C<sub>1</sub> erasure flags

Figure 3.16 CD players employ different CIRC decoding strategies. Three strategies are shown here. The lowercase  $f$  is the number of erasure flags input into the C<sub>2</sub> decoder.  $L_c$  is the number of erasure flags agreeing with the error locations calculated from the syndrome. (From Vries and Odaka)

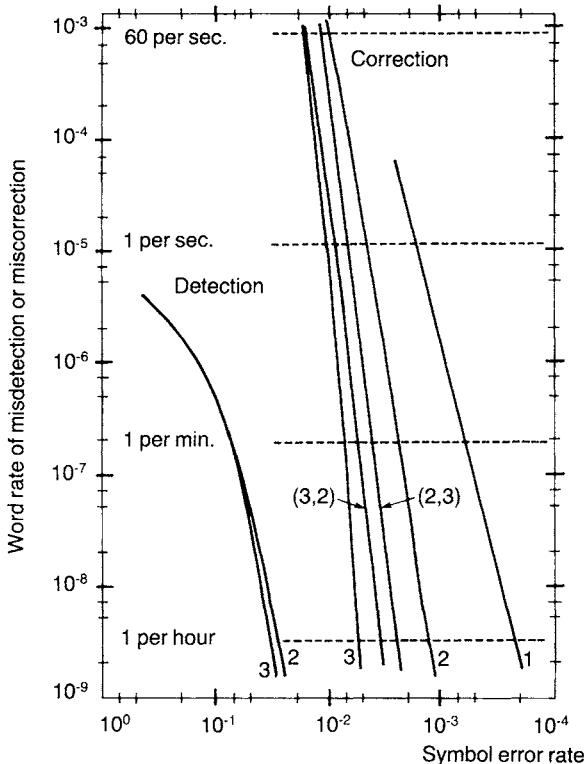


Figure 3.17    *Performance curves for three decoding strategies for random errors.* (From Vries and Odaka)

is called a word, and it is split into two 8-bit symbols.) The audio program may consist of stereo or dual monaural left and right channels. Although some recordings have used independent left and right monaural channels, the CD standard does not support this.

After grouping audio data into symbols, error correction encoding takes place. The CD system, of course, employs the Cross Interleave Reed-Solomon Code (CIRC). As we've seen, it uses a combination of interleaving and parity to make the data more robust against errors encountered during storage.

### ***Subcode***

Following CIRC encoding, an 8-bit subcode symbol is added to each frame. The eight subcode bits are designated as P,Q,R,S,T,U,V, and W, some-

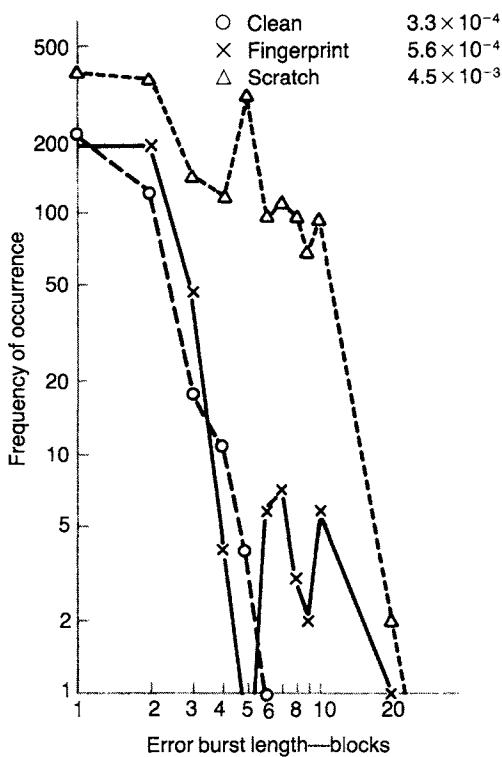


Figure 3.18 Examples of error correction performance for damaged discs. (From Doi)

times referred to as the PQ code. The CD player collects subcode symbols from ninety-eight consecutive frames to form a subcode block, with eight 98-bit words. Only the P and Q bits are used in audio CDs. Included is information specifying the total number of selections on the disc, their beginning and ending points and timings, the index points within a selection, the program lead-in and lead-out points, and updated information on the pickup's position as the disc is played. The other 6 bits (R,S,T,U,V, and W) are available for encoding other information on audio CDs. Subcode is discussed in detail later in this chapter.

#### EFM Modulation

After the audio, parity, and subcode data are assembled, the bit stream is modulated using eight-to-fourteen modulation (EFM). Blocks of 8 data

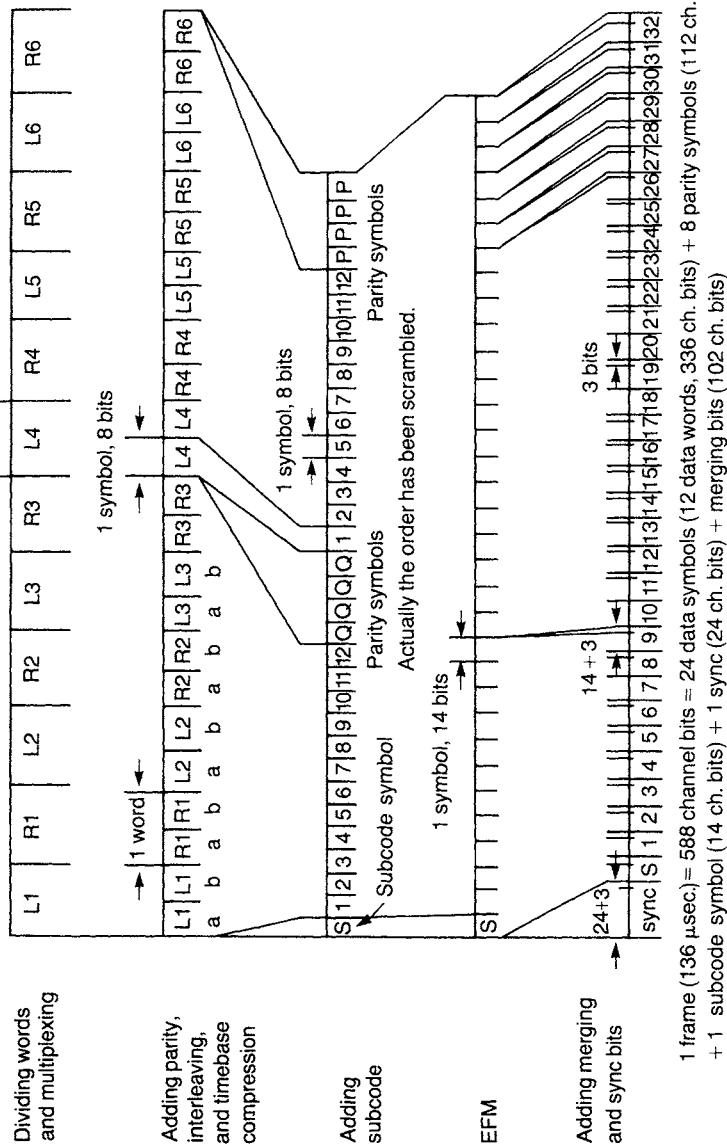


Figure 3.19 Complete CD encoding process.

<b>Data bits</b>	<b>Channel bits</b>
00000000	01001000100000
00000001	10000100000000
00000010	10010000100000
00000011	10001000100000
00000100	01000100000000
00000101	00000100010000
00000110	00010000100000
00000111	00100100000000
00001000	01001001000000
00001001	10000001000000
00001010	10010001000000

---

Table 3.4     *In EFM, each 8-bit word is translated to a 14-bit word selected for its specific bit pattern.*

bits are translated into blocks of 14 bits, known as channel bits, using a ROM dictionary which assigns an arbitrary and unambiguous word of 14 bits to each 8-bit word; a part of the lookup table is shown in table 3.4.

In EFM, each 8-bit word is translated to a 14-bit word selected for its specific bit pattern.

By choosing the right 14-bit words, the bit patterns with a low number (and known rate) of transitions between 0 and 1, greater data density can be achieved. It would be inefficient to store the 8-bit symbols directly on the disc; the large number of 0/1 transitions would demand many pits. The 8 data bits require  $2^8$  or 256 different code patterns. However, the 14-bit channel word can offer 16,384 combinations. In addition, 8-bit symbols have many similar patterns; for example, if one bit were wrong, 00001011 would be confused with 10001011. With 14-bit words, more unique patterns can be selected. EFM thus provides a kind of error correction.

To achieve pits of controlled length, only those combinations are selected in which more than two but less than ten 0s appear continuously. Only 267 combinations satisfy these criteria. With these patterns, the minimum inversion width of the signal is limited by the recording density, and the maximum inversion width is limited by the clock bit extraction. The conversion is thus a compromise between several conditions. Because only 256 patterns are needed, 11 of the 267 patterns are discarded (two of them are used for subcode synchronization words).

### **Merging Bits**

Blocks of 14 bits are linked by 3 merging bits. Two merging bits (always 0s) are required to prevent the possibility of successive 1s between serial words (a violation of the coding scheme). The additional merging bit (either a 1 or a 0, depending on the preceding and succeeding patterns) is added to each code pattern to aid in clock synchronization and to suppress the signal's low-frequency component.

The signal's low-frequency component is suppressed by selecting merging bits which maintain the signal's average digital sum value at zero, as shown in figure 3.20. The digital sum value is determined by assigning +1 and -1 to the high-level (1) and low-level (0) amplitudes respectively. The merging bits may take either state; the processor selects a 1 or 0 to reduce the overall digital sum value. In figure 3.20a two consecutive EFM patterns result in a digital sum value offset. With a transition introduced into the merging bits, the digital sum value is reduced as shown in figure 3.20b. Note that the data content itself is not affected. The ratio of high to low bits before and after modulation is 8:17. During demodulation, only the 14 data bits will be processed; the 3 merging bits are discarded.

### **Pit Length**

At this stage of the encoding, the signal exists as a non-return-to-zero (NRZ) signal. This is a form commonly employed for processing digital data in which the signal level is high at 1 and low at 0. After EFM modulation, the signal is converted to non-return-to-zero-inverted (NRZI) form. In this form, the signal inverts at the center of each 1. NRZ and NRZI are reviewed in figure 3.21. NRZI gives the waveform fewer transitions and thus simplifies the pit structure on the disc.

The resultant channel stream produces pits and lands which are at least 3 channel bits and no more than 11 channel bits long. The 3T pit is thus the highest frequency signal (720 kHz), while the 11T pit is the lowest (196 kHz). T signifies one period. Pit lengths thus vary from 0.833 to 3.054 micrometers at 1.2 meters per second, and from 0.972 to 3.56 micrometers at 1.4 meters per second. It is the combination of these varying dimensions which physically encode the data. The selection of EFM bit patterns defines the physical relationship of the pit dimensions. The pit and intervening reflective land on the CD surface do not directly designate 1s and 0s. Rather, as shown in figure 3.22, each pit edge, whether leading or trailing, is a 1, and all increments in between, whether inside or outside a pit, are 0s. This is a much more efficient storage technique than coding the binary bits directly with pits.

With EFM there are more bits to accommodate, but with modulation the highest frequency in the output signal is decreased; therefore, a lower

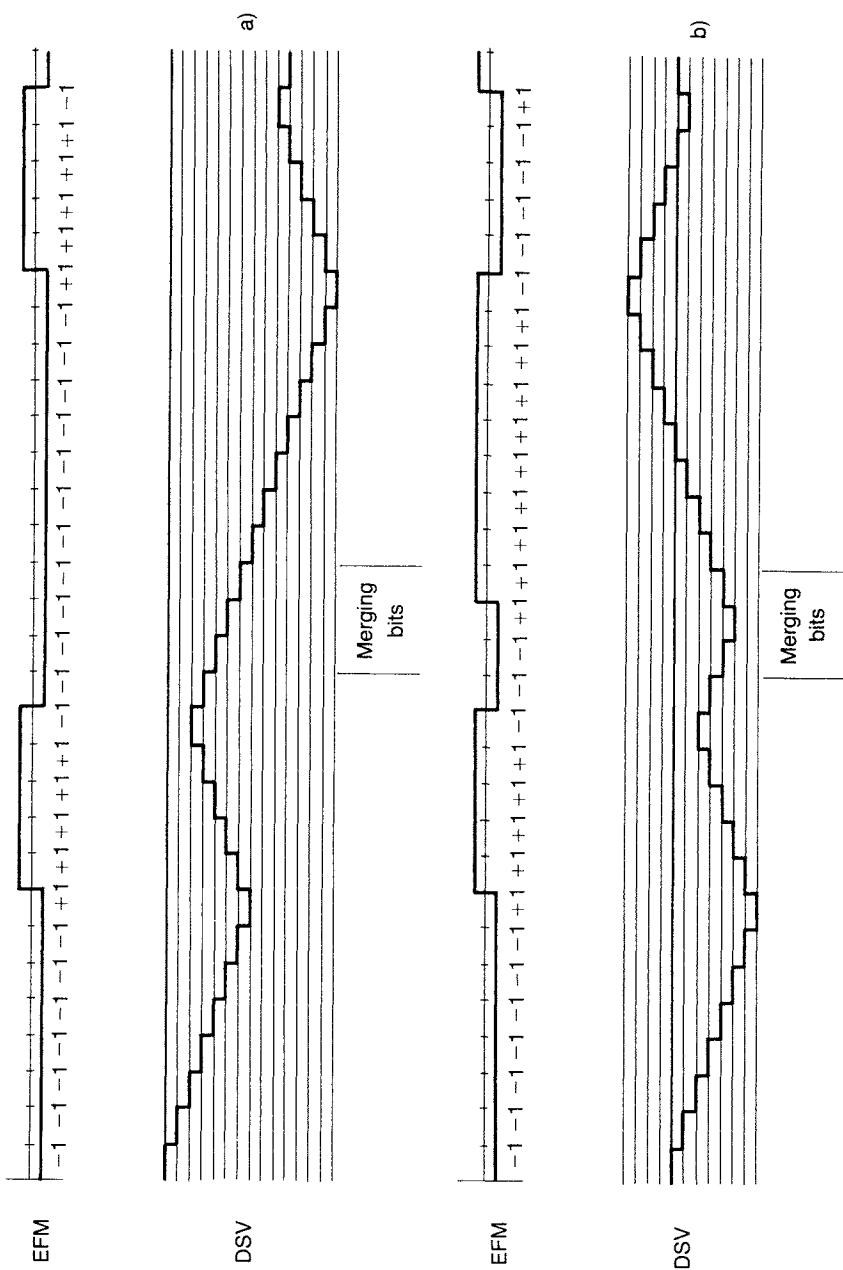


Figure 3.20 If merging bits were not used (a), the DSV would accumulate an offset. With merging bits (b), the DC component is controlled.

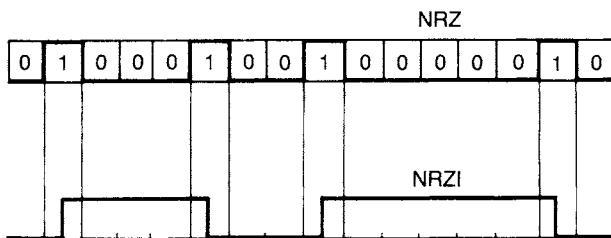


Figure 3.21 After modulation, the signal is converted from NRZ to NRZI form.

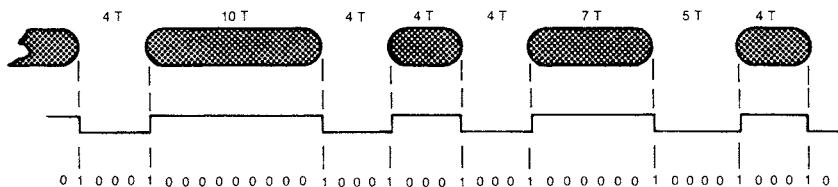


Figure 3.22 Pit edges encode logical 1 channel bits.

track velocity can be utilized, which results in longer playing time. This is a very efficient encoding method because the number of bits transmitted, divided by the number of transitions needed on the medium to convey them, is high.

## *Frame Assembly*

The individual frames in the resulting EFM bit stream must be delineated. Thus a synchronization pattern is added prior to each frame. This is necessary to make the bit stream self-clocking; that is, no modulation clock is needed to maintain an accurate readout rate. The synchronization word is uniquely distinguishable from any other possible data configuration; it is 3 transitions separated by  $11T$ . Specifically, the 24-bit synchronization word is 100000000001000000000010 plus 3 merging bits.

The total number of channel bits per frame after encoding is 588, comprising the following: 24 synchronization bits,  $336(12 \times 2 \times 14)$  data bits,  $112(4 \times 2 \times 14)$  error correction bits, 14 subcode bits, and  $102(34 \times 3)$  merging bits, as shown in figure 3.23. When the data manipulation is completed, the original audio bit rate of 1.4112 million bits per second has

1 sample:	1 word: 16 bits, 2 symbols	Sample rate:	44.1k samples/sec/channel
1 data or parity symbol:	14 bits (8 bits before EFM)	Data bit rate:	1.4M bits/sec
1 subcoding symbol:	14 bits (Note that meaningful information is on the accumulated 98 frames.)	Frame rate:	7.35k frames/sec (75 subcode-frames/sec)
1 sync signal:	24 channel bits	Control bit rate:	58.8k bits/sec
1 frame	6 samples/channel: 12 data words	Transfer rate:	2.03M bits/sec before EFM
		Bit rate: (clock frequency):	4.3218M bits/sec

Table 3.5 *Data formats and data rates.*

augmented to 4.3218 million bits per second. CD data format and rates are summarized in table 3.5.

It is important to remember that the only information in a CD signal is the position in time of the pit transitions. In other words, all data is represented as timing patterns. This means that there are more data bits recorded than there are transitions on the disc. (Specifically, there are 24 data bits for every 17 disc transitions.) That's efficiency.

On the other hand, this efficiency is tempered by a certain amount of overhead. In fact, audio data is in a minority position on a compact disc. For example, a CD holding 74 minutes, 33 seconds of music contains 6.3 billion audio bits. However, after frame assembly and EFM coding, there are 19.3 billion bits to be stored. Thus only 32.65 percent of the disc holds audio bits. The rest is given to overhead: 33.67 percent EFM, 17.35 percent merging bits, 10.88 percent CIRC, 4.08 percent synchronization, and 1.36 percent subcode. Fortunately, the vast capacity of the CD excuses the overhead; for example, even that tiny 1.36 percent of subcode represents over 32 megabytes of data.

### ***Emphasis***

The CD system includes an emphasis characteristic standard. Using this option, a master tape may be pre-emphasized, that is, audio high frequencies are boosted, so that all replicated discs will likewise carry this equalization. The pre-emphasis is flagged in the CD's subcode. Every CD player is equipped with a de-emphasis circuit in its output stage. Upon playback, the CD player will detect the pre-emphasis flag in the subcode and automatically switch in its de-emphasis circuit to produce the inverse equalization. By attenuating high frequencies without affecting overall

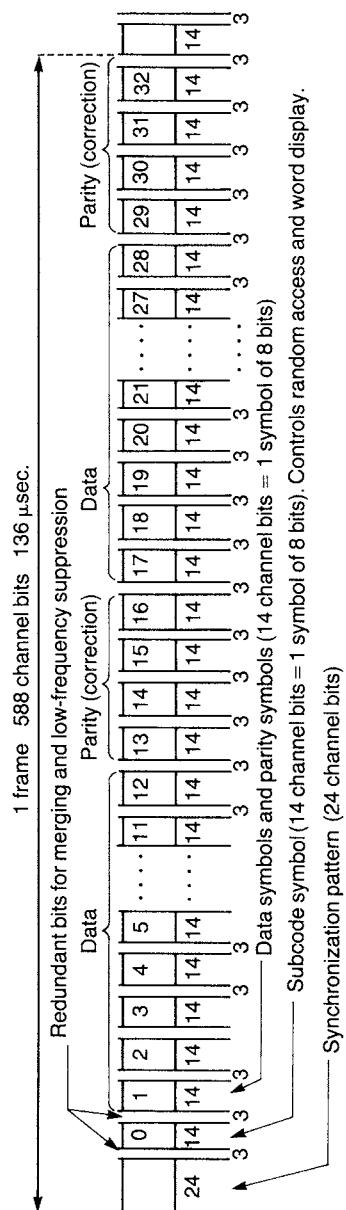


Figure 3.23 Summary of frame format after EFM modulation.

signal frequency response, the signal-to-noise ratio of the signal may be improved. The pre-emphasis and de-emphasis curves are specified with time constants of 50 and 15 microseconds, corresponding to frequency points on the equalization curve at 3,183 and 10,610 Hz, with a 6 dB/octave slope between these points. The standard pre-emphasis and de-emphasis curves are shown in figure 3.24.

## DATA DECODING

The pits on the CD data surface, as we have seen, are the physical manifestation of complicated data encoding, including multiplexing, interleaving, parity, error protection, and EFM, all of which take place at the lathe when the master disc is cut. The decoding of these bits requires complex signal processing. The player must demodulate the data, de-interleave it, and correct data errors, while observing the data format originally encoded on the master disc. The complexity of the task accounts for the design differences between players, but the overall scheme remains similar.

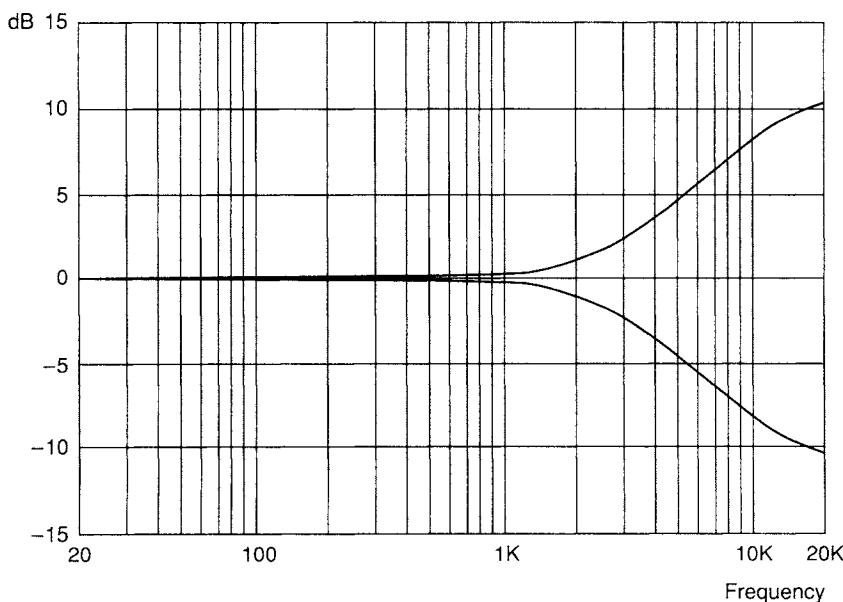


Figure 3.24 The CD standard pre-emphasis (top curve) and de-emphasis (bottom curve) characteristics. (Courtesy Robert Dunn)

### ***EFM Demodulation***

When the laser beam is reflected at the disc surface during playback, the varying intensity of the returned beam is detected by a photodiode sensor. It is the voltage from the sensor which is ultimately transformed into the analog audio signal output from the player. However, the encoded data from the pickup must first be decoded. The signal encoded on the disc utilized eight-to-fourteen modulation (EFM), which specified that the signal be composed of not less than two 0s or more than ten 0s between logical transitions (pit edges). This results in pit lengths expressed in a variety of combinational patterns from a pit length 3 units long to a pit length 11 units long; this sets the frequency limits of the EFM signal read from the disc. As we've seen, this range is sometimes referred to as a  $3T$ - $11T$  signal, with  $T$  referring to the period of one bit. The audio output is ultimately derived from the lengths (read as timing intervals) of the data pits and lands.

The photodiode and its processing circuits produce a signal resembling a high-frequency sine wave called the EFM signal. Because of its high frequency, it is sometimes referred to as an RF (radio frequency) signal. Figure 3.25 shows the RF signal from a series of pits. A collection of data periods is often called the eye pattern. (The diamond-shaped area in the middle of the pattern, along the axis, looks like an eye, if you've got a lively imagination.) An eye pattern is shown in figure 3.26.

The varying periods of these sine waves ( $3T$ ,  $4T$ , etc.) reflect the periods of time required to read the disc pits. The eye pattern is thus an electrical equivalent of the disc's data. The digital data can always be recovered from the eye pattern if it can be determined when the signal crosses the zero axis relative to the  $3T$ - $11T$  time patterns. Whenever a player is tracking data, the eye pattern is present, and the quality of the signal may be observed from the pattern. For example, a skewed disc would result in a distorted eye pattern.

The zero-crossings of the eye pattern contain the sole information content of the EFM signal. Because pit walls are not vertical, the resulting signal is not square. In addition, pits can be longer or shorter than the actual pulse width of the recorded EFM signal, depending on cutting con-

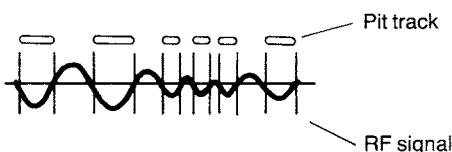


Figure 3.25 Data pits are used to produce a representative RF signal.

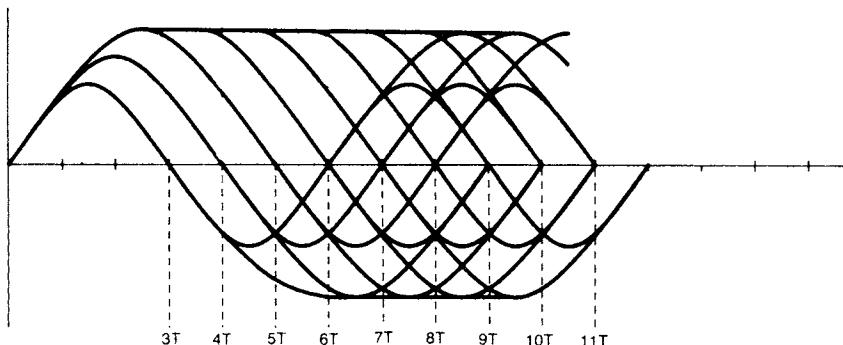


Figure 3.26 *A collection of RF signals forms the eye pattern.*

ditions. This causes the RF signal to be asymmetrical to the DC level. Figure 3.27 shows a threshold control circuit which adjusts the threshold level, eliminating DC offset in order to achieve correct pulse width. The resulting binary waveform now has a more normal digital appearance; the varying widths of the waveforms still contain the EFM information.

Because the eye pattern at the photodiodes has a very small amplitude, it must be boosted by a preamplifier. An equalization network or transversal filter must be used to ensure that the smallest signal duration (high frequency)  $3T$  is equal in amplitude with the longest duration (low frequency)  $11T$ . The filter must compensate for the fact that shorter period signals ( $3T$ ,  $4T$ , etc.) are lower in amplitude than the others. The overall signal is processed through an automatic gain control circuit; no matter what the reflectivity of the CD (within certain tolerances), this circuit optimally adjusts the level of the eye pattern.

Although the EFM signal is composed of a varying sinusoidal waveform, it is truly digital; it undergoes processing to convert it into a series of square waves more easily accepted by the digital circuits to follow. This does not affect the encoded data, since it is the width of the EFM periods which holds the information of interest. Following the conversion of the eye pattern to square waves, the NRZI (non-return to zero inverted) signal is returned to NRZ (non-return to zero), the representation in which a high level is a 1 and a low level is a 0.

The transitions are used to phase-lock a separator clock with period  $T$ . This clock identifies the number of  $T$  periods between transitions and thus distinguishes between the  $17T$  symbols (each of the thirty-three symbols of a data frame is  $17T$  long). Thanks to our pains in the encoding process, the data off the disc is thus self-clocking. The pulse generator has a slope-

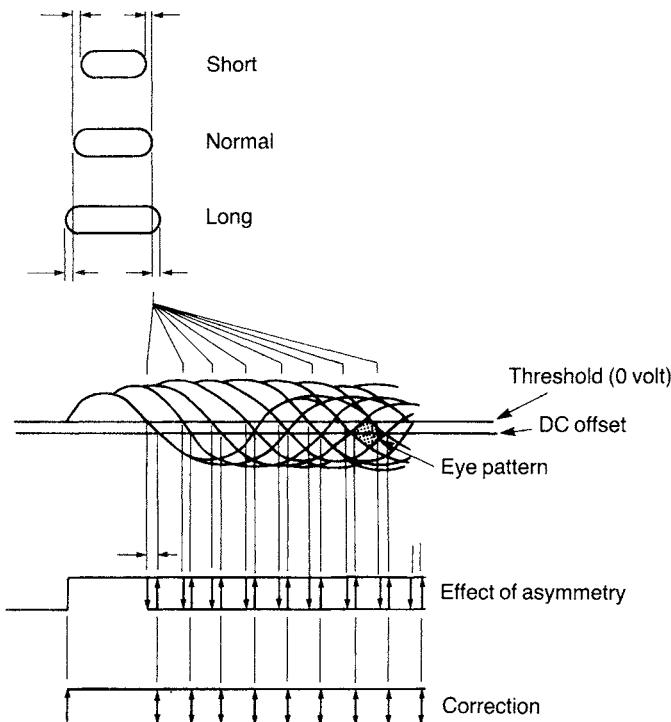


Figure 3.27     *A threshold control circuit eliminates DC offset, due to asymmetry in pit length, in the RF signal.*

sensing window of one period ( $1T$ ) for correctly reading a signal with any jitter or timebase error.

The first piece of data to be extracted from the NRZ signal is the synchronization word, three transitions separated by  $11T$  periods. The start of each data frame can thus be identified. This synchronization information is used to synchronize the thirty-three symbols of channel information in each frame, and a synchronization pulse is generated to aid in locating the zero crossing of the EFM pattern and to generate a transition at these points to produce a binary signal. The  $3T$  merging bits have served their purpose of preventing EFM code violations and suppressing DC. They are separated from the  $17T$  symbols, leaving  $14T$  symbols.

Because the circuit which controls the disc's precise rotational speed has a limited range, disc speed must be controlled approximately. This is often accomplished by examining the durations of transitions in the EFM signal. The  $3T$ - $11T$  range means that transitions cannot occur faster than

approximately 0.66 microseconds or slower than 2.44 microseconds. When a CD player first starts to play a disc, or after it jumps to a different track radius (with a different track speed), a correction signal is generated to bring the disc rotation to within the capture range of the timebase corrector (explained below) for precise regulation. This is one reason why C1 and C2 parity bytes are inverted in the CIRC encoder; 1s are placed in the bit stream when the data is muted with all 0s. Thus minimum and maximum pit lengths are recorded during the disc lead-in where the player must quickly regulate disc speed.

The EFM signal is now demodulated so that every 14-bit EFM word again becomes 8 bits. Depending on player design, demodulation is effected either by logic circuitry or a lookup table, which is a list stored in memory that uses the recorded data to refer to the original patterns of 8 bits.

The data stream now contains audio samples, redundancy, and subcode. The EFM demodulator sends subcode data to the display circuit to convey timing and other information to the player front panel. Other control signals contained in the subcode are used, among other things, to alert the player's control microcomputer to the exact location of the pickup.

### ***Timebase Correction***

A buffer is used to remove the effect of disc rotational irregularities; it is sometimes called a timebase corrector. A RAM (random access memory) is commonly used. Data input to the buffer may be irregular, but the buffer output is controlled by a crystal clock, to minimize jitter.

Timebase correction can be performed with about 16 kbytes of storage, and de-interleaving may be accomplished with the same memory. Timebase correction clocking ensures that the data is output at a constant rate despite any jitter in the data stream entering the signal processing circuits. Such variations are introduced by the disc's rotating speed. Timebase correction clocking thus eliminates two of the inherent problems of analog systems—speed variations called wow and flutter. In addition to preventing mechanical error, care must be taken to minimize jitter in the decoded data stream; phase-locked loops are often used.

Error conditions such as buffer overflow or underflow exceed the limits of the system. To prevent these, the buffer's level is constantly monitored by examining the write address. This information is supplied to the circuit which controls the speed of the disc motor and hence the bit rate entering the player. Alternatively, the data separator clock, counting samples from the disc, can be compared to a fixed reference in order to generate a rotation control voltage. Whichever method is used, it is interesting to note that the actual disc speed is unimportant; the player will rotate the disc at

any speed necessary to achieve the proper data rate. As a result, CDs may be mastered with a linear track velocity between 1.2 and 1.4 meters/second; however, the velocity will be constant on a given disc.

### ***Output Processing***

Following demodulation, interleaved audio and parity data are routed to the CIRC error correction circuit for de-interleaving, error detection, and error correction. Interpolation and muting circuits, previously discussed, follow the CIRC decoder. Using error flags from the decoder, the player's signal processing circuits determine whether to output data directly, to interpolate it, or to mute the sound. Interpolation and muting performance may also differ from player to player.

Following error correction, the left and right audio channels must be demultiplexed and their respective samples joined together in the same sequence and at the same rate in which they were recorded.

After these signal processing chores are accomplished, the data has been restored to its original state: 16-bit samples of digital audio information. In other words, the audio data has been reconstituted to the player's best ability. After a quick trip through D/A converters and low-pass analog filters, or oversampling filters which precede D/A conversion, the analog signals are output.

The compact disc bit stream is thus a little trickier than one might assume. In the interest of density and robustness, the data must undergo sophisticated processing during both encoding and decoding. All the more reason to appreciate the sound a CD player delivers. Next time you listen to Schubert's *Trout Quintet*, remember the bit stream it came from.

## **SUBCODE**

A compact disc player does some amazing things, but some of the credit should rightfully go to the disc itself. Many of the player's operational features derive their ability from special subcode control information encoded on every music disc. Although the user cannot access the subcode directly, it provides much of the information for proper disc playback as well as front panel controls and display. All players contain the circuitry required to interpret and utilize this information.

### ***Subcode Blocks***

As we have seen, a frame is the smallest recognizable section of data on a disc. Each frame contains audio data, parity, the synchronization word,

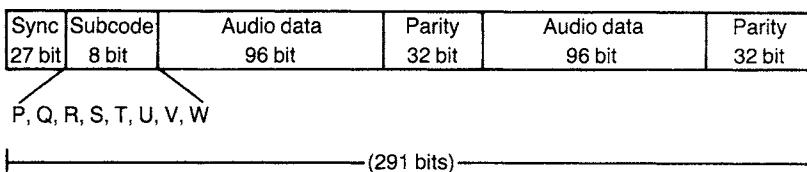


Figure 3.28 A CD frame is composed of 291 bits before EFM modulation, including 8 subcode bits per frame.

and an 8-bit field called the subcode, shown in figure 3.28. Following modulation, each subcode field is conveyed by one 14-bit EFM symbol. Upon playback the subcode data is demodulated and separated from the audio data to be decoded independently.

The 8 subcode bits are contained in every frame and are designated as P,Q,R,S,T,U,V, and W. Only the P and Q subcode bits are used in the audio format. (There is no relation to the P and Q codes in CIRC.) Since the number of bits available in each frame is small, the entire number of subcode bits available from 98 frames is collected to form a subcode block. In other words, the subcode is constructed a byte at a time from 98 successive frames. Thus the 8 subcode bits (P through W) are used as 8 different channels, with each frame containing 1 P bit, 1 Q bit, 1 R bit, and so forth. This interleaving helps minimize the effect of disc errors on subcode data.

A little mathematics determines just how fast subcode blocks accumulate. A CD contains 44,100 left and right 16-bit audio samples every second, so the byte (8-bit) rate is  $44,100 \times 4$ , or 176.4 bytes/second. There are 24 audio symbols in every frame, so the frame rate is  $176.4/24$  or 7350 Hz. Because it takes 98 frames to make one complete subcode block, the subcode block rate is  $7350/98$  or 75 Hz; that is, 75 subcode blocks appear every second. Incidentally, if you take 7350 frames per second and multiply it by the number of channel bits, 588, you wind up with 4.3218 MHz, the overall bit rate.

A subcode block is complete with its own synchronization word, instruction, data, commands, and parity. The start of each subcode block is denoted by the presence of  $S_0$  and  $S_1$  sync patterns in the first symbol positions of two successive blocks. Because  $S_0$  and  $S_1$  are two patterns (00100000000001 and 00000000010010, respectively) not used in the EFM code, they can be uniquely identified. (You might recall that there are 267 legal EFM patterns, but only 256 of them are used to encode regular data on a disc.)

### ***The P Channel***

The P channel contains a flag bit. It designates the start of a track, as well as the lead-in and lead-out areas on a disc, as shown in figure 3.29. The music data is denoted by 0 (all zeros recorded in the P bits), and the start flag as 1 (all ones recorded). The length of a start flag is a minimum of 2 seconds but equals the pause length between two tracks if this length exceeds 2 seconds. Lead-in and lead-out signals tell the player where the music program on the disc begins and ends. A lead-in signal is encoded in the P channel; it consists of all 0s appearing just prior to the beginning of the music data. At the end of the lead-in, a start flag 2 to 3 seconds long appears just prior to the start of music. During the last music track, preceding the lead-out, a start flag of two to three seconds appears. The end of that flag gives the start of lead-out, and the P channel remains at 0 for 2 to 3 seconds. Following that time, a P signal consisting of alternating 1s and 0s (at a 2 Hz rate) appears. These various sets of identity signals can be used by players of basic design to control the optical pickup. For example, a player could count start flags placed in the blank interval between tracks to locate any particular track on a disc. In practice, virtually all players use only the more sophisticated Q code. During master disc encoding, the P code is generated automatically from the Q code.

### ***The Q Channel***

The Q channel contains four basic kinds of information: control, address, Q data, and an error detection code. Figure 3.30 illustrates the Q channel data format. The control information (4 bits) handles several player functions. The number of audio channels (two or four) is indicated; this distinguishes between two- and four-channel CD recordings. Although none have yet been manufactured, a four-channel CD player would thus automatically switch between stereo or quadraphonic output. The digital copy (permit/deny) bit regulates the ability of other digital recorders to record the CD's data digitally. Pre-emphasis (on/off) is also indicated. The player reads the code and automatically switches to the appropriate de-emphasis circuit (see figure 3.24).

The address information consists of 4 bits designating the three modes for the Q data bits. Primarily, Mode 1 contains number and start times of tracks, Mode 2 contains a catalog number, and Mode 3 contains the International Standard Recording Code (ISRC). Each subcode block contains 72 bits of Q data, as described below, and 16 bits for cyclic redundancy check code (CRCC), a code used for error detection on the control, address, and Q data information.

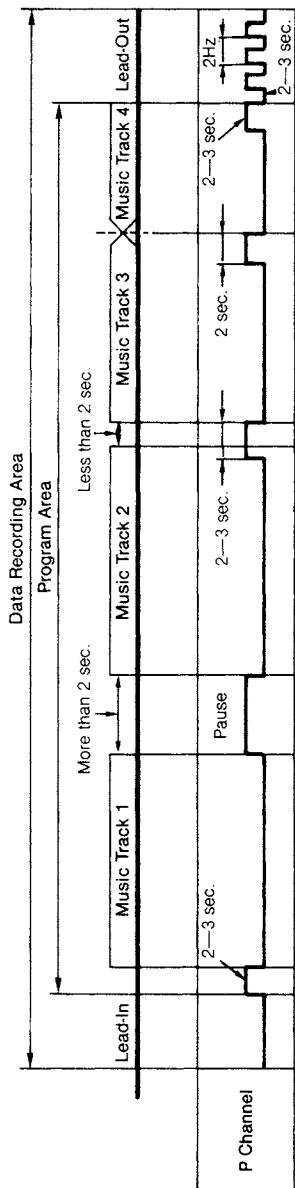


Figure 3.29 The P channel contains a flag bit designating lead-in, program, and lead-out areas on the disc. (Courtesy Sony)

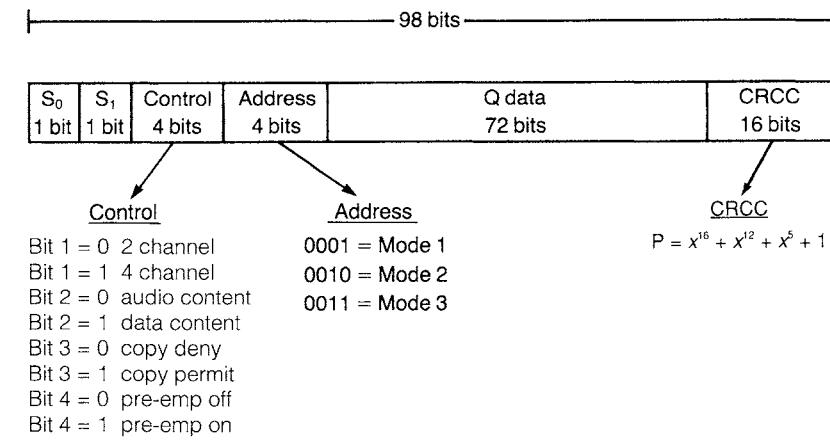


Figure 3.30    *The Q channel contains control, address, Q data (for three data modes), and CRCC.*

As noted, there are three modes of Q data. Mode 1 stores information in the disc lead-in area, program area, and lead-out area. The data content in the lead-in area differs from that in the other areas, as illustrated in figure 3.31a. Mode 1 lead-in information is used to construct the table of contents (TOC) using the 72-bit Q data, stored as 9 bytes, each containing two hexadecimal digits. The TOC stores data indicating the number of music selections (up to ninety-nine) as a track number (TNO) and the starting times (P times) of the tracks. The TOC is read during disc initialization, before the disc begins playing, so that the player can respond to any programming or program searching that is requested by the user. In addition, most players can display this information.

In the lead-in area, the TOC field is set to 00 (two binary coded decimal (BCD) digits), indicating that the data is part of the table of contents. The table of contents is assembled from the Point field; it designates a track (or control) number and points to the absolute starting time of the track in minutes, seconds, and frames (75 frames per second), stored as Pmin, Psec, and Pframe data (each holding two BCD digits). For example, table 3.6 shows a two-disc set with four tracks on each disc. Each subcode block contains one track starting time, and the starting time is repeated in three successive blocks on the disc to guard against errors (not shown in the table). The table of contents is repeated continuously through the lead-in area.

The Point number can range from 00 to 99. The Point number can also be set to A0, A1, or A2, instead of a track number. When it is set to A0,

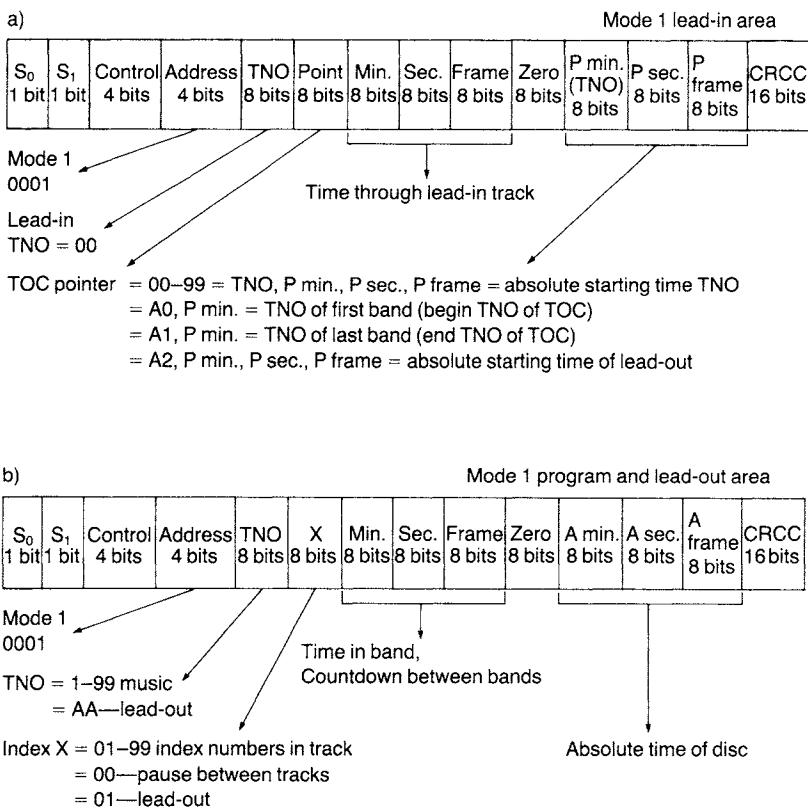


Figure 3.31 The information in the Q channel, Mode 1, varies between lead-in area, where TOC contents such as track numbers and times are stored (a), and program and lead-out areas, where track number, running and absolute time, and index information are stored (b).

the Pmin field shows the number of the first track on a disc (Psec and Pframe are set to zero). When the Point number is set to A1, the Pmin field shows the number of the last track on a disc (Psec and Pframe are set to zero). When the Point number is set to A2, the Pmin, Psec, and Pframe fields show the absolute time of the start of the lead-out track, that is, total disc time. When a disc is initialized, most players use these fields to display the total number of tracks (A1) and the total disc time (A2). In addition, these fields help accommodate multiple-disc sets. Track numbers run successively through the set using the Point fields, but the A values correspond to the tracks on each disc. In the example in table 3.6, the first Point

**Multi-Disc Table of Contents**

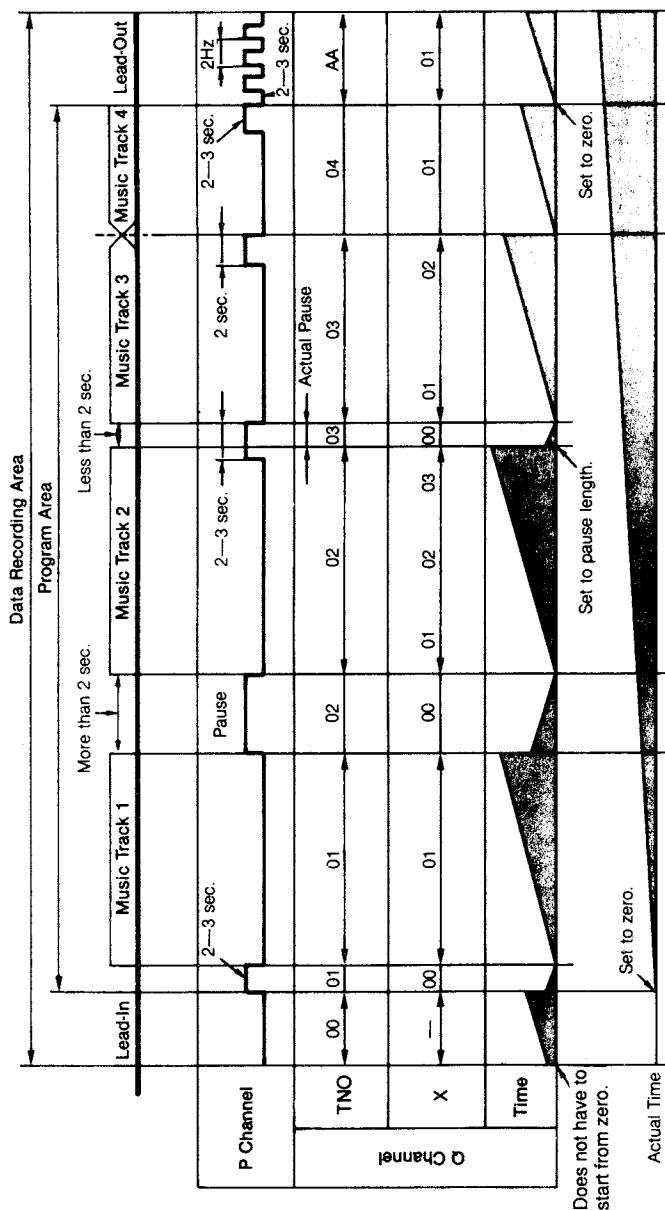
Disc One (four tracks)						Disc Two (four tracks)					
Frame no.	Point	Pmin	Psec	Pframe	Frame no.	Point	Pmin	Psec	Pframe		
0,1,2	01	00	02	13	0,1,2	05	00	02	21		
3,4,5	02	12	53	74	3,4,5	06	14	34	12		
6,7,8	03	25	22	67	6,7,8	07	34	45	71		
9,10,11	04	65	34	52	9,10,11	08	53	18	11		
12,13,14	A0	01	00	00	12,13,14	A0	05	00	00		
15,16,17	A1	04	00	00	15,16,17	A1	08	00	00		
18,19,20	A2	73	25	72	18,19,20	A2	68	53	45		
21,22,23	01	00	02	13	21,22,23	01	00	02	21		
etc.					etc.						

Table 3.6     *The table of contents coded in a two-disc set with four tracks on each disc.*

number of the second disc is 05, the A0 field is set to 05, and the A1 field is set to 08.

In the program and lead-out area, shown in figure 3.31b, Mode 1 contains track numbers (TNO), index numbers (X) within a track, time within a track (Min, Sec, Frame) and absolute time (Amin, Asec, Aframe). TNO designates individual tracks and is set to AA during lead-out. Track running time is set to 0 at the beginning of each track (including lead-in and lead-out areas) and increases to the end of the track. At the beginning of a pause, time decreases, ending with 0 at the end of the pause. The absolute time is set to 0 at the beginning of the program area (the start of the first music track) and increases to the start of the lead-out area. Index numbers both separate and subdivide tracks. When set to 00, X designates a pause between tracks, and countdown occurs. Non-zero X values set index points inside tracks. A 01 value designates a lead-out area. Using indexing, up to ninety-nine locations within tracks can be indexed. Mode 1 information occupies at least nine out of ten successive subcode blocks. Figure 3.32 summarizes the timing relationships contained in Mode 1 Q channel information.

In Modes 2 and 3, the program and time information is replaced by other kinds of data. These are displayed in figure 3.33. Mode 2 contains a catalog number of the disc, such as the bar-code of UPC (Universal Product Code). This data remains unchanged throughout the disc. Mode 2 also continues absolute time count from adjacent blocks. Mode 3 is used to give ISRC for identifying each music track; a different code can be written for each track. The ISRC consists of twelve alphanumeric characters (five 6-bit bytes with 000000 through 001001 representing 0 through



**Figure 3.32** The Q channel stores track numbers, index points, and program and absolute times in order to form an information map across the disc. (Courtesy Sony)

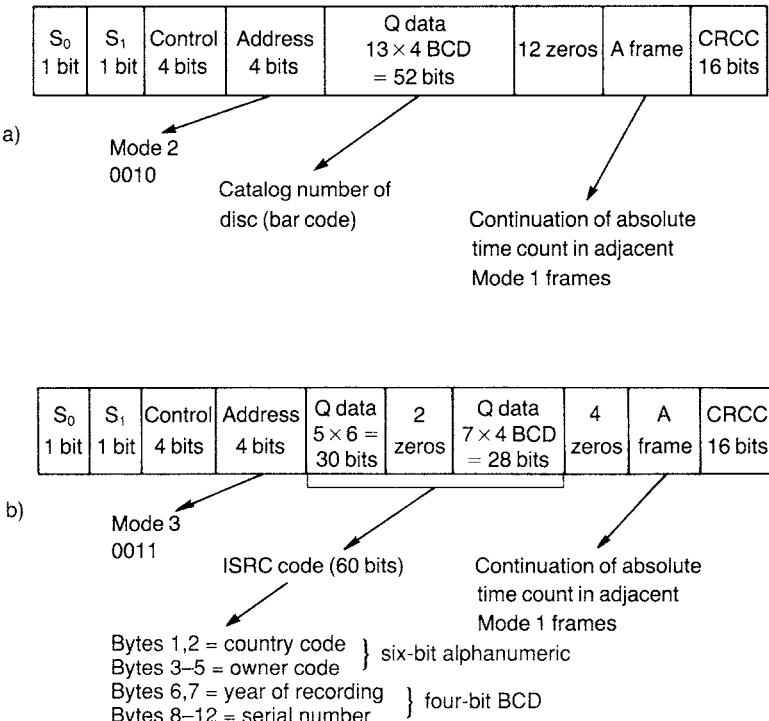


Figure 3.33 The Q channel Mode 2 contains the disc catalog number (a), and Mode 3 contains an ISRC code (b).

9 and 010001 through 101010 representing A through Z) and seven BCD bytes (4 bits each). It can identify a recording by country, owner, year of recording, and serial number. ISRC is defined according to the ISO 3901-1986 and DIN-31-621 standards. Mode 3 also continues absolute time count.

Although a system is not currently implemented, the ISRC provides an almost ideal method for broadcasters to log CD playback for copyright purposes. Currently, broadcasters must manually log which songs are played so that royalties can be paid to composers and performers. Using CD players modified for broadcasting, the ISRC could be used to trigger computerized payment logs, to be delivered to the appropriate copyright organizations. At this time, however, most discs are not coded with ISRC.

Modes 2 and 3 can be deleted from the subcode if they are not required. If they are used, Mode 2 and Mode 3 must occupy at least 1 out

of 100 successive subcode blocks. In addition, Mode 2 and 3 data can be present only in the program area.

Every audio compact disc contains additional data capacity in the subcode bits. The other 6 bits (R,S,T,U,V, and W), which account for about 24 megabytes of 8-bit storage per disc, are available for other data, such as video information. The R-W subcode area was reserved as contingency storage by the original disc designers. Manufacturers promote this feature of the CD standard through the CD + G/M (Compact Disc + Graphics/MIDI) format, which will be discussed in Chapter Six.

### **Twenty-Bit CDs?**

Although the 16-bit word length of the compact disc is certainly more than adequate for most consumer applications, a higher-resolution disc, with larger signal-to-noise ratio, would be useful for professional applications (particularly if the disc were also recordable and erasable). Given the existing CD format, it would be possible to encode 20-bit words at the standard sampling frequency, using data reduction and the subcode channel for added capacity. During encoding, the 20-bit input signal would be compressed with a floating-point compander; the new nonlinear samples would be represented by 16-bit mantissas and 3-bit exponents. The 16-bit mantissa would be recorded like a standard CD signal, and the 3-bit exponent would be recorded in the subcode field. During each subcode block rate of 75 Hz, 588 stereo samples would appear. These could be divided into twelve 49-sample blocks. Each of these blocks would be assigned a 3-bit exponent, one per stereo channel, according to the absolute value of the largest sample. This 6-bit exponent information could be contained in the 6-bit R-W subcode area without affecting the PQ subcode. Because only  $12/64$  of the overall R-W area is needed, other R-W bits could be used for additional error protection of the exponent information. In practice, noise-shaping is used to code 20-bit master recordings into a CD's usual 16-bit data field. This takes advantage of psychoacoustics to place quantization noise in relatively inaudible high-frequency regions.

### **FOR FURTHER READING**

Berlekamp, E. R. *Algebraic Coding Theory*. New York: McGraw-Hill, 1968, and Laguna Hills, Calif.: Aegean Press, 1983.

\_\_\_\_\_. "Error Correcting Codes for Digital Audio." In Blessler, B., B. Locanthi, and T. G. Stockham, Jr., eds. *Collected Papers of the AES Premiere Conference—Digital Audio*. New York: Audio Engineering Society, 1983.

- Carasso, M. G., J. B. H. Peck, and J. P. Sinjou. "The Compact Disc Digital Audio System." *Philips Technical Review* 40(6):151–55, 1982.
- Doi, T. "Error Correction for Digital Audio Recordings." In Blessler, B., B. Locanthy, and T. G. Stockham, Jr., eds. *Collected Papers of the AES Premiere Conference—Digital Audio*. New York: Audio Engineering Society, 1983.
- Hamming, R. W. *Coding and Information Theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1980.
- Heemskerk, J. P. J., and K. A. Schouhamer-Immink. "Compact Disc System Aspects and Modulation." *Philips Technical Review* 40(6):157–64, 1982.
- Hoeve, H., H. H. J. Timmermans, and L. B. Vries. "Error Correction and Concealment in the Compact Disc System." *Philips Technical Review* 40(6):116–73, 1982.
- IEC (International Electrotechnical Commission). "Compact Disc Digital Audio System." Standard BNN15-83-095.
- Lin, S. *An Introduction to Error-Correcting Codes*. Englewood Cliffs, N.J.: Prentice-Hall, 1970.
- Nöldeke, C. "Compact Disc Diffraction." *The Physics Teacher* 28(7):484–85, 1990.
- Ogawa, H., and K. A. Schouhamer-Immink. "EFM—The Modulation Method for the Compact Disc Digital Audio System." In Blessler, B., B. Locanthy, and T. G. Stockham, Jr., eds. *Collected Papers from the AES Premiere Conference—Digital Audio*. New York: Audio Engineering Society, 1983.
- Peterson, W. W. "Error-Correcting Codes." *Scientific American* 206(2):96–108, February, 1962.
- Pohlmann, K. C. *Principles of Digital Audio*. Indianapolis: Sams, 1989.
- Reed, I. S., and G. Solomon. "Polynomial Codes Over Certain Finite Fields." *J. Soc. Indust. Appl. Math.* 8:300–304, 1960.
- Shenton, D., E. Debenedictis, and B. Locanthy. "Improved Reed-Solomon Decoding Using Multiple-Pass Decoding." *Journal of the Audio Engineering Society* 33(11):878–87, November, 1985.
- Vries, L. B., and K. Odaka. "CIRC—The Error Correcting Code for the Compact Disc Digital Audio System." In Blessler, B., B. Locanthy, and T. G. Stockham, Jr., eds. *Collected Papers from the AES Premiere Conference—Digital Audio*. New York: Audio Engineering Society, 1983.

- Watkinson, J. R. "Inside CD, Part 1." *Hi Fi News and Record Review* 31(11):31–37, November, 1986.
- \_\_\_\_\_. "Inside CD, Part 3." *Hi Fi News and Record Review* 32(1):37–39, January, 1987.
- \_\_\_\_\_. "Inside CD, Part 6." *Hi Fi News and Record Review* 32(4):45–47, April, 1987.



## *Chapter Four*

# CD Player Design

### **INTRODUCTION**

A compact disc player is perhaps the most sophisticated and electronically complex piece of home entertainment equipment to ever reach the consumer. Yet even the most advanced products often have humble origins. For example, when Sony set out to design the first portable CD player, chief engineer Kozo Ohsone made a block of wood about 5 inches square and 1½ inches high and presented it to the design team, informing them that the portable player would be that size. To make sure that none of the wily engineers tried replacing it with a bigger block, he signed his name on the bottom.

### **PLAYER OVERVIEW**

Every compact disc player contains two primary systems: the audio data processing system and the servo and control system. The data path directs the signal from the pickup (carrying information derived from the modulated light received from the disc) through a series of processing circuits, ultimately yielding a stereo analog signal. The data path of a contemporary CD player consists of elements such as the data separator, the buffer, the de-interleaving RAM, the error correction and concealment circuits, the oversampling filter, the D/A converters, and the analog output filters as shown in figure 4.1. The servo, control, and display system must direct the mechanical operation of the player, including spindle drive, auto-tracking, auto-focus, and the user interface to the player's controls and displays.

A CD player employs a readout system using a semiconductor laser. A noncontact means is required to read data, maintain the laser beam's focus, and track the pit spiral. The result is a highly sophisticated pickup,

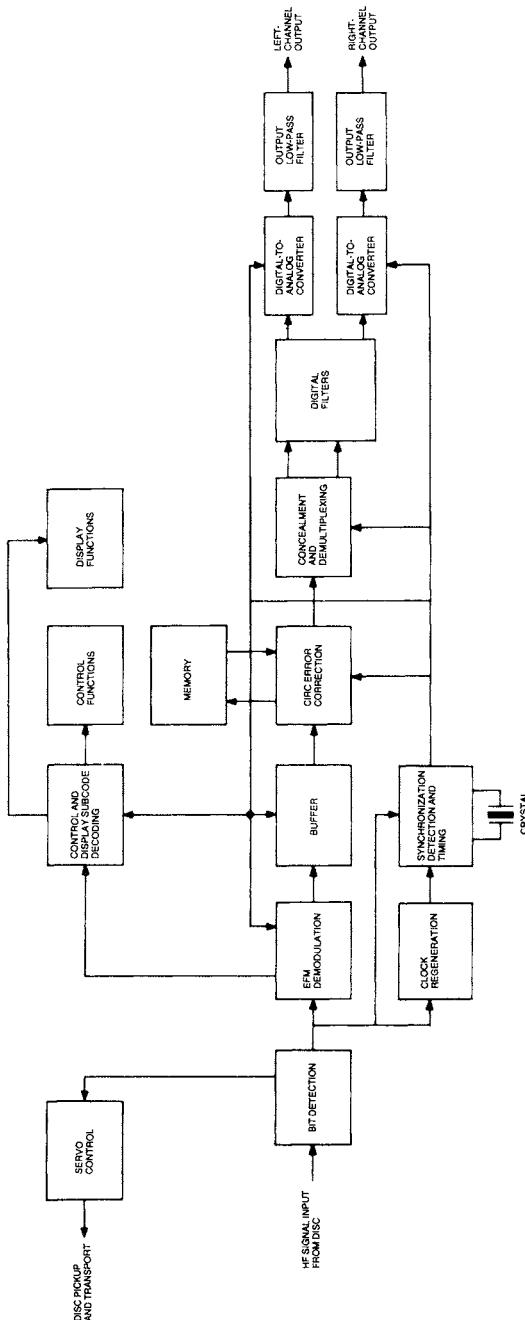


Figure 4.1 *Block diagram of a CD player showing audio path as well as servo and control functions.*

utilizing complex optical devices within servo loops. The servos use electrical signals from the pickup to control motors to mechanically adjust the pickup's position horizontally and vertically, relative to the disc surface.

A spindle motor is used to rotate the disc with constant linear velocity. Thus the disc must vary its speed depending on where the pickup is located underneath the surface. In another servo loop, information from the data itself is used to determine the correct rotating speed and maintain the proper data output rate.

A microprocessor monitors user controls and their interface to the player's circuitry. Various modes of player operation are controlled by software. Subcode data plays an important role in directing the pickup to the proper disc location. For example, a time code is used to locate the start of any track.

Once the data is recovered from the disc, the player must decode the audio information to reconstruct the audio signal. The EFM data is demodulated, errors are detected and corrected by decoding the error correction algorithm, and the audibility of gross errors is minimized through interpolation and muting. In many players, this corrected data is output through a digital serial transmission interface before also being processed into an analog waveform. Digital filters are used to oversample the signal so that image components may be easily removed with gentle analog output filters. The digital data must be converted to a stereo analog signal. This requires one or two D/A converters and analog low-pass filters.

A final output circuit is nothing more than an interface to the outside world. It ensures that the player's line level output is appropriate to drive external amplifiers with a minimum amount of analog distortion.

It is interesting to note that the various subsystems in a CD player are closely interrelated with a tightly interlocked timing relationship, as shown in figure 4.2. A few key numbers unlock the relationship: The audio data rate is 176.4 kbytes per second (44.1 kHz sampling of four 8-bit words). Because there are 24 audio bytes in a frame, the frame rate is 7350 Hz ( $176.4/24$ ). Because there are 588  $T$  periods in a frame, the master clock  $T$  is 4.3218 MHz ( $7350 \times 588$ ). A total of ninety-eight frames are needed to form one subcode block so the subcode rate is 75 Hz ( $7350/98$ ). Similarly, minimum and maximum recorded frequencies, EFM rate, and the subcode rate can be derived.

The remainder of this chapter examines all of these topics in more detail, focusing on the laser pickup, low-pass filters, D/A converters, and serial transmission output interface. Upstream intermediate signal processing was examined in Chapter Three.

It is important to remember that the electrical design of CD players has undergone three evolutionary steps since the format's introduction. First-generation players used multibit D/A converters followed by analog brick-wall filters to perform conversion and filtering of the signal. Second-

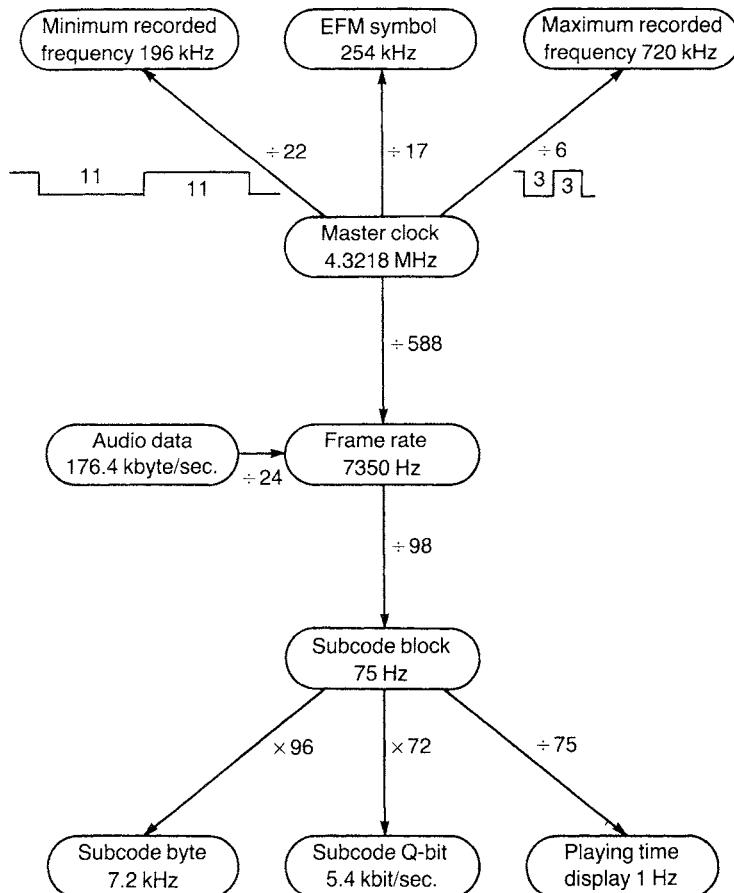


Figure 4.2     *Timing interrelationships in CD player design.*

generation players retained multibit D/A converters, but replaced analog brickwall filters with digital oversampling filters, with the filters placed prior to the converters. Third-generation players retain digital oversampling filters but use low-bit D/A converters instead of multibit converters. The two great evolutionary steps—the oversampling filters that replaced brickwall analog filters and the low-bit converters that replaced multibit converters—were made possible by the advent of powerful DSP (digital signal processing) technology.

## PICKUP DESIGN

The compact disc is certainly one of the most advanced storage media available. One reason is its optical pickup. A disc might contain three billion pits precisely arranged on a spiral track. The optical pickup must focus on, track, and read that data spiral. The entire lens assembly, a combination of the laser source and the reader, must be small enough to glide laterally beneath the disc, moving in response to tracking information and user random-access programming. Furthermore, the pickup must maintain focusing and tracking even under adverse playing conditions, such as the playback of a dirty disc or impact and vibration. It is perhaps the difficulty of this task that has encouraged considerable engineering diversity in optical pickup design to attempt to achieve the best way to retrieve data from CDs.

### *Auto-Tracking*

Intensity-modulated laser light carries the data, and nothing (except light) touches the medium surface. This poses an interesting engineering problem. How do we track a spiral pit track if there is no groove to guide the pickup? The answer, of course, is the auto-tracking system found in all CD players.

The spiral pit track on a compact disc, running from a center circle nearly to the outer edge, makes 22,188 laps over a 35.5 millimeter radial distance. A disc that is off center might exhibit track eccentricity of as much as 300 micrometers, which is the equivalent of sixty pit tracks. In addition, vibration can challenge the pickup's ability to track within a  $\pm 0.1$  micrometer tolerance. Engineers have devised four methods for tracking the pit spiral: (1) one-beam push-pull, (2) one-beam differential phase detection, (3) one-beam high frequency wobble, and (4) three-beam.

### *Auto-Focusing*

With 600,000 pits per second flying past, the optical pickup has its work cut out for it. Moreover, even the flattest disc isn't perfectly flat; the disc specifications acknowledge this by allowing for a vertical deflection of  $\pm 600$  micrometers. Meanwhile, the laser beam must stay focused within a  $\pm 2$  micrometer tolerance. Otherwise, the phase interference between the directed and reflected light is lost along with the audio data, as well as the tracking information, and, ironically, the focusing information. The objective lens must therefore be able to refocus as the disc surface deviates vertically. A servo-driven auto-focus system manages this, utilizing control electronics and a servo motor to drive the objective lens. Three techniques are available for generating a focusing signal: (1) a cylindrical

lens using astigmatism, (2) a knife edge using Foucault focusing, and (3) critical angle focusing.

Any pickup must, of course, perform both tracking and focusing simultaneously. Thus a completed design would utilize a combination of these techniques. Fortunately, concerned audiophiles are spared the task of analyzing all the permutations. When the focusing and tracking mechanisms are combined, two pickup designs stand out from the rest: (1) one-beam push-pull tracking with Foucault focusing and (2) three-beam tracking with astigmatic focusing.

These designs have been commercialized, with each manufacturer choosing its preference: one-beam or three-beam pickups. One-beam pickups are usually mounted on the end of a pivoting arm which swings the pickup across the disc in an arc. Three-beam pickups are always mounted on a sled which slides linearly across the disc. Let's consider both types of pickups, beginning with the three-beam design.

### ***Three-Beam Optical Design***

Although design particulars vary among manufacturers, three-beam pickups are generally similar in function. Figure 4.3 illustrates the optical path of a three-beam pickup.

To achieve sharp focus on the data surface and proper intensity modulation from the pit height, it is necessary to use a laser as the light source. Laser is an acronym for *light amplification by stimulated emission of radiation*. A laser uses an optical resonator to stimulate atoms to a higher energy level that induces them to radiate in phase. Laser light differs significantly from white, visible light. A light bulb, for example, radiates all the frequencies of the spectrum at all different phases. A laser light is monochromatic (composed of a single frequency), and is coherent in phase. Phase coherency is vital, of course, to implement interference cancellation in the beam reflected by disc pits so that disc data can be read. CD pickups use an aluminum gallium arsenide (AlGaAs) semiconductor laser with, for example, a 0.5 milliwatt optical output radiating a coherent-phase laser beam with a 780 nanometer wavelength (some manufacturers use 790 nanometers). The beam is thus composed of near-infrared light.

The laser diode is placed at the focal point of a collimator lens with a long focal distance. Its purpose is to make the divergent light rays parallel. A monitor diode is placed next to the laser diode to control power to the laser. It compensates for temperature changes and prevents thermal runaway. The monitor diode conducts current in proportion to the laser's light output. In other words, it stabilizes the semiconductor laser's output.

A three-beam pickup uses two additional beams for tracking. To generate these beams, the light from the laser passes through a diffraction grating, a screen with slits spaced only a few laser wavelengths apart, as

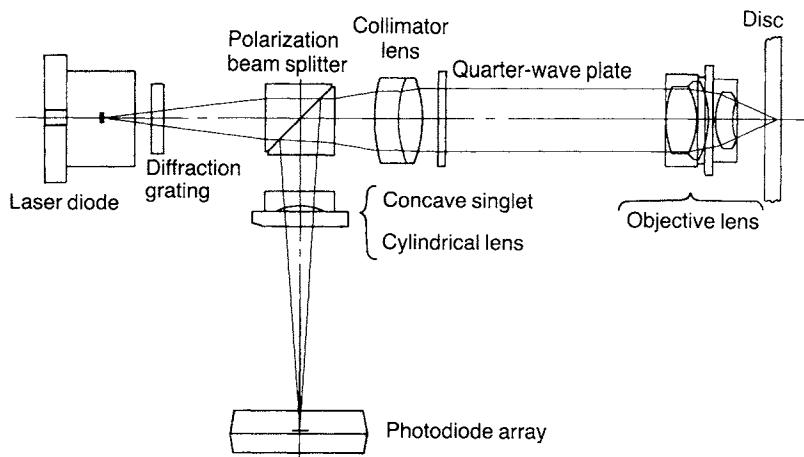


Figure 4.3    *Optical path of a three-beam pickup.*

shown in figure 4.4. As the beam passes through the grating, the light diffracts; when the resulting collection is again focused, it will appear as a single, bright, centered beam with a series of successively less intense beams on either side. It is this diffraction pattern that actually strikes the disc. A three-beam pickup uses the center beam for reading data and focusing and two secondary beams for tracking only. A one-beam pickup accomplishes all of these tasks with one beam.

The next part of the optical system, a polarization beam splitter (PBS), directs the laser light to the disc surface and angles the reflected light to the photo sensor. (In some designs, a half-silvered mirror is used.) The PBS consists of two prisms with a common 45 degree face acting as a polarizing prism. The collimator follows the PBS (in some designs, it precedes it). The light then passes through a quarter-wavelength plate (QWP), an anisotropic material that rotates the plane of polarization of the light beams (required to make the PBS work). Light which has passed through the QWP and been reflected from the disc back again through the QWP will be polarized in a plane at right angles to that of the incident light. Because the PBS will pass light in one plane, (e.g., horizontally polarized) but reflect light in the other plane (e.g., vertically polarized), the PBS will properly deflect the reflected beam toward the photodiode sensor to read the digital data.

The final piece of optics in the path to the disc is the objective lens that is used to focus the beams on the disc data surface, taking into account the refractive index of the polycarbonate substrate (see figure 3.4). The objective lens focuses the laser light into a convergent cone of light. The

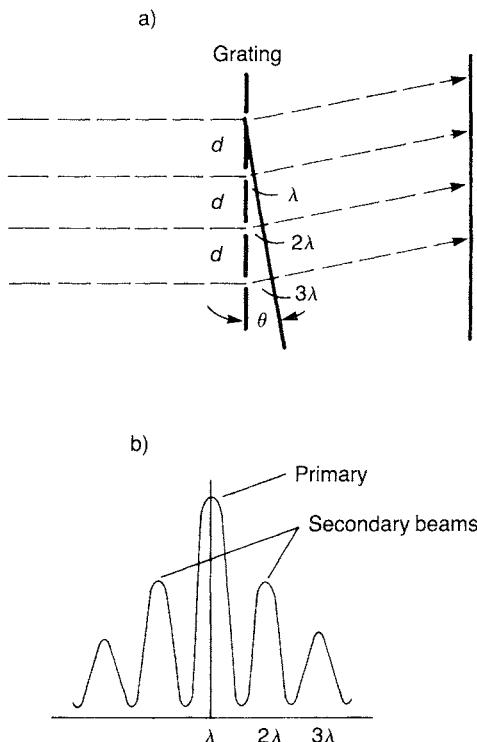


Figure 4.4     In a three-beam pickup a diffraction grating (a) splits the beam into successively less intense secondary beams (b).

convergence is a function of the numerical aperture (NA) of the lens. Most pickups use an objective lens with an NA of about 0.45 (this corresponds to f/1.0 in a photographic lens). The main spot is about 800 micrometers in diameter on the outer surface of the disc's transparent polycarbonate substrate. The refractive index of the polycarbonate substrate is 1.55 and its thickness about 1.2 millimeters, so the spot is narrowed to 1.0 micrometer at the reflective surface (see figure 3.4), two times wider than the pit width of 0.5 micrometer and comparable in width to the wavelength of the light itself.

The data encoded on the disc now determines the fate of the laser light. When the spot strikes land, the smooth interval between two pits, the light is almost totally reflected. When it strikes a pit (seen as a bump by the laser) with a depth in the transparent layer of about a quarter wavelength of the laser light, diffraction and cancellation due to interference cause less light to be reflected (see figure 3.5). All three intensity-modulated

light beams return through the objective lens, the QWP, the collimator lens, and the PBS. Finally, they pass through a singlet lens and a cylindrical lens en route to the photodiode.

### ***Three-Beam Auto-Focus***

Auto-focus is an absolute prerequisite in an optical pickup system. Disc warpage and other irregularities can cause vertical deviations in the disc data surface. This movement would place the data out of the pickup's depth of focus, making it impossible to create the necessary interference pattern with the pit height and land.

In a three-beam player, the unique properties of astigmatism are used to achieve automatic focusing. The cylindrical lens just prefacing the photodiode (see figure 4.3) is used to detect an out-of-focus condition. As the distance between the objective lens and disc reflective surface varies, the focal point of the optical system also changes, and the image projected by the cylindrical lens changes its shape. This is illustrated in figure 4.5. The change in the image on the photodiode generates the focus correction signal. When the disc surface lies precisely at the focal point of the objective lens, the image reflected through the intermediate convex lens and the cylindrical lens is unaffected by the astigmatism of the cylindrical lens, and a circular spot strikes the center of the photodiode. When the distance between the disc and the objective lens decreases, the image projected by the objective lens, the convex lens, and the cylindrical lens moves further from the cylindrical lens, and the pattern becomes elliptical. Similarly, when the distance between the disc and the objective lens increases, the image moves closer to the lens, and an elliptical pattern again results, but it is rotated 90 degrees from the first elliptical pattern.

The four-quadrant photodiode uses the light's intensity level from each of the quadrants to generate focus correction voltages. An elliptical pattern mainly striking quadrants A and C, as shown in figure 4.5b, indicates the disc is too near, and a positive voltage is created. When the disc is in focus, no net voltage is created from the round pattern. An elliptical pattern striking quadrants B and D indicates the disc is too far, which creates a negative voltage. This varying voltage, as shown in figure 4.5d, is used to correct the focusing mechanism continually towards a zero voltage, thus maintaining a focused laser beam. The center beam is used to convey data from the disc, using a signal summed from the four-quadrant photodiode ( $A + B + C + D$ ).

The correction voltage also generates a control signal in the form of  $(A + C) - (B + D)$ . This electrical input controls the mechanical motion of a servo motor. In this case the servo's job is to move the objective lens along the optical axis in response to vertical disc motion and thus maintain focus. This servo motor consists of a coil and permanent magnet structure

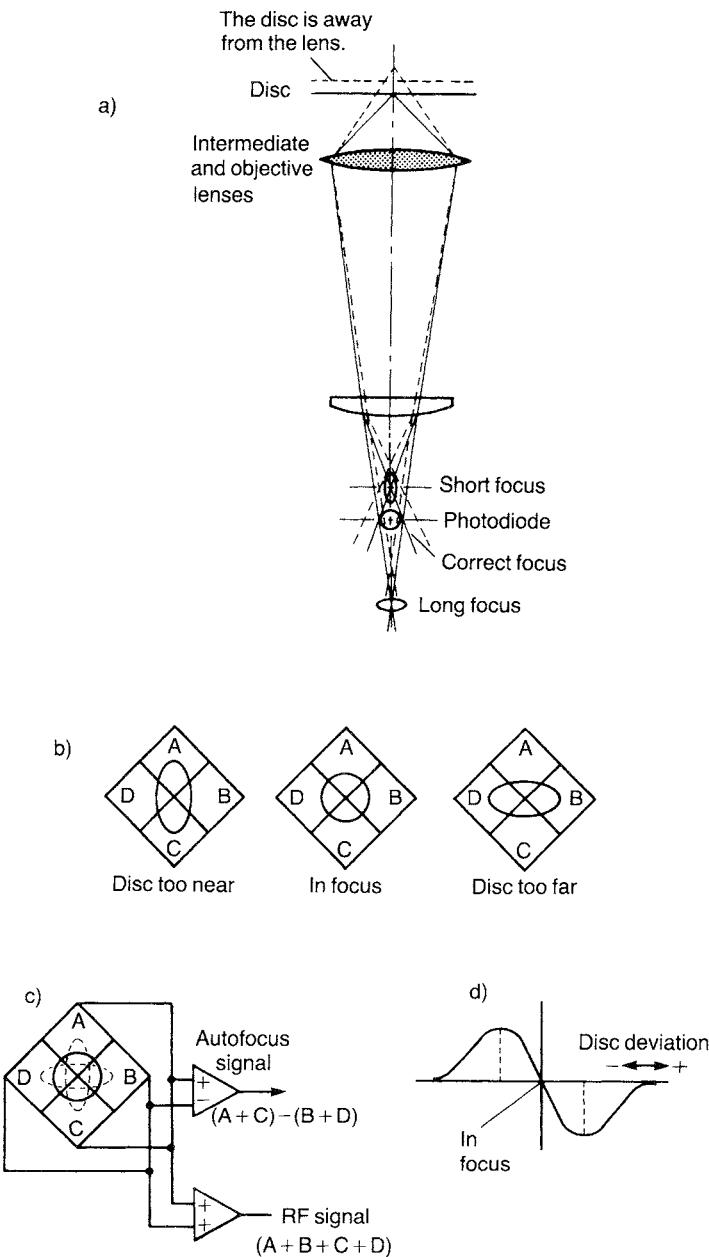


Figure 4.5 *A three-beam pickup uses the properties of astigmatism to generate an auto-focus correction signal.*

similar to that used in a loudspeaker, with the objective lens taking the place of the speaker cone, as shown in figure 4.6.

When a disc is first loaded, initial focus is achieved by injecting a signal into the system which causes the laser to turn on and the objective lens to move up and down several times until a focus condition is found. Then, the automatic focusing system takes over. If no disc is detected, the system typically tries again, shutting off if it fails again. When the auto-focus is not operative (e.g., disc tray is open) the system pulls the objective lens back to prevent damage to the lens or disc. The auto-focus system is slightly complicated, its task is complex, and it performs well at keeping the pickup floating underneath the spinning disc, maintaining focus to within a tolerance of approximately  $\pm 0.5$  micrometer.

### **Three-Beam Auto-Tracking**

The track pitch, the distance between adjacent laps of the pit spiral, is 1.6 micrometers on a CD, which is too fine a tolerance for any mechanical tracking system to be cost-effective. An auto-tracking system is required. In three-beam pickups, the two secondary beams are used for auto-tracking. The center beam, which carries the information from the disc, is accompanied by the two secondary beams from the point of generation at the diffraction grating. The tracking beams also strike the data surface and are reflected. Their varying intensities are used at two separate photodiodes mounted alongside the four-quadrant photodiode.

The central beam spot covers the pit track, while the two tracking beams are aligned above, below, and to either side of the center beam. Their relative position is firmly fixed. When the beam is tracking the disc properly, part of each tracking beam is aligned on the pit edge; the other part

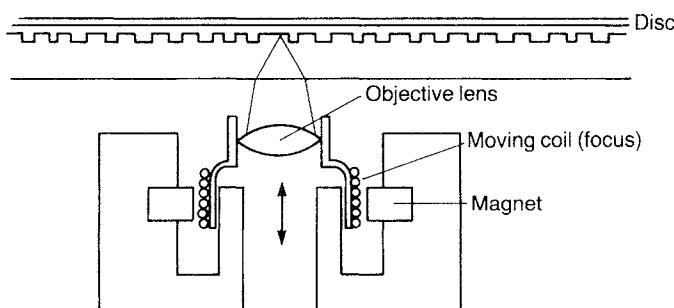


Figure 4.6    A coil and magnet structure is used to move the objective lens in the optical path.

covers the mirrored land between pit tracks. The three beams are reflected through the quarter-wave plate and polarizing beam splitter; the main beam strikes the four-quadrant photodiode, and the two tracking beams strike two separate photodiodes (E and F) mounted to either side of the main photodiode, as illustrated in figure 4.7.

As the three spots drift to either side of the pit track, the amount of light reflected from the tracking beams varies. There is less average light intensity reflected by the beam which encounters more pit area and greater

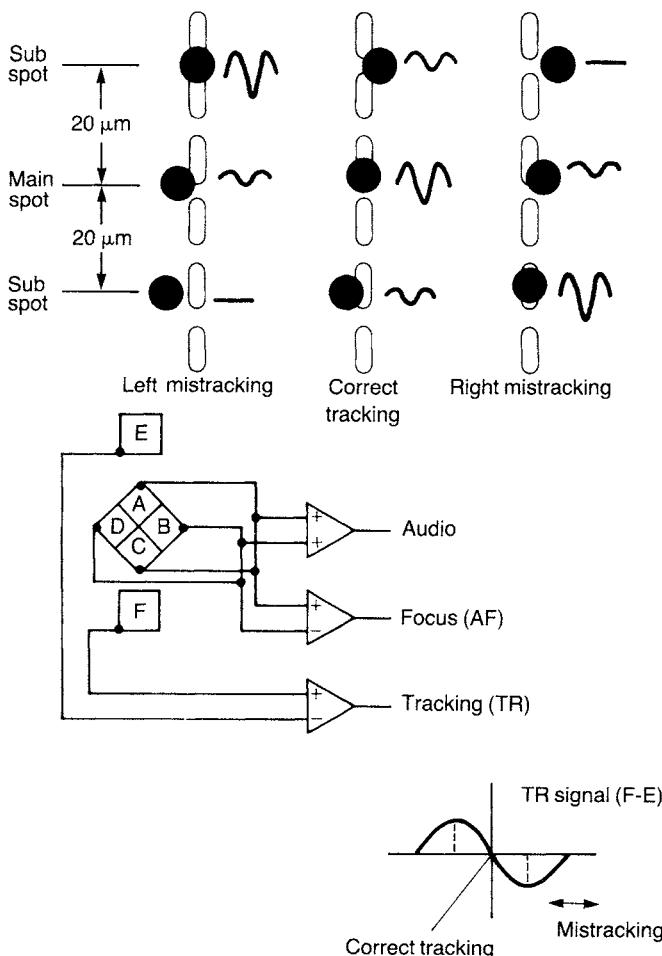


Figure 4.7    In a three-beam pickup, two tracking beams are precisely aligned to either side of the center of the pit track. When reflected, they strike outrigger photodiodes to generate an auto-tracking correction signal.

reflected light intensity from the beam which encounters less pit area. The relative output voltages from the two tracking photodiodes thus form a tracking correction signal (F-E).

The varying output from the tracking photodiodes is converted into a useful correction signal. Because the tracking beams are aligned to different areas of the disc, the signal from the leading beam is delayed 30 microseconds to permit comparison on the basis of the same pit, as if they were reading the same pit simultaneously. The outputs are constantly compared and used to control a servo system. If the tracking is precisely aligned, the difference is zero. If the beams drift, a difference signal is generated, varying positively for a left drift, for example, and negatively for a right drift.

### ***Three-Beam Pickup Mechanism***

In many designs, both the tracking and focus difference signals are applied to a two-axis actuator, with a focus/tracking coil and a permanent magnet. The top assembly of the pickup is mounted on a base with a circular magnet ringing it. A circular yoke supports a bobbin containing both the focus and tracking coils. Control voltages from the focus circuit are applied to the bobbin focus coil, and it moves up and down with respect to the fixed magnet. The objective lens thus maintains its proper depth of focus. The other axis of movement, from side to side, is used to achieve tracking accuracy. When the tracking difference signal is applied to the coil, the bobbin swings around a shaft to move the objective lens laterally, in the direction dictated by the polarity of the signal, so the main laser beam is again centered and the tracking signal is again zeroed.

A three-beam pickup requires a linear movement because the three beams must stay in a fixed position relative to the pit track. The three-beam pickup is thus mounted in a sled that moves radially across the disc. The sled can be driven by a lead screw; the pitch of the screw's thread is a compromise between the fine thread required for normal tracking (35 millimeters per hour) and the coarse thread needed for fast forward and reverse (perhaps 35 millimeters per second). In some designs, a moderate thread is used to achieve disc access, and a movable mirror is inserted into the laser's path to deflect the beam quickly to handle small differences, such as track eccentricity. Alternatively, many designs use linear motors to provide smooth and rapid movement of the pickup across the disc surface.

### ***One-Beam Optical Design***

The optical components of a one-beam design are shown in figure 4.8, along with the photodiode array used to generate tracking and focusing

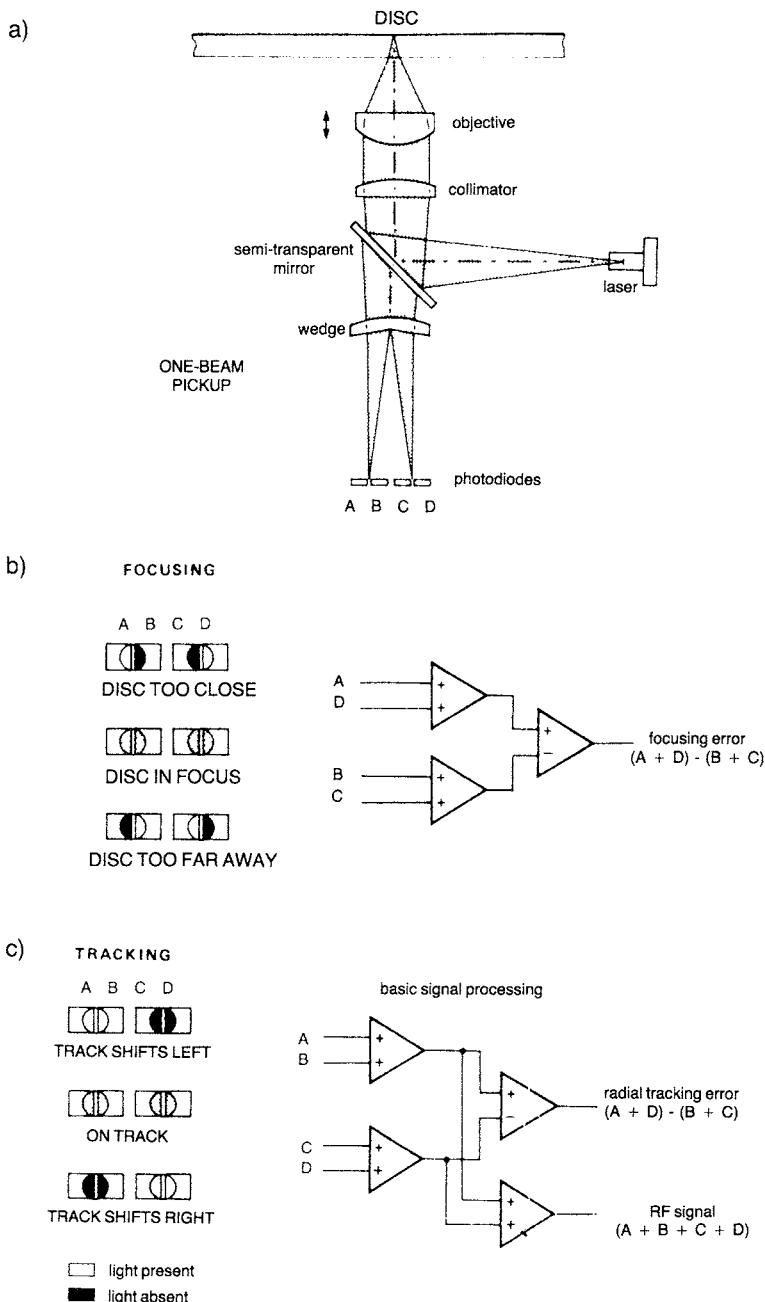


Figure 4.8 A one-beam pickup (a) derives data, focus, and tracking signals from one beam. Auto-focusing (b) utilizes the shift in split beams. Auto-tracking (c) utilizes asymmetry in the beam, caused by beam interference. (Courtesy Philips)

signals, as well as the data signal. A laser diode generates the laser beam which strikes a semitransparent mirror (used in place of a PBS and QWP). A collimator and objective lens focus the beam on the disc's reflective pit track surface. As the pits fly by, the data encoded as pits and land results in a reflected laser beam with varying intensity.

On its return path, some of the reflected laser light passes through the semitransparent mirror and strikes a wedge lens. It is then split into two beams, adjusted to strike an array of four horizontally arranged photodiodes. The outputs of all the photodiodes are summed to provide the data signal ( $A + B + C + D$ ), which is demodulated to yield audio data and speed control signals.

### ***One-Beam Auto-Focus***

Auto-focus is accomplished by the four photodiodes. As noted, Foucault focusing is often used. When correct focus is achieved, two sharp images are located between photodiode pairs. When focus varies, the focal point of the system is shifted, and the split beams draw closer together (when the disc is too far) or further apart (when the disc is too near), as shown in figure 4.8b. The difference between diode pairs A/D and B/C forms a focus error signal  $(A + D) - (B + C)$ , used to maintain the focus of the servo-driven objective lens. The capture range of the one-beam focus system is sufficient to accommodate normal tolerance without special start-up focusing procedures, as in three-beam designs.

### ***One-Beam Auto-Tracking***

The diodes also generate error signals for auto-tracking. When the laser spot strikes the center of the pit track, a symmetrical beam is reflected. As the laser beam wanders from the pit track, interference creates asymmetry in the beam, resulting in an intensity difference between the two split beams, as shown in figure 4.8c. For example, as the beam moves off track, one side of the beam encounters more reflective land; consequently, less interference occurs on that side of the beam. Thus reflected light is more intense there. As a result, the split beam derived from that side of the beam will be more intense, and the photodiode's output will be greater. Either photodiode pair A/B or C/D generates a greater signal from the increased laser light. The difference between the pairs  $(A + B) - (C + D)$  produces an error signal used to maintain the pickup on track.

Aging or soiling of the one-beam pickup could cause the reflected beam to become increasingly asymmetrical. This asymmetry would generate a constant offset in the tracking correction signal, causing the pickup to remain slightly off track, requiring readjustment. To prevent this, a sec-

ond tracking error signal is generated. A low-frequency (e.g., 600 Hz) alternating voltage is applied to the coil that controls the tracking. The output signal from the four photodiodes is thus modulated by the alternating tracking signal. As the pickup wanders off track, a deviation in the tracking signal occurs in the modulated signal, as shown in figure 4.9. When the modulation signal is rectified, a drift-free tracking error signal is produced. This tracking signal is used to correct the primary tracking signal with a direct voltage. The data signal from the four photodiodes is thus always returned to its maximum value. In practice, of course, the four photodiodes must simultaneously perform all three tasks of data reading, tracking, and focusing.

Although most one-beam pickups are mounted on a pivoting arm, a linear sled may be used as well. In the former, the pivoting arm describes a radial arc across the data surface. A coil and permanent magnet are mounted around the pivotal point of the arm. When the coil is energized, the pickup may be positioned anywhere on the pit track and its position corrected by the auto-tracking system. The objective lens is provided with a coil and a permanent magnet for displacement in the direction of the optical axis to achieve auto-focusing.

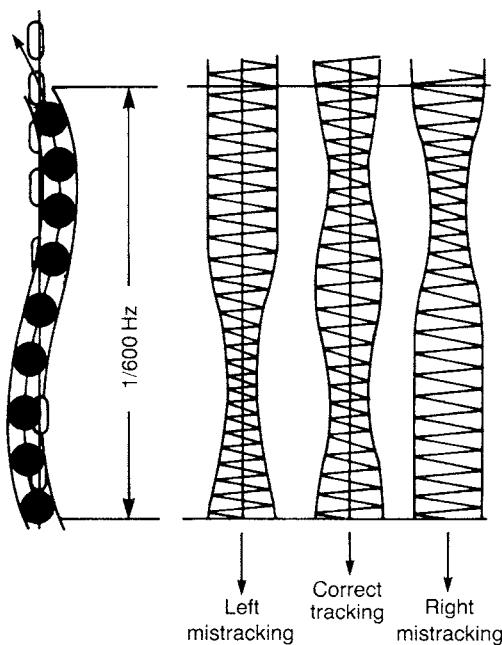


Figure 4.9     *In a one-beam pickup, a modulation signal is sometimes injected into the servo system to compensate for any tracking offset.*

### **Pickup Control**

Aside from the tracking accuracy needed to keep the laser beam on track, the pickup must properly move across the disc surface to track the entire pit spiral. The pickup must also be able to jump from one place on the disc to another.

In the search mode, the player's microprocessor takes command to provide faster motion than is possible during normal tracking. Control signals are directed to the pickup for accelerated movement in the forward or reverse direction. Subcode data added to each CD frame is used to determine the pickup's location on the pit spiral.

For forward or reverse jumps to programmed locations on the disc, the tracking signal is disabled, and the microprocessor controls the pickup. Just prior to alignment, a brake pulse is generated to compensate for the pickup's inertia. When the correct location is reached, the microprocessor informs the system that proper alignment is imminent, the tracking signal is again activated, the pickup comes to rest on the correct track, and the system resumes auto-tracking.

It is interesting to note that the movement of a CD player's pickup is similar to that of a stylus in an LP groove. With the aid of the auto-tracking system, the pit signal "pulls" the pickup across the disc in the same way that the phono cartridge is pulled across an LP by the groove. Of course, a CD pickup uses a lot more technology than a phonograph cartridge to achieve tracking, but most would agree that the outcome is much more satisfactory, especially in terms of the longevity of CDs compared to LPs.

## **DIGITAL SIGNAL PROCESSING**

Digital signal processing (DSP) is a powerful technology, and, in fact, without DSP the CD system would not be a viable reality. It is DSP that permits error correction, digital filtering, low-bit D/A conversion, and other crucial CD player operations. Before we tackle the important topics of digital filtering and low-bit conversion, let's take a look at DSP.

DSP is a technology used to analyze, manipulate, or generate signals in the digital domain. A signal is a time-based sequence in which the ordering of values is important. For example, a digital audio signal makes sense, and can be processed properly, only if the sequence is properly preserved. DSP is thus a special application of general data processing.

DSP is the workhorse of modern engineering. Its extensive applications in telecommunications includes modems, cellular systems, data transfer between computers, vocoders, and transmultiplexers. DSP image processing is used to process photographs received from orbiting satellites and deep

space vehicles. Television studios use digital signal processing techniques for manipulating picture signals. Other applications include robotics, machine vision, pattern recognition, digital television, computer animation, flight simulators, missile/torpedo control, vibration analyzers, navigation systems, automobile anti-lock brakes, active suspension, and motor and emissions controls.

Of course, digital audio systems also present rich possibilities for DSP. Error correction, multiplexing, sample rate conversion, speech and music synthesis, data reduction, volume/fader/balance, filtering, adaptive equalization, dynamic compression and expansion, crossovers, reverberation, ambience processing, time alignment, acoustic noise cancellation, mixing and editing, and acoustic analysis can all be performed with digital signal processing. An audio DSP system is in reality a computer dedicated to the processing of audio signals.

### *Advantages and Disadvantages*

In theory, digital signal processing, like digital audio in general, provides many performance advantages over analog signal processing. Digital processing is more precise and repeatable, and it can perform operations that are impossible with analog techniques. Noise and distortion can be much lower with DSP; thus auto fidelity is much higher. In addition, whereas analog circuits age, lose calibration, and are susceptible to damage in harsh environments, DSP circuits cannot lose calibration and are much more robust. In that respect the technology is well suited for consumer applications where everyday use over long periods puts great wear on products.

However, DSP is not a simple technology to implement. The challenge becomes apparent when design engineers begin to amass the hardware and software needed to process the signal. Aside from the sheer volume of computations needed, a number of difficult problems can arise. For example, truncation errors occur when the numerical result of a calculation such as addition is limited to a certain number of bits without regard to the value of the number. If a number such as 2.88 was simply truncated to 2.0, a considerable error would occur, and the error would be compounded as many calculations take place, each using truncated results. The resulting numerical error would be manifested as a distortion of the output signal. To avoid this, all DSP computations on audio signals must be of extremely high accuracy. This requires long word lengths; many DSP chips employ digital words that are 32 bits in length or longer.

In addition, even simple DSP operations may require several intermediate calculations, and complex operations may require hundreds of operations. To accomplish this, the hardware must execute the steps very quickly. Because all computation must be accomplished in real time—

that is, within the span of one sample period (perhaps 22 microseconds)—the processing speed of the system is crucial. DSP chips must often process twenty or thirty million instructions per second; in fact DSP chips are among the fastest processors in use.

## FILTER DESIGN

Classic digital audio design dictates an analog brickwall filter at the input and output of the digitization system. Today, most systems employ digital filters. However, to understand the purpose of filtering and the advantages of digital filtering, our discussion must begin with a look at analog filters.

### *Analog Filters*

Filtering is a fact of life for digital audio systems. An input anti-aliasing filter must precede the sampler to uphold the Nyquist theorem's criteria for bandlimited, and thus lossless, sampling. Similarly, the output anti-imaging filter must filter out all frequencies above the half-sampling frequency. The analog signal at the output of the digital-to-analog converter is a pulse amplitude modulation waveform whose staircase appearance is easily spotted by waveform watchers. These sudden shifts in amplitude represent high-frequency components not present in the original analog waveform; such artifacts of sampling must be removed to create a smooth waveform that eliminates high-frequency components. Even when digital filters are employed, a gentle analog smoothing filter is still required at the output.

It is the output filter that reconstructs the audio waveform. Although the idea of smoothing the output samples to remove high frequencies is entirely correct, a more mathematical analysis shows the inner workings. A low-pass filter has a  $\sin(x)/x$  impulse response. This means that a single rectangular audio sample (of infinitesimal width) passing through an ideal low-pass filter would emerge with a waveform in the shape of this  $\sin(x)/x$  curve, as shown in figure 4.10a. If the low-pass filter has a cutoff frequency at the half-sampling point ( $f_s/2$ ), then the  $\sin(x)/x$  curve passes through zero at multiples of  $1/f_s$ .

When a series of audio samples passes through the filter, the resulting waveform is the summation of all the individual  $\sin(x)/x$  contributions of the individual samples, as shown in figure 4.10b. Each sample's impulse response is zero at the  $1/f_s$  position of any other sample's maximum response. The output waveform has the value of each sample at sample time and the summed response of the impulse response of all samples in between sample points. This superposition of individual impulse responses

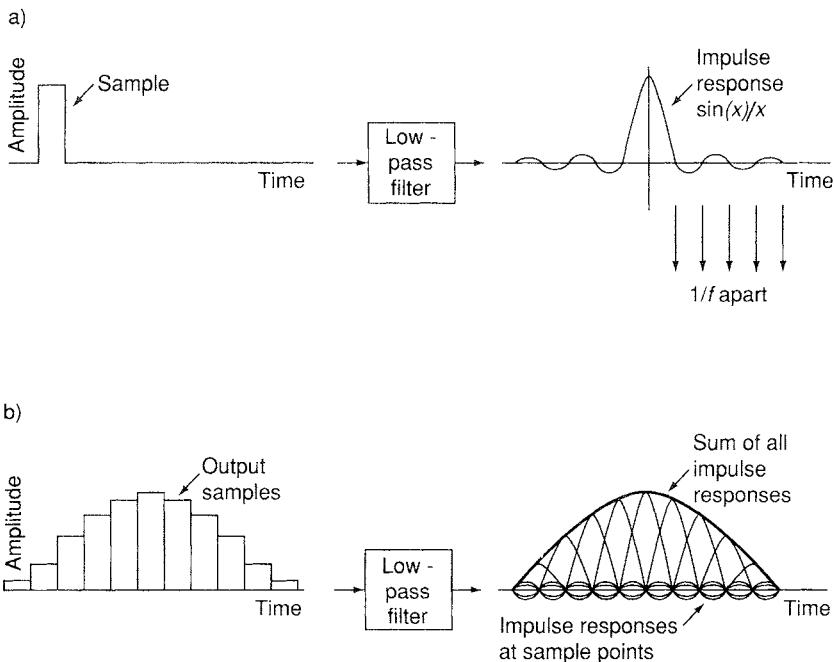


Figure 4.10     *The impulse response of an ideal low-pass filter is a  $\sin(x)/x$  function (a). Only the center portion of the impulse response is shown. A series of impulses (samples) results in the summing of the individual impulse responses, which forms the output waveform (b).*

results in an output waveform that is identical to the input low-pass filtered audio waveform. It is thus the filter's impulse response to the audio samples that reconstructs the original waveform.

Sharp-eyed readers might question the need to worry about frequencies such as 88.2 or 176.4 kHz, since they lie so far above the limits of audibility. Why waste money on a filter, when the ear itself is rather effective at filtering everything beyond 20 or 25 kHz? The original waveform is reproduced without filtering, but the accompanying image bands could stress the player's output amplifier or cause intermodulation in downstream analog equipment through which the signal passes. For example, the high-frequency components might create intermodulation components with base-band frequencies or with the high-frequency bias current of an analog tape recorder. That, in turn, could add audible distortion to the signal. Filtering is therefore mandated.

### Ideal Filters

Because engineers have been designing analog filters for a long time, we would expect little trouble with this particular assignment. Both the input and output filters can share similar designs, and the design criteria can be easily summarized. Ideally, we would like to attenuate all audio frequencies above the half-sampling frequency without affecting the lower audio-band frequencies. Moreover, we would like the transition to occur instantaneously so that the useable bandspace is extended as far as possible in order to yield an extended and flat frequency response. Thus an ideal filter, such as the one in figure 4.11, would have a flat pass-band (the audio range the filter passes), an immediate or brickwall filtering characteristic, and a stop-band (the frequencies the filter eliminates) attenuated to below the system's quantization resolution. In addition to these frequency response criteria, an ideal filter would not affect the phase or any other time-domain characteristic of the signal.

### Actual Filters

Although an ideal filter may be approximated, in practice its implementation presents a number of engineering challenges. A brickwall design requires compromise in other specifications, such as flat pass-band and low phase distortion. To alleviate these problems inherent in a brickwall response, we could design filters with more gradual cutoff which would not exhibit phase nonlinearities. However, the frequency of the half-sampling point would have to be increased to make sure that it was placed in a sufficiently attenuated part of the filter characteristic. This is the case

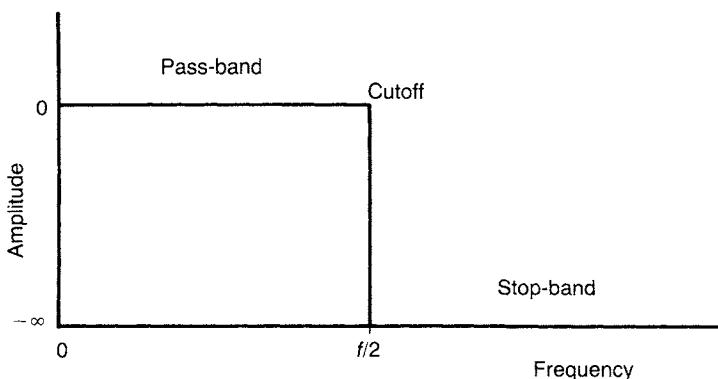


Figure 4.11 An ideal low-pass filter has a flat pass-band, instantaneously sharp cutoff, and infinitely attenuated stop-band.

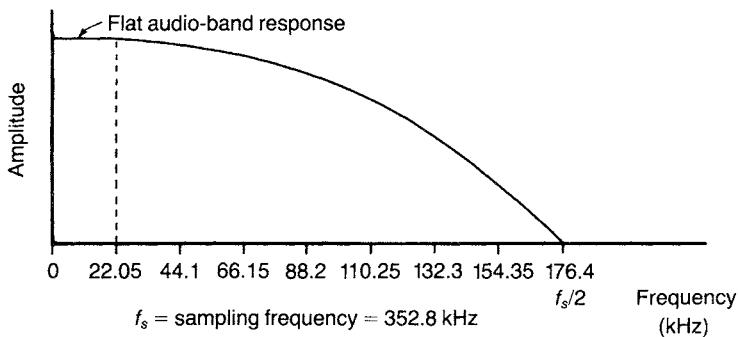


Figure 4.12 *Problems associated with brickwall filters can be overcome with more gentle filter characteristics. However, a higher sampling rate would be required to ensure a flat audio band.*

shown in figure 4.12. Consequently, a higher sampling frequency would be needed to achieve the same flat audio frequency response. To limit the sampling rate and make full use of the bandspace below the half-sampling point, a brickwall filter, at both input and output of the system, is the only alternative unless, as we shall see later, the technique of oversampling is employed to increase the sampling rate.

Let's consider an analog output filter design, such as those found in some early compact disc players. With a sampling rate of 44.1 kHz, the output filters (one for each channel) are usually designed for flat response from DC to 20 kHz. This provides a guard-band of about 2 kHz to ensure that attenuation is sufficient at the half-sampling point. The pass-band undoubtedly exhibits some frequency response irregularity, called ripple, which is typically specified to be less than  $\pm 0.1$  dB, as represented in figure 4.13. The stop-band's attenuation is designed to be equal to or greater than the system's dynamic range, as determined by word length. Ideally, a 16-bit system would require a stop-band attenuation of greater than 95 dB. The stop-band also typically exhibits ripple.

### **Cascade Filter Design**

Although there are many types of filter design, only a few are able to achieve the performance criteria required of a high fidelity digital audio system. One filter type often employed is an elliptical filter, sometimes called a Cauer filter. The design generally uses a low-pass filter combined with a band-reject notch filter to sharpen the cutoff. The elements of an elliptical filter interact with each other, making them challenging to design. High precision parts are also required.

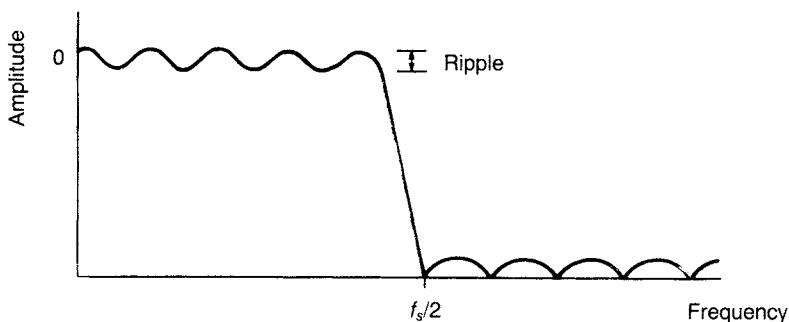


Figure 4.13 Actual low-pass filters exhibit ripple in both the pass- and stop-bands. In addition, cutoff is not instantaneous.

Basic filter elements of various types can be cascaded (repeated in series) to sharpen the cutoff. As the number of cascaded stages is increased, the filter characteristic steepens until the ideal filter frequency response is approximated. Unfortunately, as the cutoff steepens, the phase shift increases as well. Compact disc players might require ninth-order filters with characteristics such as those indicated in figure 4.14. The cutoff shown in figure 4.14a looks like the north rim of the Grand Canyon, but the phase shift shown in figure 4.14b might exceed 90 degrees at 20 kHz. This phase shift introduces nonlinear time delay across the audio band.

The resulting group delay shown in figure 4.14c, which measures the change of phase shift with respect to frequency, causes high frequencies to be shifted in time (delayed) relative to lower frequencies. The delay increases toward the cutoff frequency of an analog filter in a CD player. This creates a kind of time smear in the signal. (Ultrasonic frequencies are theoretically delayed “forever,” which is how they are filtered out.)

Absolute, non-frequency-dependent phase shift is inaudible. Recordings of Enrico Caruso may be ninety years old, which is a lot of time delay, but the CD suffers no distortion on that account, because all of the frequencies of the recording have been delayed by the same amount. But frequency-dependent group delay is different. Conceptually, we might compare such delay to the worst case example of an extremely strange loudspeaker in which the tweeter is located a mile behind the woofer. If properly equalized, steady-state tones would sound fine. But when the tone ended, we would hear sound emanating from the tweeter a full 5 seconds after the sound from the woofer had ended. Conversely, at the onset of a tone, the tweeter’s attack would arrive 5 seconds late. Such long frequency-dependent delays would obviously be audible. What about a

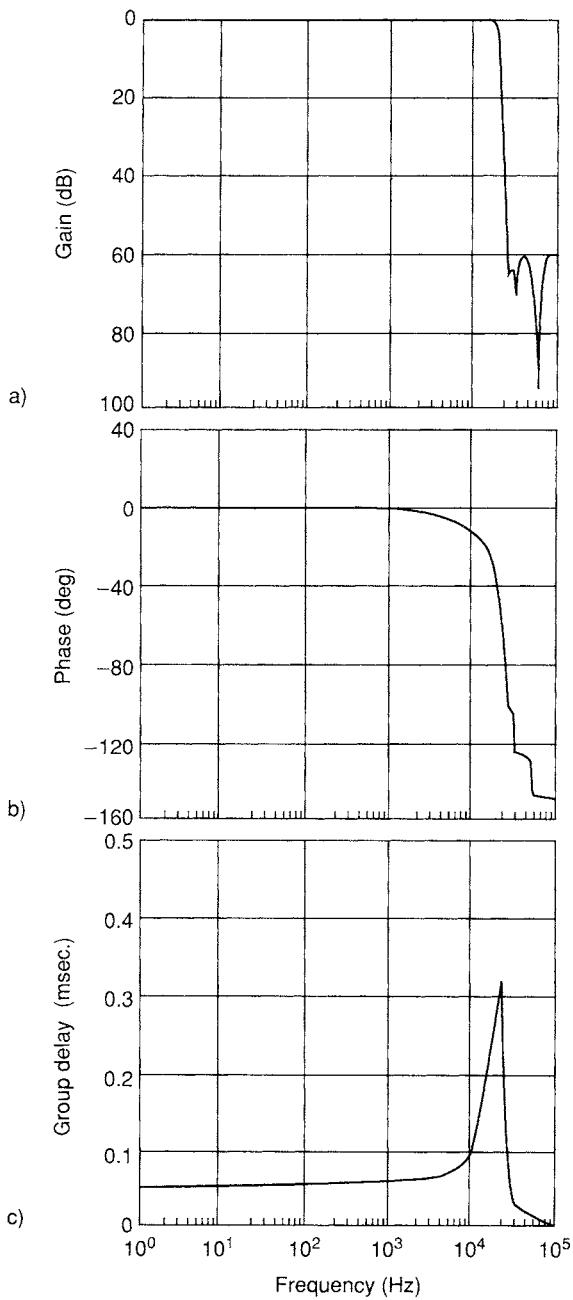


Figure 4.14    A brickwall filter will have a sharp cutoff (a), high-frequency phase shift (b), and high-frequency group delay (c). (Courtesy Sony)

real brickwall filter, in which delay at 20 kHz might be 300 microseconds relative to 0 Hz? Is that audible? The answer has yet to be fully determined, but manufacturers have certainly hesitated to feature the relevant specification in full color advertisements. Obviously, we would prefer to avoid any such group delays in frequency response. The point is that the analog filters necessary in CD players are not trivial design exercises.

## DIGITAL FILTERS

The digital filter used in today's CD players is a circuit which accepts audio samples and outputs audio samples. In other words, it is a DSP application. Filtering takes place precisely because the values of the output samples have been altered to produce filtering. In this case, the digital filter simulates the process of low-pass analog filtering and the waveform reconstruction it provides. Moreover, a technique called oversampling is utilized in which additional sample values are computed by interpolating between original samples. Because additional samples have been generated (perhaps two, four, or eight times as many), the sampling rate of the output signal is greater than the input signal.

Each intermediate sample is multiplied by the appropriate  $\sin(x)/x$  coefficient corresponding to its contribution to the overall impulse response of the filter. The multiplication products are summed together to produce the output filtered sample. It thus digitally simulates the impulse response of an analog filter. The spectrum of the signal is changed, with the images appearing at multiples of the new (oversampled) sampling rate. Because the distance between the baseband and sidebands is larger, a gentle analog filter can be used to remove the images without causing phase shift or other artifacts.

### *Time Delay*

Rather than using an analog brickwall filter to suppress the high-frequency image components after the signal has been converted to analog form, it is possible to process the digitized signal before D/A conversion using a digital filter. The basic mechanism used in the digital filters found in CD players is time delay. A delayed signal added back to itself results in a wholly new frequency response called a comb filter response. If a signal is delayed by time  $T$ , as shown in figure 4.15, and mixed with the original undelayed signal, the two will be out of phase, causing cancellations at odd multiples of the frequency, that is,  $1/(2T)$ . For example, a 1 millisecond delay will produce notches at 500 Hz, 1500 Hz, 2500 Hz, etc.

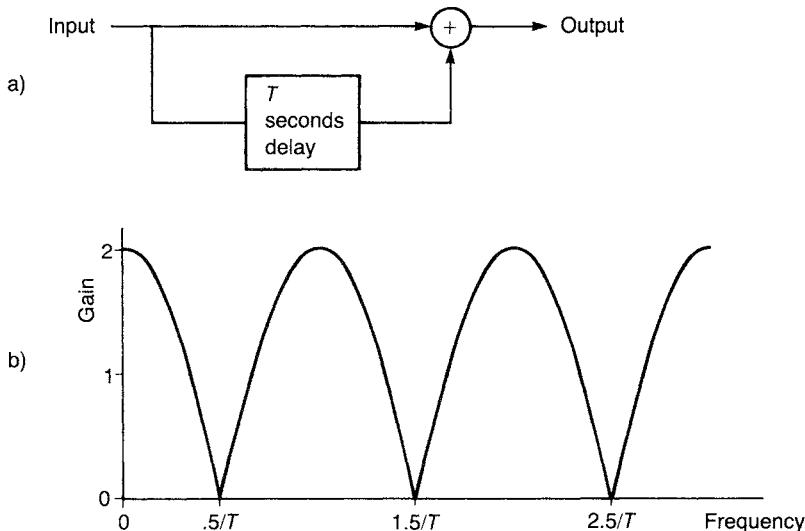


Figure 4.15    *A delayed signal added to itself (a) results in a filtered frequency response (b).*

In the same way, digital filters use summed delays to produce phase cancellation in the signal and consequently a filtered frequency response.

Although the notion of filtering a signal with a delay line might seem unusual, the phenomenon is common in the acoustical world. For example, a sound traveling to a microphone can take a direct path or a reflected (delayed) path (such as a bounce off the floor). The summed difference in path lengths creates a comb filter response.

In a digital filter, delay lines composed of shift registers are used to create time differences in the signal. Furthermore, multiple delays are taken from the line. For example, delays of 3, 4, and 5 microseconds will yield notches at odd multiples of 166, 125, and 100 kHz. In addition, the notches will overlap, attenuating all frequencies in that band. In the case of a CD player's digital filter, the delays are greater in number and specifically calculated to attenuate image spectra above the audio band.

### *Transversal Filters*

Remembering that each sample point of the filtered output signal is the sum of many filtered samples, we can demonstrate how a digital filter works. We may employ a shift register with output taps after each delay element, as shown in figure 4.16. Input samples enter the shift register. The output of each tap is multiplied by a coefficient associated with the

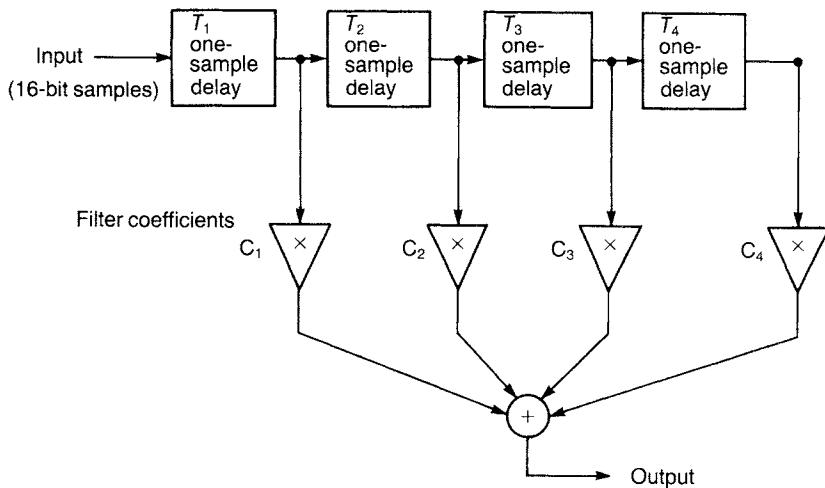


Figure 4.16 A transversal filter uses a shift register, coefficient multipliers, and an adder to produce a filtered output.

impulse response, and the product is summed with other products to yield the new output sample. The coefficients are taken from the  $\sin(x)/x$  curve and thus impose that characteristic on the audio data when it is multiplied by the coefficients. The sum of many such multiplied samples results in low-pass filtered data. To access many such samples, each time a new sample is input, the samples in the shift register are shifted one delay to the right, and the new sum of products is recalculated. Because of the movement of the samples across the shift register, this design is often called a transversal filter.

A step-by-step example may illustrate the filter's operation. First (referring again to figure 4.16), assume that data has already shifted through the filter. We observe that the output sample is equal to the  $T_1$  input sample times the  $C_1$  coefficient, plus the  $T_2$  input sample times the  $C_2$  coefficient, plus the  $T_3$  input sample times the  $C_3$  coefficient, plus the  $T_4$  input sample times the  $C_4$  coefficient. When the next new sample is entered, the previous samples are shifted one delay to the right to make room for it, and the calculation is repeated. Thus, in this example, each sample shifted through the filter requires four multiplications and four additions.

In practice, a number of considerations determine the design of implemented filters. The analog signal is the sum of the  $\sin(x)/x$  waveforms resulting from each sample. The  $\sin(x)/x$  waveform extends to infinity in both the positive and negative directions, so theoretically all of the values of that infinite waveform would be required to reconstruct the analog

signal. (In our example above, only four coefficients were used.) Fortunately, we can find a point on the waveform where neglecting any further response results in an error less than the system error. We can disregard the rest of the response values. Such a design is called a finite impulse response (FIR) filter.

As noted, the number of samples must be increased through oversampling. To achieve oversampling, a transversal filter could, for example, be constructed so that each delay is now one-fourth sample period. This would result in four times the number of output samples, with the filter calculating three new sample values for each input sample. To provide enough impulse response values to permit a good response after summation, perhaps ninety-six delay elements would be required. It would input new data four times each sample period. Every fourth sample would be an audio sample from the disc, and the other three would be zero intermediate samples. Thus only twenty-four of the ninety-six elements would be filled with non-zero samples at any time. The filter would output data at a rate of 176.4 kHz, each new sample the sum of twenty-four non-zero multiplications. The filter thus calculates three new sample values at the location of the zero input samples. However, this oversampling design is inefficient; the same result can be obtained with fewer delay elements.

A more cost-effective approach is shown in figure 4.17. It shows the architecture of a practical four-times oversampling digital filter, generating three intermediate samples between each input sample. The filter consists of a shift register of twenty-four delay elements, each delaying a 16-bit sample for one sampling period. Thus each sample remains in each

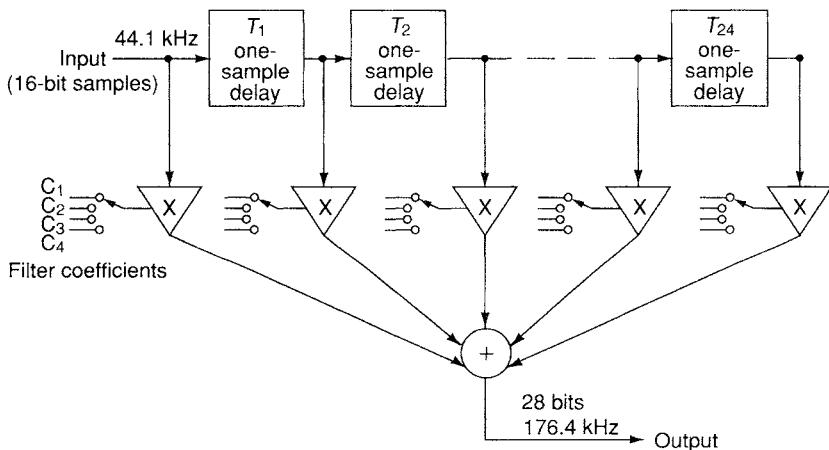


Figure 4.17 A practical four-times oversampling transversal filter may use twenty-four delay elements and four sets of coefficients.

element for a sample period. During this time each 16-bit sample is multiplied four times by a 12-bit coefficient stored in ROM, a different coefficient for each multiplication. In total, the four sets of coefficients are applied to the samples in turn, thus producing four output values. The twenty-four multiplication products are summed four times during each period and are output from the filter. The data is shifted one place, and the process is repeated. After multiplication and summation, the weighted average of a large number of samples is generated. Four times as many samples are present after oversampling, with new intermediate values calculated by the filter. The sampling frequency is increased four times, to 176.4 kHz.

Readers still uncertain as to the workings of digital filters, or suspicious of their ability to simulate the effect of analog filters in terms of waveform reconstruction, are directed to figure 4.18, a graphical summary of oversampling. It shows how intermediate samples are computed in a four-times oversampling filter. The input samples  $I_3$ ,  $I_2$ ,  $I_1$ , and  $I_0$  are treated as  $\sin(x)/x$  impulses placed relative to the center of the filter. Their  $\sin(x)/x$  impulse response amplitudes are equal to the original sample amplitudes, and the width of the impulse responses are determined by the response of the filter, in this case a filter with a cutoff frequency at the Nyquist frequency of the input samples. The summation of their unique contributions forms the interpolated samples (in practice, as noted, many more than four samples would be present in the filter). Each of the three interpolated samples is formed by adding together the four products, as shown in the figure. Original disc samples pass through the filter unchanged by using one set of filter coefficients that contains three zero coefficients and one unity coefficient. In this way, the output sample that coincides with the input sample is unchanged. By multiplying each group of samples (four in this case) with this coefficient set and three others, interpolated samples are output at a four-times rate.

The filter's coefficients produce a transition region between 20 and 24.3 kHz and again around 154 kHz. Figure 4.19 shows the effect of the filtering. The values of the intermediate samples, obtained by the calculation process, determine the filter characteristic; the bands centered at 44.1, 88.2, and 132.3 kHz have been largely suppressed.

Following the multiplications occurring during oversampling, the word length is much longer than the original 16 bits. Oversampling reduces the audio-band quantization noise by distributing it over a wider bandwidth, four times the bandwidth in the case of four-times oversampling. However, the output word cannot simply be truncated to 16 bits; this could increase distortion. Sometimes employed is a special rounding-off mechanism known as noise shaping. Quantization noise and the distortion from truncating are distributed over the entire oversampling spectrum. Thus the distortion in the audio band is a fraction of the total. With noise shap-

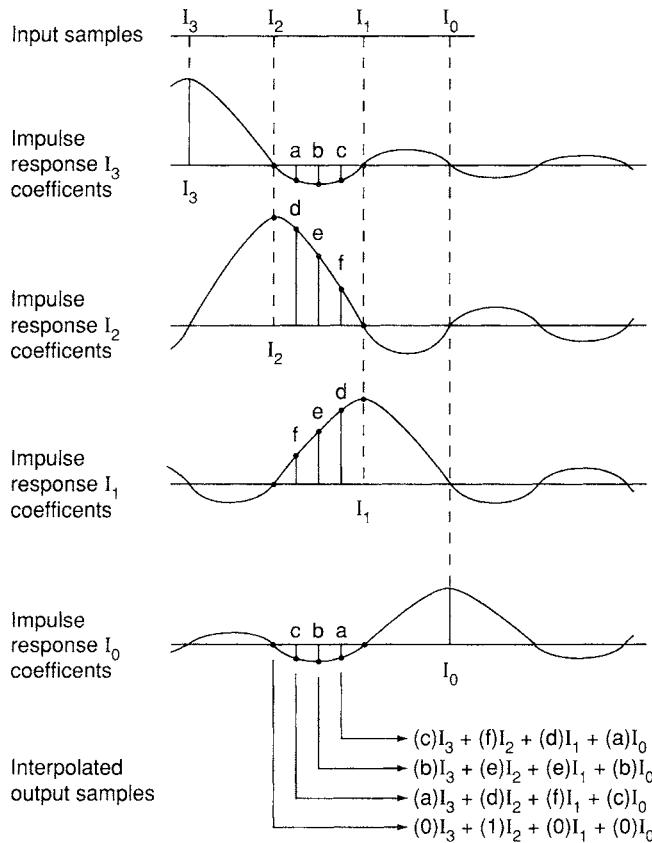


Figure 4.18 A four-times oversampling filter treats input samples as points on a  $\sin(x)/x$  curve, thus synthesizing the effect of an ideal analog low-pass filter. In practice, many more input samples are used.

ing, the dynamic range of the system can be extended. Noise shaping is discussed in more detail later in this chapter.

## DIGITAL-TO-ANALOG CONVERTERS

All compact disc players have digital-to-analog converters, otherwise known as D/As or DACs. Increasingly, however, CD players optionally output their digital information directly to preamplifiers and power amplifiers with their own D/A converters. Ultimately, all-digital stereo systems will appear. The analog signal that our ears require will not appear until the

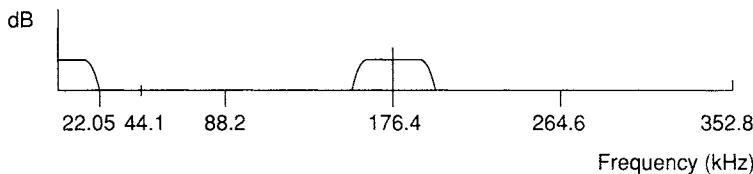


Figure 4.19 After four-times oversampling, an image spectrum remains at 176.4 kHz.

last possible moment, at the loudspeaker. Until then, players contain D/As to convert the digital bit stream to an analog signal, so that analog preamplifiers and power amplifiers can process it.

A D/A must produce an analog voltage corresponding to each digital audio sample. The sum total of these sample voltages creates an analog waveform—actually, one waveform for each stereo channel. The signal output from the D/A (following the output sample-and-hold) is thus equivalent to the signal appearing after the sample-and-hold circuit in the recording side of the chain. There are several excellent ways to accomplish this task.

### Ladder Network D/A Converter

D/As have traditionally used resistor ladder networks, such as the one in figure 4.20. The reference voltage source  $V_{ref}$  produces a voltage along the ladder, which is converted to a current by the resistors. Each successive rung of the ladder taps off part of the current: That is, each current is a binary power of two less than the preceding current. This corresponds to the values of the bits of the digital word. The bits of the word are used to open and close switches, turning their appropriate currents on and off.

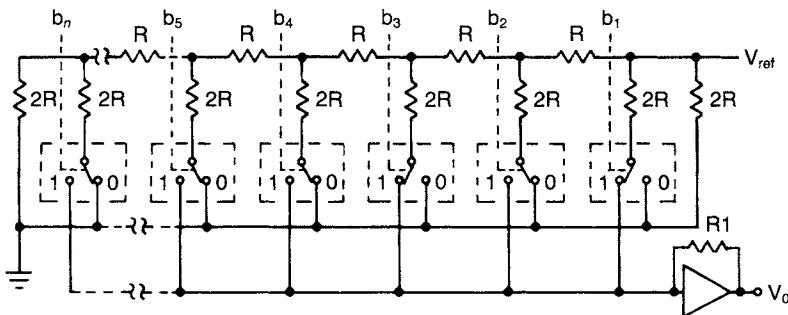


Figure 4.20 Circuit diagram of an R-2R ladder network D/A converter.

For example, a word with all 0s would keep the switch open, and no current would flow. A word with all 1s would close all the switches, and maximum current would flow. When the output current is converted to a voltage, the result is a voltage from the circuit which corresponds to the input binary word. Our conversion is thereby accomplished.

### ***Integrating D/A Converter***

Other circuit designers have devised D/As that differ significantly in design but accomplish the same purpose. For instance, an integrating D/A uses a constant current that flows for a period of time controlled by the digital input. Figure 4.21a illustrates the concept behind this design: A current source produces a changing voltage by means of an integrator. The input word determines how long the current is applied to the integrator, thus establishing the value of the output voltage. A small digital value allows a small amount of time for current flow, whereas a larger value would allow a long time for current flow. By measuring the amount of current that has flowed and converting that current to a voltage, we receive a voltage that corresponds to the input word.

In practice, a single current source and timer are not used, because the timer cannot operate quickly enough to achieve high resolution. To achieve conversion speed without sacrificing accuracy, two current sources are used, as shown in figure 4.21b. This design is called a dual slope integrating D/A converter. The two current sources have a precise ratio of 256:1; the larger current, for every time period it flows, is equivalent to the smaller source operating over 256 periods. The two current sources are thus assigned to the eight least significant bits and the eight most significant bits of the input word respectively. The high-order bits first generate a coarse ramp output. Then the low-order bits take over to generate an accurate ramp output. The result is a highly accurate output voltage, quickly achieved.

### ***Dynamic Element Matching D/A Converter***

Another D/A design, called dynamic element matching (DEM), uses the principle of current adding. This is illustrated in figure 4.22a. In practice, as shown in figure 4.22b, an accurate current source is obtained by dividing a reference current source with a pair of resistors, then by averaging their outputs with a pair of switches and capacitors. By cascading these or other types of current-dividing stages, a binary weighted series of currents can be obtained. Their selected summation accomplishes the task of conversion. Although a DEM converter may require more

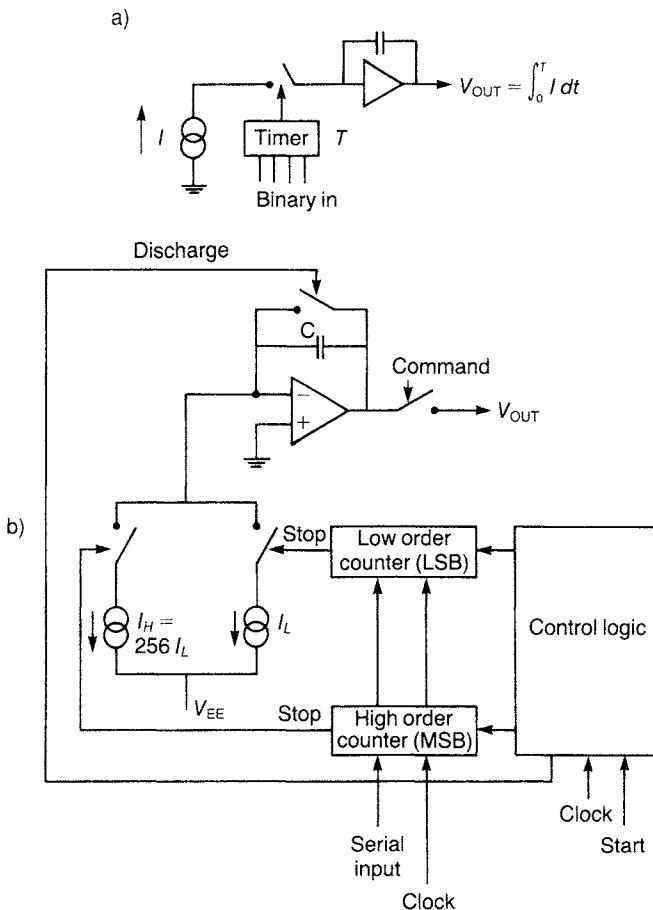


Figure 4.21 An integrating D/A converter (a) uses a current source, integrator, and timer to produce an output voltage. A dual slope integrator (b) separates low- and high-order bits and uses two current sources to produce a segmented output voltage.

external parts than other designs, it is cost-efficient because it requires no calibration.

Whichever method is employed, the D/A converters used in most CD players do not pose insurmountable design problems. Contemporary players utilize one or two converters, typically converting a 16-bit word. In some designs, 18- or 20-bit D/A converters are employed; the longer word lengths present following oversampling permits this technique, which yields an increase in the system's S/N ratio.

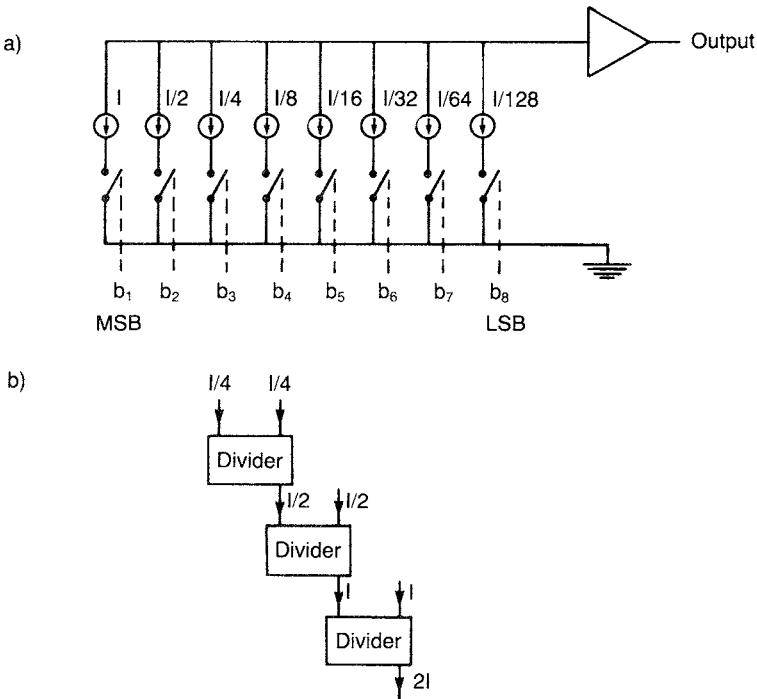


Figure 4.22 A dynamic element matching D/A converter uses current adding (a). The current sources are obtained from a current source and current dividers (b).

### D/A Converter Resolution and Linearity

The number of bits that a converter processes is not a good means of determining quality. In reality, the accuracy of the converted bits is far more important. For the most part, 16-bit D/A converters are being superseded by converters that provide 18 or even 20 bits of resolution. In fact, such resolution is considered to be *de rigueur* in even medium-quality players.

The choice of storing 16-bit words as the CD standard was determined primarily by the availability of linear 16-bit D/A converters. However, new conversion methods offer a chance to improve conversion quality. While no D/A technology can yield more than 16 bits of fidelity from a 16-bit medium, the use of new conversion methods can provide more linear conversion of those 16 bits. When correctly done, alternative conversion can improve the amplitude resolution of the player by ensuring an accurate conversion of the medium's 16-bit signal, yielding faithful reproduction of the recorded signal. Looked at in another way, it may be said that in practice linear 16-bit D/A conversion is insufficient for 16-bit

storage or processing. The reason for this lies in flaws inherent in D/A converters. Except in theory, 16-bit converters cannot fully decode a 16-bit signal without a degree of error.

A D/A linearity test measures a converter's ability to reproduce analog signals, particularly low-amplitude signals. Ideally, for example, a -80 dB tone on a disc would be reproduced at -80 dB. However, a nonlinearity of +2 dB at this level would result in a tone reproduced at -78 dB. A deviation such as this manifests itself as distortion in the audio signal. Results of linearity tests are shown in Chapter Five.

Converter nonlinearity can be traced to a number of causes, but the most critical factor is the accuracy of the converter's most significant bit (MSB). It accounts for a change in fully one-half of the analog signal's amplitude, while in comparison the least significant bit (LSB) in a 16-bit converter accounts for an amplitude of one part in 65,536. Many D/A converters thus provide a means to calibrate the MSB, and in some cases several of the most significant bits can be calibrated. This permits the CD player manufacturer to compensate for differences in chips. Because these most significant bits steer the converter's output, proper calibration of individual converters greatly improves low-level linearity, and accuracy. This calibration must be done with great care. For example, when trimming the MSB of a 16-bit converter, an error of 0.01 percent would negate the contribution of the two least significant bits. An error of 0.1 percent in the MSB would offset the combined value of the five least significant bits.

Unfortunately, many audio critics have found cases of calibration trim pots consistently set in the middle of their adjustment range—a very suspicious alignment. Even if these adjustments are accurately performed, calibration is often lost over time as the player is exposed to changes in temperature, humidity, and simple aging. Given the accuracy needed in the calibration of the MSBs, this is not surprising.

One solution to the limitations of 16-bit conversion is a D/A converter with more bits' worth of conversion. An 18-bit D/A converter has 262,144 levels, and a 20-bit converter has 1,048,576 levels, sixteen times as many output levels as a 16-bit converter. Any nonlinearity would be correspondingly smaller, and increasing the quantization word length at the conversion stage results in an increase in S/N ratio. Simultaneously, any quantization artifacts are diminished. In other words, a higher-bit converter would give a better 16-bit conversion. In fact, the extra bits of a linear high-bit converter would not even have to be connected to yield improved 16-bit performance.

The intent of 18- or 20-bit D/A conversion technology is similar to that of oversampling: While the sampling rate is increased, the method doesn't create new information. It merely attempts to make better use of existing information. Not coincidentally, it is oversampling that makes high-bit

conversion of 16-bit data possible. It solves the obvious dilemma of coming up with the extra bits, when the output from the medium is only 16 bits. When a 44.1 kHz, 16-bit signal is oversampled, both the sampling frequency and number of bits are increased, the former because of oversampling and the latter because of the multiplication which must take place. For example, the output of an oversampling filter may be 176.4 kHz and 28 bits. Nominally, only the 16 most significant bits are used for conversion through a 16-bit D/A converter, and the rest are discarded or used for noise shaping. (The latter is a particularly clever idea, which we'll return to later in this chapter.)

The use of extra bits for output conversion requires either a longer word length D/A converter or a switched 16-bit converter. In the latter design, the parallel bits from the oversampling filter's output are wired through switches to the inputs of a 16-bit D/A converter. An example of an 18-bit shifted bit D/A converter system is shown in figure 4.23. When all 16 bits are used to convey a signal (a high-amplitude signal), the upper 16 bits are applied to the 16-bit converter. However, when the two upper bits from the oversampling filter are not being used to convey a signal (a lower-amplitude signal) the 18 bits are shifted downward, the unused bits are ignored, and the 16 lower bits are utilized instead. This adaptive scheme makes sense because in music recording the 2 upper bits are rarely used, and then often for a brief period of time. Through bit shifting, a 16-bit converter may thus handle an 18-bit input.

However, because we have shifted the word by 2 bits, the signal's amplitude changes. In the binary system, a shift of one place (one bit) results in a doubling of value. A shift of two places quadruples the value. Thus when the lower bits are shifted up, the amplitude is four times too large, thus the output must be attenuated by one-quarter. An analog gain block downstream of the D/A handles this chore.

This adaptive 18-bit conversion scheme may be considered a noise reduction scheme because the signal is expanded at the D/A converter. The benefits result from the fact that the residual noise of the converter as well as its conversion nonlinearities will be proportionally reduced. Looked at another way, a four-times (12 dB) higher analog output has been achieved without increasing the D/A's residual noise and conversion error. When the gain is reduced by one-quarter to bring things back to normal, the noise and conversion errors are reduced by one-quarter. There is an increase in S/N of 12 dB and distortion reduction by one-quarter, both equivalent to using an 18-bit converter.

Of course, as with any clever scheme there is a price to be paid for the benefit accrued. Specifically, when the bits are shifted, it is difficult to immediately and simultaneously shift the gain of the analog output to compensate. Furthermore, any static offset introduced by component tolerances in the attenuator could become apparent when the switching takes

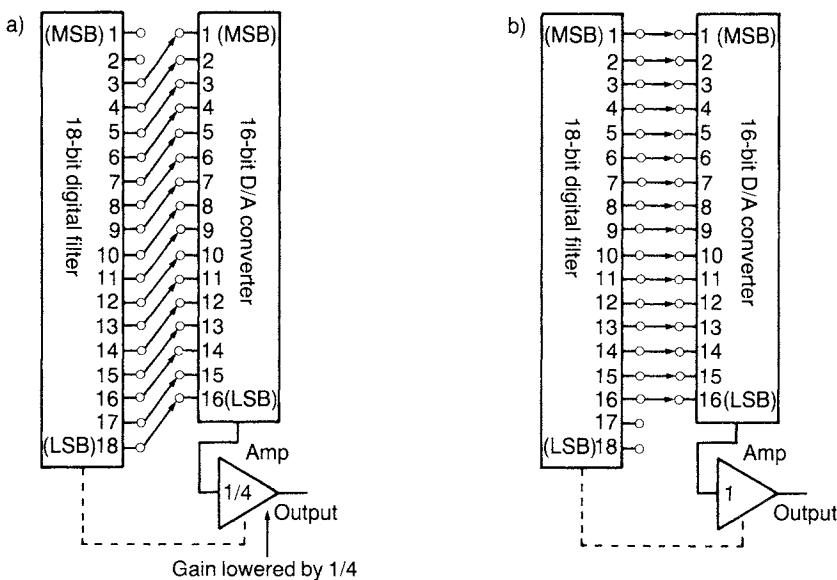


Figure 4.23 Eighteen-bit performance can be achieved with a 16-bit D/A converter by using two lower bits from the oversampling filter during low audio levels (a). The bits are shifted when high levels occur (b). (Courtesy Yamaha)

place. It would not be significant over the first 14 bits, affecting only the attenuation ratio. But when switched off (or later, on), the difference between the attenuator and the absolute value of the output could create a glitch in the waveform where the attenuated and nonattenuated signals are joined. The glitch would be present at  $-12$  dB regardless of signal frequency. In practice, the increase in S/N ratio far outweighs the disadvantages of switching nonlinearity.

Recent work has resulted in linear 18-bit D/A converters, which are both very affordable and highly sophisticated. For example, one D/A converter contains a test circuit to monitor settling time of the current source. This minimizes glitches to the extent that resultant distortion is below measurable levels and can be analyzed only through computer modeling. Although a conventional binary-weighted architecture is employed in this particular converter, several special design considerations were instituted to attain the desired accuracy. The three most significant bits are made up of seven individual current sources; this reduces thermal errors. Bits 4 through 16 are made up of unit-valued current sources which feed the R-2R ladder network. Currents for bits 17 and 18 are derived from the unit-valued source. The relative gain of the 3 upper bits can be

adjusted against the total weight of the 15 lower R-2R bits by trimming the scale-down network.

The chip itself is a 40-pin hybrid with a special divided layout to provide full control over laser trimming of the upper and lower bits. Supporting circuitry is located outside the chip to help maintain the chip's thermal balance, a critical concern in D/A design. Accuracy to the least significant bit requires precision equivalent to four parts per million.

The architecture of a linear 18-bit D/A design is markedly simpler than a pseudo 18-bit method. The use of 18-bit D/A chips results in player specifications clearly superior to 16-bit designs. As described above, an 18-bit converter has four times greater resolution than a 16-bit converter. In other words, 18-bit conversion of a 16-bit signal provides for a 12 dB increase in S/N ratio while processing the data. This improvement is clearly measurable with 18-bit conversion. Of course, there is no risk of gain matching errors as with pseudo 18-bit conversion methods.

## OUTPUT SAMPLE-AND-HOLD

The sample-and-hold (S/H) circuit following the D/A, which is sometimes contained on the D/A chip, performs a hold function to create a pulse amplitude staircase. The impulses in the output staircase waveform are the width of a sampling period, but reconstruction theoretically requires impulses of infinitely short duration. This is impossible to achieve because it would require, among other things, infinitely large current flow. Because of the finite duration of the output samples, a filtering effect occurs in which the amplitude response (a  $\sin(x)/x$  curve) gradually falls to zero at the sampling frequency, which can be seen in figure 4.24. This is beneficial because image spectra are attenuated. However, the filtering results in a slight loss at the Nyquist frequency. This is sometimes called aperture error. To minimize this loss, the impulses can be narrowed closer to their original width. Such shorter duration impulses result in a frequency characteristic closer to the originally flat response for the pulse amplitude signal. Figure 4.25a shows the sample pulses narrowed by four

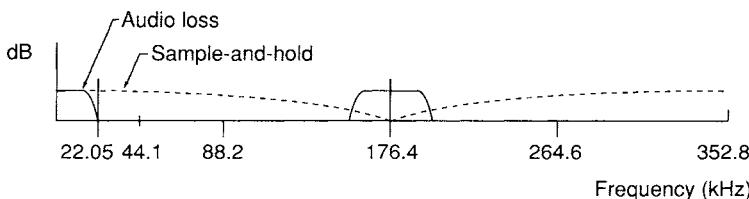


Figure 4.24     *Aperture error results in a slight attenuation of high audio frequencies. This occurs even in the four-times oversampling system shown here.*

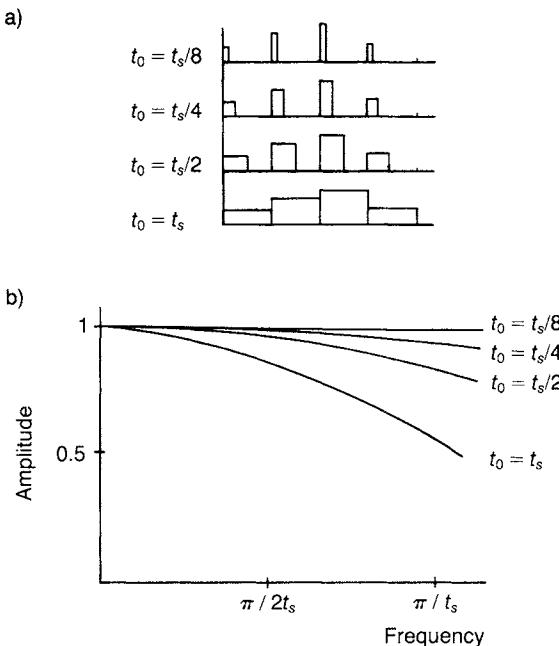


Figure 4.25 *Shorter duration output samples minimize aperture error, resulting in a flat frequency response.*

amounts. The aperture time  $A$  is defined by  $A = t_0/t_s$  where  $t_s$  is the sample time and  $t_0$  is the duration of the pulse. Figure 4.25b shows the resulting frequency roll-off due to aperture effects. The digital filter design can compensate for any slight remaining high-frequency attenuation with an inverse, or  $x/\sin(x)$  characteristic.

In addition to its role in correcting aperture error, the sample-and-hold circuit is also useful in removing glitches from the staircase waveform. The D/A converter output is not a smooth transition from one analog level to the next. Instead, inaccurate transition voltages known as glitches are often generated. Although they are momentary and disappear when the D/A's output value settles to its correct voltage, these glitches would increase distortion in the output signal. A sample-and-hold circuit can be used to gate them out. This circuit is timed to wait for them to settle before it samples the D/A output. In this way, only accurate values are propagated. Other types of de-glitcher circuits have also been devised to solve this problem. For example, a shunting circuit could switch the output to ground during glitching transitions.

Following the S/H circuit, some image spectra energy will still remain. This must be suppressed by an output analog filter. This anti-imaging filter

is tame compared to brickwall filters. Since the remaining band is so high in frequency, we can use a filter with a gentle, 12 dB/octave response and a  $-3$  dB point between 30 and 40 kHz. It is a noncritical design, and its low order guarantees good phase linearity; phase distortion can be reduced to  $\pm 0.5$  degrees across the audio band.

## A CHIP SET

When circuitry is reduced to an integrated circuit, several interesting things happen. First, everything's a lot smaller. Instead of a king-size circuit board populated with several thousand transistors and other discrete components, we have something considerably more compact. Furthermore, power consumption is dramatically reduced, decreasing the size of the needed power supplies or batteries.

Cost is the bottom line and the most interesting asset of integration. While the development cost of a new chip can be very expensive, the manufacturing costs are quite low. If these costs are amortized over the market life of the chip and divided over the number of chips sold, the cost per chip can be about as small as the microcircuitry itself. In addition, labor costs, packaging costs, and shipping costs of the finished product are all lower. The result is a product that is smaller, lighter, less sensitive to environmental changes, less power hungry, more reliable, and cheaper—all thanks to chips.

Most of the signal processing in a compact disc player is handled by chips. In fact, CD players are prime examples of the benefits of digital microelectronics. A CD player without chips would be enormously large, power-hungry, costly, and wouldn't sound very good. Its hundreds of circuit boards would be extremely unstable, particularly with respect to temperature changes. Even with a full-time engineer crawling around inside to adjust potentiometers, fluctuations would make a joke of operating tolerances.

### *A Philips Chip Set*

For these and other reasons, manufacturers are anxious to reduce CD player circuitry to the fewest possible number of chips. The decrease in the number of parts marks the evolution of player design. For example, Philips, coinventor of the CD system, has achieved a high degree of integration. Its second-generation chip set is shown in figure 4.26. The SAA7210 handles demodulation, error correction, and basic interpolation. The SAA7220 contains circuitry for enhanced interpolation and a phase-linear digital FIR (finite impulse response) filter. The TDA1541 is a dual 16-bit D/A converter chip.

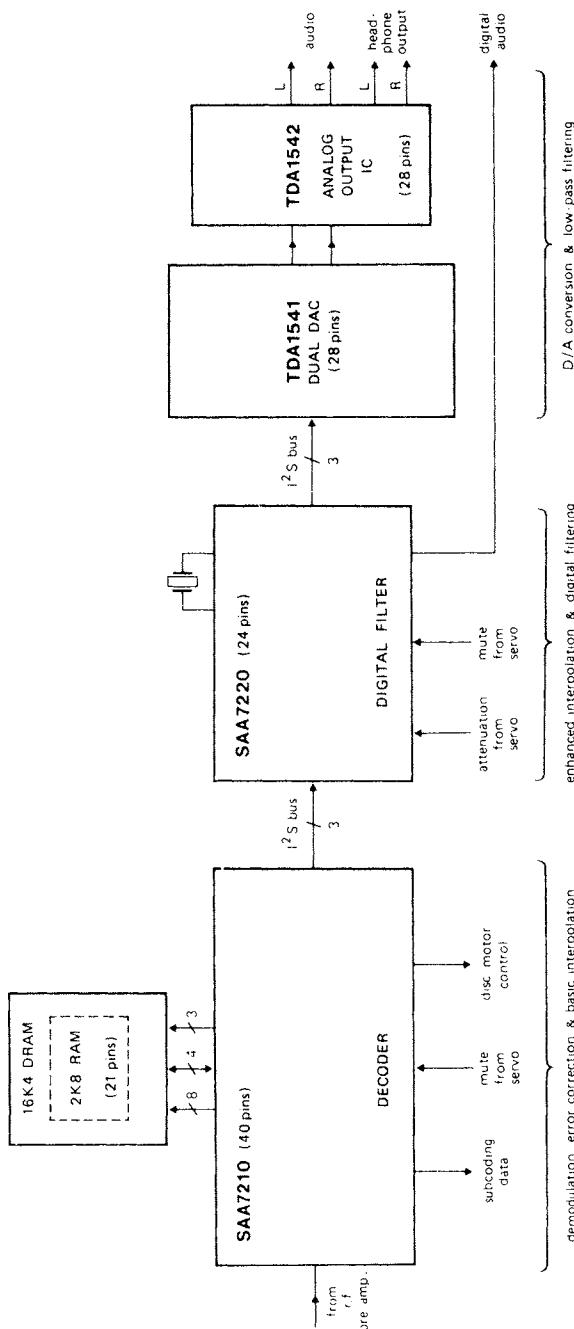


Figure 4.26 Philips chip set with digital filter and dual channel 16-bit D/A converter. (Courtesy Philips)

For design flexibility, the I<sup>2</sup>S (inter-IC signal) format is used between the SAA7210, SAA7220, and TDA1541 chips. It provides a three-line bus comprising serial data line, clock, and left/right channel select line. Using this bus, the SAA7210 and TDA1541 could be connected directly, omitting the SAA7220.

### *Features of SAA7210 Decoder*

Of course, the time and trouble required to compress circuitry into a few chips provides the opportunity for greater performance. In this case, these three decoding chips offer significant improvements over first-generation chips. The SAA7210 contains a demodulator with adaptive data slicer, a fully integrated phase-locked loop demodulator for bit-clock generation, and a motor speed control to stabilize the input data rate.

Part of the SAA7210's job is to process subcode information from the disc—the nonaudio data representing track numbers, playing times, titles, and so forth. In some designs the control processor wastes a lot of time working with the subcode. The SAA7210 continually collects the subcode, sending a full frame of subcode only when the processor requests it. If the processor requires only a portion of the subcode—for instance, only a track number (contained in the first 16 bits of the subcode block)—it can prematurely reset the SAA7210, returning it to subcode collecting.

The SAA7210 decoder also handles error correction chores, using CIRC. The CIRC is standard for all compact discs, but decoding strategies vary considerably. Prior to the SAA7210, no error correction chip made full use of the CIRC's correcting capabilities. The SAA7210 can make four erasure corrections (the position of the error is known) per 32-symbol block and up to four corrections when the position of the errors is not known. In some other chip sets, only two erasure corrections are possible.

Furthermore, unlike other decoders, the SAA7210 uses adaptive error correction to discriminate between the types of errors found on a disc. This makes better use of the CIRC correction capability, permitting correction of longer burst errors and more reliable corrections. All decoders use error flags to pinpoint troubled data, but the SAA7210 uses extra flags. They are generated by the EFM decoder to mark symbols that violate the 3T–11T criterion.

In designing the chip, laboratory testing of defective discs resulted in the identification of types of defects and the development of three 2-bit error flags: a hard error flag (the most reliable flag), a medium error flag, and a soft error flag (the least reliable flag). A no-error flag identifies valid data. The optimum error-correction strategy is selected from one of 60 possibilities by a flag processor, as shown in figure 4.27, depending on the type (soft, medium, or hard) and number of flags caused by disc defects.

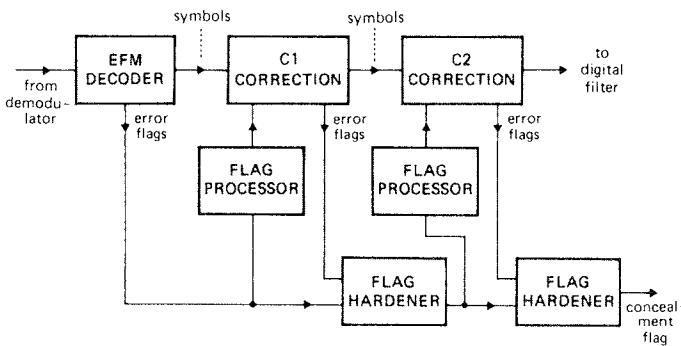


Figure 4.27 Flag processor for selecting error correction algorithms. (Courtesy Philips)

After correction, more flags are generated and compared to incoming flags to fine-tune the strategy by updating and hardening the flags. Progressive flag-hardening improves the reliability of the error correction, so that very reliable corrections can be made. The various strategies are programmed in the SAA7210's ROM and can be altered for nonaudio CD applications which might call for different strategies.

The SAA7210 must also communicate with RAM; a FIFO (first in, first out) memory is needed to de-interleave data and to buffer data rate irregularities. The SAA7210 is designed to operate with a  $4 \times 16$  kbytes RAM which accommodates 64 frames of data. Among other benefits, the large FIFO helps smooth data disruptions occurring in car and portable players. The SAA7210 delivers its data to the SAA7220 digital filter.

#### *Features of the SAA7220 Filter*

The SAA7220 is a dual-channel linear-phase FIR digital filter chip. The SAA7220 chip employs a four-times oversampling design. The number of filter coefficients has been increased over the first-generation filter by 24 per channel to a total of 120. Pass-band ripple has been reduced to 0.02 dB to reduce the possibility of that artifact's audibility. (Experiments have shown that pass-band ripple of  $\pm 0.2$  dB can be heard.) In addition, the SAA7220 mutes the signal smoothly so that no clicks or pops occur when the player starts, stops, or pauses. Uncorrected data errors stemming from burst errors are masked with an 8-sample interpolator on-board the SAA7220 to create a smooth bridging waveform over the bad samples, as shown in figure 4.28. Improved interpolation and error correction result in significantly better tolerance to burst errors compared to earlier chip sets. The

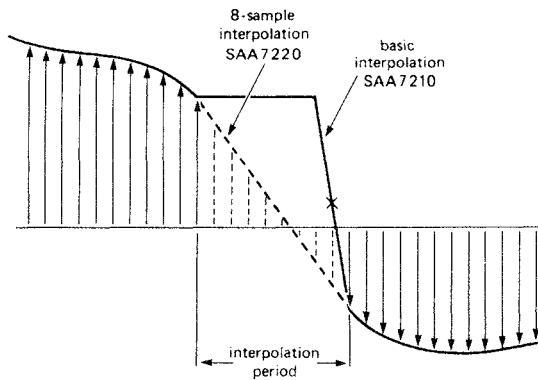


Figure 4.28 An 8-sample interpolation on the SAA7220 is used for error concealment. (Courtesy Philips)

SAA7220 can deliver a digital-format audio output for interfacing to out-board digital signal processors.

#### *Features of the TDA1541 D/A Converter*

The TDA1541 employs a dynamic element matching design technique. Moreover, the TDA1541 is a dual 16-bit D/A chip. Since the TDA1541 offers two D/As on one chip, there is no chance of mismatched performance or thermal drift as with single D/A chips. Internal bit switching is performed with a diode transistor configuration for fast and accurate switching. When the filter's coefficients were selected, the designers compensated for both the  $\sin(x)/x$  response of the hold function on the D/A converter and the analog output Bessel filter's response. The combined conversion function is thus flat in the audio band.

The TDA1541 also permits use of digital signal processing, such as digital tone controls. Such processing requires increasing the dynamic range of the output signal when select frequencies are boosted. For example, a digital tone control using the TDA1541 could boost a frequency band by 12 dB while maintaining a 16-bit dynamic range over the audio band. When the 16-bit TDA1541 is coupled to noise shaping circuits to shift quantization noise out of the audio band, resolution equivalent to 18 bits can be achieved.

#### *Output Stage*

The audio output is handled by the TDA1542 dual channel analog filter chip. As we have seen, digital filtering shifts frequency components away

from the audio band, but they reappear again at higher frequencies. The TDA1542 removes those artifacts with an active third-order filter. Choice of external components determines the filter type (Bessel, Cauer, and so forth). In Philips players, a Bessel filter is used.

These second-generation chips offer significant improvements over first-generation chip sets. These improvements result in improvements in player cost, size, and power consumption, while improving reliability and audio performance.

## LOW-BIT D/A CONVERSION

The advent of multibit 18-bit D/A conversion technology set a new standard for digital audio conversion. In many cases, the primary sonic limitation in CD reproduction using linear 18-bit D/A converters is the quality of the equipment used to make the master recording itself.

On the other hand, although 18-bit converters can reduce nonlinearity error below levels mandated by 16-bit converters, the technology is not perfect. For example, the reference voltages supplied to an 18-bit converter must be highly accurate and stable, the converter must be carefully calibrated, and extreme care must be taken to prevent drifting. The difficulty can be assessed by recalling that a 16-bit converter must be calibrated to an accuracy of one part in 65,536, whereas an 18-bit converter demands one part in 262,544. In other words, although it is more precise than a 16-bit converter, an 18-bit converter has all the real-world problems of a 16-bit converter, only more so.

One response to the limitations inherent in multibit converters lies in a clever conversion method which follows the philosophy that sometimes less is more. Instead of converting parallel digital words at the output sampling rate, shorter word lengths are converted at much faster rates. In the same way that oversampling removes the deficiencies of analog filtering by performing the operation in the digital domain, low-bit (sometimes called 1-bit) conversion methods help alleviate the deficiencies of parallel conversion—an analog process—by performing conversion in the digital domain using DSP. Today, many CD players use this third-generation converter technology.

Fundamentally, a multibit converter represents the analog waveform as an amplitude signal, processing information that measures the amplitude sample by sample. However, the method is flawed when quantization introduces errors in amplitude representation. Moreover, because a multiplicity of bits are used to form the representation, and because each bit has an error unequal to the others, the overall error varies with each sample, making compensation almost impossible.

By contrast, a low-bit system takes a radically different approach. It employs only a small number of bits (perhaps 1 bit) at a very high rate to represent the signal. Instead of an amplitude representation, a low-bit system uses time division to represent the signal. When this representation is low-pass filtered, the result is an analog waveform, just as in the multibit method. Because only a few bits are used, any amplitude error is essentially constant, and hence there is no relative error. In theory, neglecting jitter, an error would produce only an offset and thus could easily be removed from the signal. Because time division is inherently more accurate than an amplitude representation, low-bit converters are inherently more accurate than multibit converters.

The way in which a 1-bit (high- or low-level) pulse signal can accurately reconstruct an audio signal is not intuitively obvious. A simple analogy helps illustrate the method: If a light switch is on, the room is brightly lit, and if the switch is off, the room is dark. But if the switch is turned on and off quite rapidly, any intermediate light intensity can be created. Similarly, a 1-bit D/A converter can reproduce an audio waveform. All 1s would be a full positive signal, all 0s would be a full negative signal, evenly alternating 1s and 0s would be a zero-level signal, and other variations would create intermediate levels. Figure 4.29 shows how a single bit, with either a high or low level, can represent, in this example, a smooth sinewave. In practice, as described below, noise shaping is required to reduce in-band noise, and in some designs several bits are used instead of one bit.

### ***Noise Shaping***

A principal objective of low-bit conversion is reduction in the word length required to represent a signal. However, the bit reduction required to form a low-bit signal from a 16-bit source greatly degrades the signal's dynamic range. For example, to accurately convert a maximum-amplitude 16-bit word, a 1-bit system would have to change polarity  $2^{16}$  times per conversion period; with a sampling rate of 44.1 kHz, this would demand a rate of approximately 2.9 GHz. This is too fast for current technology. As the rate is slowed to accommodate current hardware limitations, 16-bit performance is lost and noise levels increase to an intolerable level. For this reason, noise shaping is a prerequisite in any low-bit system to reduce noise in the audible spectrum and hence improve dynamic range.

A low-bit converter may be viewed as a wide-band, spectrally weighted noise source in which the noise is more benign than multibit converter nonlinearity error. As its name implies, noise shaping shifts noise away from the audio band, thus lowering audio-band noise. Figure 4.30 shows a 2 kHz test tone and the noise floor with low-bit conversion. Clearly, with noise shaping, noise is decreased in the audio band (up to  $2 \times 10^4$  in the

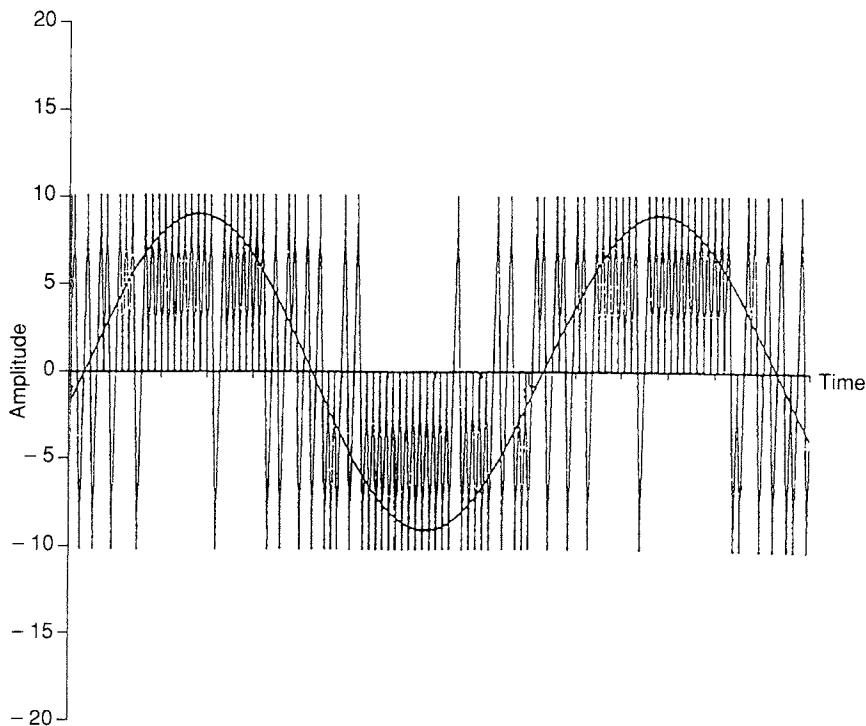


Figure 4.29 A computer simulation shows how a time-averaged 1-bit signal can represent an analog sine wave waveform. (Courtesy Philips)

figure), at the expense of boosting noise beyond the audio band. Because the noise beyond the audio band can be safely removed through filtering, with noise shaping, 16-bit performance (or much more) can be achieved with a single bit.

Noise shaping reduces quantization error by using the known characteristics of quantization error to modify the error and decrease its effect. Efficient noise-shaping algorithms use recursion to spread the noise power across a greater spectrum. In particular, the method places the error information generated in the bit reduction process back into the signal, much like negative feedback is used to reduce distortion in analog amplifiers. In effect, noise shaping attempts to cancel error by subtracting it from the input signal. Feedback in this discretely sampled system can operate only if a delay element one sample or more in duration is placed in the feedback loop. By placing the quantization error signal in a feedback loop, the frequency response of the error signal is altered, shifting much of its en-

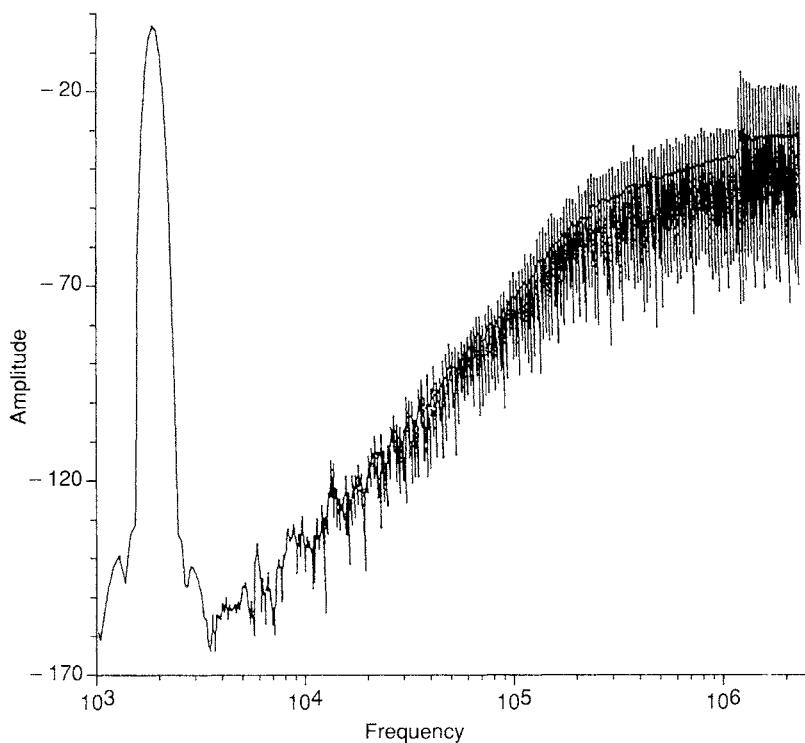


Figure 4.30 *Noise-shaping algorithms reduce in-band noise while increasing out-of-band noise. This computer simulation shows the low in-band noise floor with a 2 kHz test tone. (Courtesy Philips)*

ergy outside the audio band. Moreover, the noise is shaped by the approximate inverse of the loop transfer function; when a low-pass filter is placed in the loop, the noise spectrum rises with frequency. In other words, the noise-shaping loop low-pass filters the signal and high-pass filters the noise. If the system is designed so the signal's frequency content is less than the filter's cutoff frequency, the signal will not be affected. What's more, oversampling performed prior to noise shaping spreads quantization noise over a wider bandwidth; hence noise is further diminished in the audio band.

The real innovation associated with low-bit systems comes in the design of noise-shaping algorithms. Simply put, the more complex the algorithm, the lower the noise in the audio band. More specifically, higher orders of integration in noise shaping decrease the in-band noise level. Figure 4.31 shows four noise curves—for no noise shaping, first-order, second-order, third-order, and fourth-order noise shaping. Higher-order

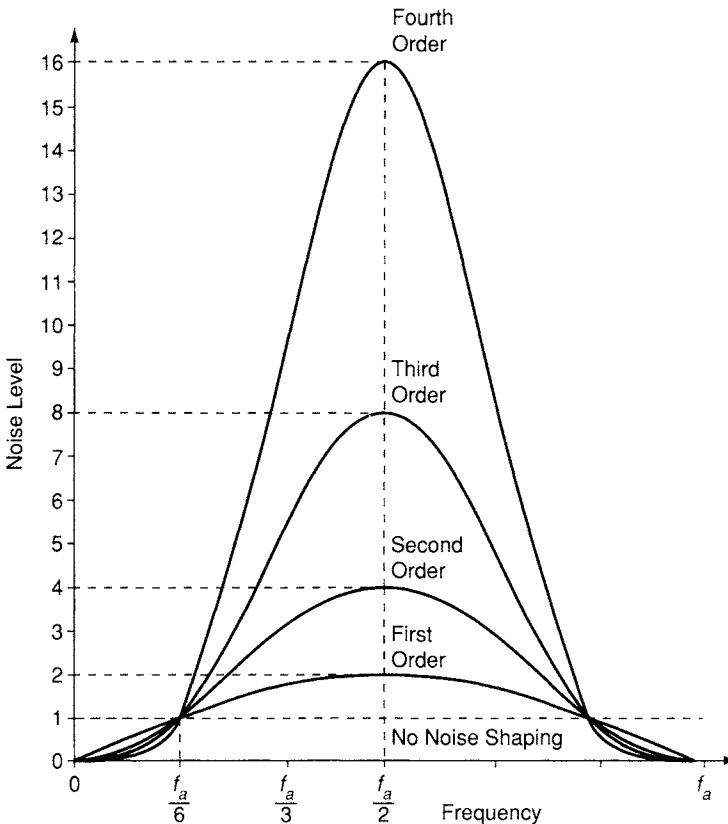


Figure 4.31 *Depending on the order of the noise shaping employed, in-band and out-of-band noise levels are relatively affected. (From Pohlmann 1991)*

noise shapers can be designed with cascaded digital differentiators. The greater the number of cascaded stages, the greater the slope of the shaping function; this in turn increases the suppression of audio-band noise. However, third- or higher-order noise shapers may suffer operational instability due to overloading of the integrators. Thus in practice, a multi-stage circuit is often employed, as described below.

As the order is increased, the noise-shaping characteristic changes, and in-band noise is decreased. However, out-of-band noise is greatly increased; this could overly burden output analog low-pass filters. The dynamic range can be optimized through a combination of oversampling rate and order of the noise-shaping filter. As the oversampling rate is increased, the portion of the curve in the audio band is relatively reduced, that is, moved toward 0 Hz. In other words, although the shape of the noise curve remains the same, high oversampling rates relatively decrease

in-band noise. A successful noise-shaping circuit thus seeks to balance a high oversampling rate with noise-shaping order to reduce in-band noise and shift it away from the audible range.

Noise-shaping processing can also result in the output of correlated patterns when the processor is idling with low-level audio signals. To overcome this, dither is added to the data so the circuit always operates with a changing signal even when the audio signal is zero or otherwise unchanging. A digital dither signal applied inside the shaping loop is identical to a high-pass filtered dither signal applied at a point outside the loop prior to the quantizer. In other words, the quantizer and the dither signal noise are both shaped by the loop.

### ***Low-Bit D/A Conversion with Second-Order Noise Shaping***

One implementation of a true 1-bit D/A conversion method is composed of three elements: oversampling filters, second-order noise shaping, and pulse density modulation output. This Bitstream converter was developed by Philips. The sampling rate is increased from 44.1 kHz to 11.2896 MHz—an increase of 256 times. At the same time, the 16-bit signal is converted to a 1-bit signal. It is this fast, 1-bit output that reproduces the audio waveform. The requantization error of the output signal is corrected by second-order noise shaping.

The output bit, operating at a rate of 11.2896 MHz, is converted to an analog signal using a simple switched capacitor network. The network's operation is very accurate, and hence the error of the signal is low. There are only positive and negative full-scale reference points; intermediate points are determined by time averaging. Any errors in the two points would produce an offset error but not a linearity error. The offset error can easily be removed.

Figure 4.32 shows the complete system. The first of the three oversampling stages uses a nonrecursive FIR interpolation filter to perform four-times oversampling to attenuate image spectra; in addition, first-order noise shaping is performed in the filter. The signal is downscaled to prevent clipping during any signal overshooting. The gain of the filter is selected to compensate for the effects of other stages in the conversion path. In particular, a high-frequency rise is used to compensate for the aperture error (-3 dB at 60 kHz) present in the output analog filter. The second stage performs 32-times oversampling through linear interpolation. A dither signal (-20 dB at 352 kHz) is added to prevent idle patterns from causing nonlinearity. Two-times oversampling is performed in the third stage through a sample-and-hold operation. This 17-bit signal (dither adds 1 bit to the original 16-bit signal) undergoes second-order noise shaping as described above, and a single bit is output from the quantizer. Finally, D/A conversion is accomplished at a 1-bit D/A converter via PDM

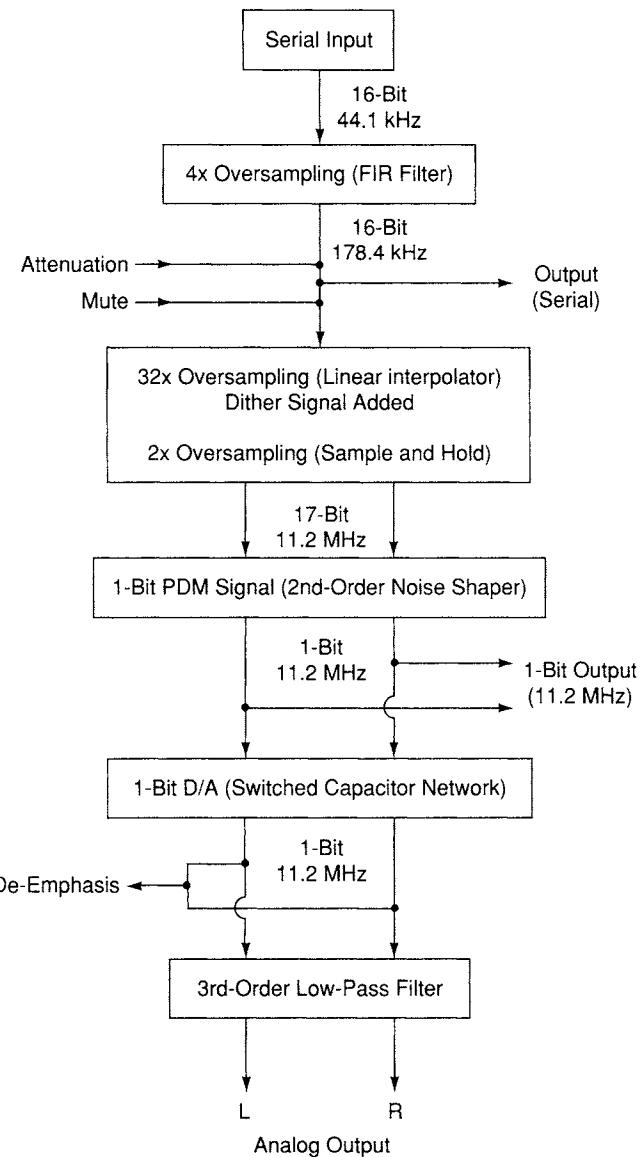


Figure 4.32 Block diagram of the signal processing in a 1-bit D/A converter employing second-order noise shaping. (Courtesy Philips)

(pulse density modulation), outputting 1-bit data at 256-times oversampling, or 11.2896 MHz. A third-order analog low-pass filter removes out-of-band high-frequency components.

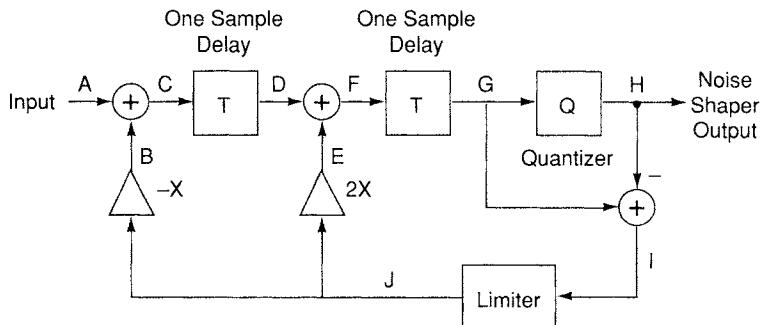
The noise shaping performed in this D/A converter is complicated (it is an example of DSP at its best), but we can get a feel for its workings through figure 4.33. This circuit performs second-order noise shaping through a feedback loop and generates a low-bit signal for conversion into an analog waveform. In particular, this noise shaper consists of two integration (filter) loops to reduce in-band quantization noise. The 1-bit code from the quantizer is the output sign bit; only one bit is output, with a logical 0 or 1 level. The remainder of each sample is fed back as a quantization error, after a limiting operation designed to prevent overflow. Second-order noise shaping is performed by adding double the error and the negative value of the error to the previous two samples.

Operation of the circuit can be explained by following the values of audio samples during several sample periods, at various points in the noise shaper. These values are shown in figure 4.33b. In this example, a maximum 0 dB value is represented by a 1 (in practice, to prevent overload, an input is scaled down in the digital filter). The successive inputs in this example (column A, going top to bottom) are a constant 0.5 (-6 dB); for correct operation, the output of the circuit should thus be a 1-bit output with an average value of 0.5. Arbitrarily, initial values for the contents of the delay registers in figure 4.33a are selected as  $D = 0.5$  and  $G = -0.6$ . Initially, the input is 0.5, and the value at  $G$  is negative; thus the 1-bit output at  $H$  is -1. The output of the quantizer is +1 if its input is positive (MSB = 0) and -1 if its input is negative (MSB = 1).

The quantization error signal is formed by taking  $G-H$ ; this error is fed back into the double integration loops. The value of  $I$  is 0.4, and  $J = -0.6 + (-) -1 = 0.4$ . The value of  $E$  is  $2 \times J = 0.8$ , and  $C = A - J = 0.1$ . The value of  $F = D + E = 1.3$ . Similarly, values pass through the circuit at each sample time; for example,  $D$  takes the previous value of  $C$ . It can be seen that the values inside the loop are larger than the unit value; in other words, wider data buses are required. Also, if large values are input to the circuit, the limiter would be needed to prevent overloading of the loops. This example uses a constant input level, but high-speed averaging applies equally to varying input music signals. In either case, the output signal from the noise shaper is applied to a 1-bit D/A converter, which generates the appropriate analog voltage.

Two values are output at 512 times the sample rate. The signal conveys the audio waveform through the density of pulses; this pulse density modulation signal is converted into an analog signal using a switched dual-capacitor network. Specifically, a capacitor is charged and discharged according to the "1" or "0" value of the data; the result is an analog waveform which reflects the encoded waveform through time averaging of the output bit.

a)



b)

	A	B	C	D	E	F	G	H	I	J
$t_0$	0.5	-0.4	0.1	0.5	0.8	1.3	-0.6	-1	0.4	0.4
$t_1$	0.5	-0.3	0.2	0.1	0.6	0.7	1.3	+1	0.3	0.3
$t_2$	0.5	0.3	0.8	0.2	-0.6	-0.4	0.7	+1	-0.3	-0.3
$t_3$	0.5	-0.6	-0.1	0.8	1.2	2.0	-0.4	-1	0.6	0.6
$t_4$	0.5	-1.0	-0.5	-0.1	2.0	1.9	2.0	+1	1.0	1.0

Figure 4.33 An example of a second-order noise-shaping algorithm (a) with numerical values in the processing loop (b). (Courtesy Philips)

This converter retains its accuracy over environmental extremes, and a stereo converter can be contained on one chip. In addition, this chip may be used in a differential configuration.

#### Low-Bit D/A Conversion with Third-Order Noise Shaping

Several manufacturers have developed D/A conversion methods which employ third-order noise shaping. However, their design details differ somewhat, particularly in the noise-shaping algorithms and the low-bit output signal. The MASH system is a multistage third-order noise-shaping method. MASH was codeveloped by Matsushita Electric Industrial and Nippon Telegraph and Telephone Corporation. One implementation of this design accepts 16-bit words at a nominal sampling frequency, and a digital filter stage performs 8-times oversampling and outputs 24-bit words. Noise-shaping circuits output data as an 11-value signal at a 32-times oversampling rate. D/A conversion is accomplished via PWM (pulse width modulation), outputting low-bit data at a 768-times oversampling rate.

As with other low-bit systems, the key to the MASH system lies in noise shaping. As noted, if noise-shaping circuits are cascaded to achieve

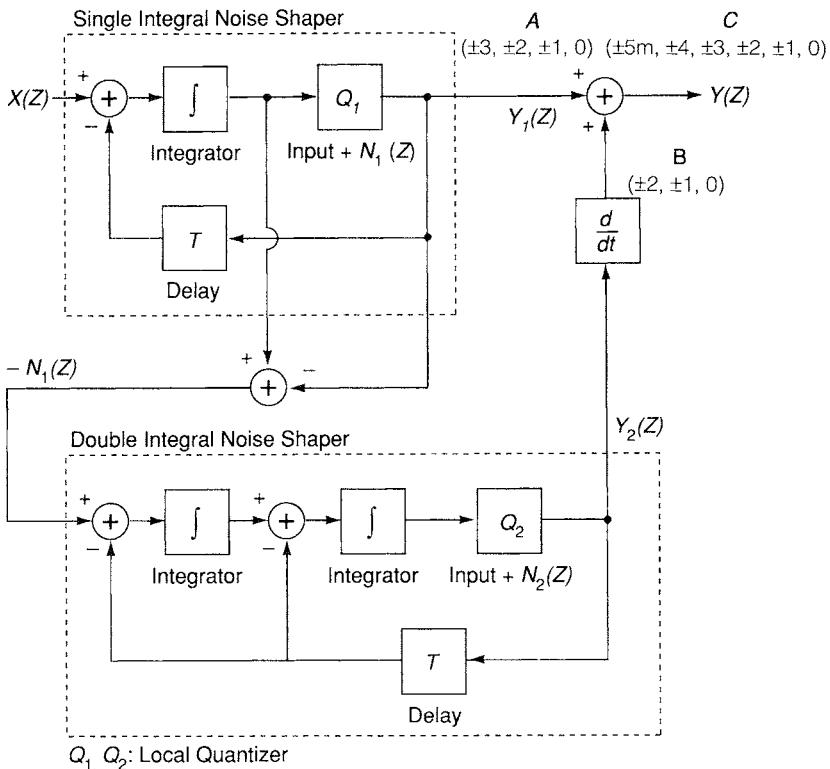


Figure 4.34 A block diagram of the signal processing in a third-order MASH noise-shaping algorithm. (Courtesy Matsushita)

more than second-order noise shaping, they can be prone to oscillation; the MASH circuit achieves third-order noise shaping but avoids oscillation through its “multistage” (noncascaded) configuration. A simplified schematic for the MASH noise shaper is shown in figure 4.34; it contains a first-order noise shaper in a main loop in parallel with a second-order noise shaper. Input data is requantized into seven values ( $\pm 3, \pm 2, \pm 1$ , and 0) at the main loop. At the subloop the requantization error is requantized into five values ( $\pm 2, \pm 1$ , and 0) used to correct the requantization error of the main loop. When the output values of the two stages are added together, the low-bit (4-bit) digital signal output from the circuit represents eleven values ( $\pm 5, \pm 4, \pm 3, \pm 2, \pm 1, 0$ ). These values represent the audio signal.

The requantization error of the output signal is shaped with a third-order differential characteristic (18 dB/octave). The final element in the system is analog conversion. The eleven-value signal is converted into

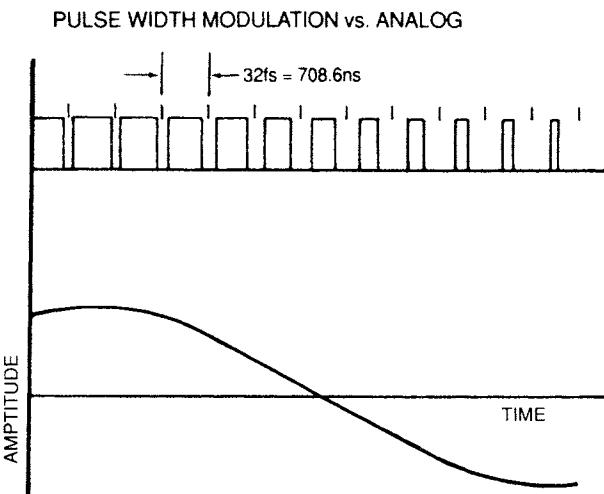


Figure 4.35 A pulse width modulation signal can be used to represent an analog waveform in the output stage of a MASH D/A converter.

pulses, each with a width corresponding to one value. This can be accomplished by applying the 4-bit output of the DSP circuit to a ROM to map eleven amplitude values into twenty-two time values with constant amplitude, each representing a “start” and “stop” time of the pulse widths. The widest pulses translate into a large positive output, while the narrowest pulses translate into a large negative output, as shown in figure 4.35. The width of the pulses carries the vital information; the amplitude of this signal can only be high or low. At this point the signal has the form of 1-bit binary data.

Because the signal is represented by a pulse width modulation waveform, final conversion is performed by simply passing the signal through a low-pass filter. Precise timing accuracy can be achieved through crystal oscillators; thus the widths are very accurate, and the error of the signal is low. Positive- and negative-going pulses may be output to cancel common noise. This  $33.8688 \text{ MHz}$  ( $768 \times f_s$ ) 1-bit data forms a PWM representation of the waveform.

Low-bit converter designs have already demonstrated the ability to provide outstanding performance. In terms of phase linearity, amplitude linearity, noise, long-term stability, and other parameters, oversampling D/A converters using low-bit architectures offer significant advantages over conventional multibit conversion methods. Moreover, the method is also applicable to A/D conversion. In short, the low-bit architecture is an exciting development in digital audio.

## SERIAL TRANSMISSION FORMAT

Analog audio signals can be conveyed from one device to another with a minimum amount of confusion. However, the transmission of digital data is a great deal more complicated on account of the potential disagreements on the sampling rate, the synchronization method, the block length, and many other factors. To permit an orderly exchange of data between professional digital audio devices, the Audio Engineering Society (AES) and European Broadcasting Union (EBU) jointly established a standard transmission format for linearly represented digital audio data. A number of digital audio devices with digital output use this format, which is known as the AES/EBU format; it has been codified as the ANSI S4.40-1985 standard and is sometimes referred to as AES3. For compatibility, many manufacturers of consumer electronics have adopted a format derived from the AES/EBU format for transmission of digital audio data. It has been described by the International Electrotechnical Commission (IEC) in a draft document, 84/WG11 as well as the IEC Publication 958, Digital Audio Interface "Consumer Use" specification.

The AES/EBU digital audio format transmits both left and right channel data using one digital cable. A time-sharing, multiplexing transmission transmits or receives frames, each containing left and right channel data alternatively. The transmission rate corresponds exactly to the source sampling frequency. When the sampling frequency is 44.1 kHz, 44,100 frames are transmitted per second. One frame consists of two subframes, labeled A (left channel) and B (right channel), each with 32 bits.

Each subframe carries data for one audio channel, as shown in figure 4.36. The first 4 bits are used for synchronization and for identifying preambles, which will be described shortly. The next 24 bits carry audio data, with the MSB transmitted last. Since most digital audio devices use 16-bit words, the last 16 bits in the field are typically used. In addition, the first 4 bits in the field are set aside for auxiliary audio or other data.

The last 4 bits form a control field which carries special information. The validity bit indicates if the previous audio sample is error free. The user data bit can be used to form a block for data associated with the audio channel, such as labeling of program number and duration; a high-level data link control (HDLC) communications protocol is specified for the user data channel. The channel status bit is used to form a data block; for each channel, one block is formed from the channel status bit contained in 192 successive frames. The start of a block is identified with special forms of the subframe preamble. The parity bit is used to provide even parity for each subframe; this permits simple error detection.

Three preambles are used to designate three types of events. One preamble marks the start of subframe A and the start of a channel status data

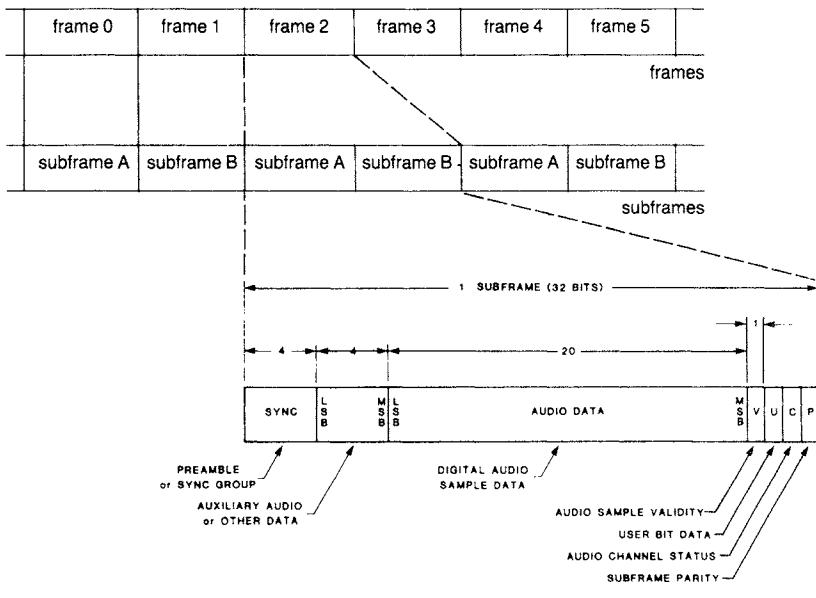


Figure 4.36 Subframe format for the AES/EBU serial transmission format. (Courtesy JAES)

block. Otherwise, a second preamble marks the start of subframe A. A third preamble marks the start of subframe B.

Each channel status data block consists of 192 bits (times 2) organized into two 24-byte sequences. Table 4.1 shows that these 8-bit bytes convey a considerable amount of information pertaining to the transmission. The first byte specifies a consumer or professional interface, use of pre-emphasis, and sampling frequency. The second byte specifies channel mode (e.g., stereo, or two-channel mono). The third byte carries word length information. The remaining channel status bytes specify information on sample address, time-of-day timecode, channel status validity, and CRCC.

The data is modulated with biphasic mark, in which a 0 corresponds to one inversion and a 1 corresponds to two inversions, as in the example in figure 4.37. However, the preambles ignore the inversion correspondences to provide unique identification, as shown in figure 4.38. The receiver input section uses this preamble portion for synchronization. The AES/EBU standard specifies that the digital waveform is transmitted over twisted pair conductors, with 3-pin XLR connectors (pin 1 shield or earth and pins 2 and 3 carrying signal).

<b>Byte 0</b>				<b>Byte 2</b>								
<b>Bits</b>	<b>Status</b>	<b>Defines:</b>	<b>As:</b>	<b>Bits</b>								
0	0	CSB usage	consumer	0–2	000	aux. bits not defined (20 bits)						
	1		professional		001	aux. bits used with main audio (24 bits)						
1	0	mode	normal audio		010	aux. bits coordination signal (20 bits)						
	1		non-audio	3–5		source word length and source recording history						
2–4	000	emphasis	not indicated	6, 7		Reserved						
	100		none			<b>Byte 3</b> multichannel modes						
	110		50/15 µs			<b>Byte 4</b>						
	111		CCITT J17			<b>Bits</b>						
5	0	sampling	locked	0, 1	AES11 sync reference							
	1	frequency	unlocked	2–7	Reserved							
<b>Bytes 5</b>				Reserved for future definition.								
<b>6–9</b>				Channel origin data. (7 bits ASCII)								
<b>10–13</b>				Channel destination data. (7 bits ASCII)								
<b>14–17</b>				local sample address code (32-bit binary)								
<b>18–21</b>				Time of day sample address code (32-bit binary)								
<b>Byte 1</b>				<b>22</b> Channel status validity flags.								
0–3	0000	mode of channel	not indicated	<b>Bits</b>	<b>Status</b>	<b>Defines:</b>	<b>As:</b>					
	0001		two-channel	0–3	0000	reserved	undefined					
0010			single-channel	4	0	bytes 0–5	are OK					
	0011		primary/secondary		1		are unreliable					
0100			stereophonic	5	0	bytes 6–13	are OK					
	1111		Byte-3 pointer		1		are unreliable					
4–7	0000	user bits management	default	6	0	bytes 14–17	are OK					
	0001		192-bit block		1		are unreliable					
	0010		AES18	7	0	bytes 18–21	are OK					
	0011		user defined		1		are unreliable					
<b>Byte 23</b> Channel status data CRCC (Cyclic Redundancy Check Character).												

Table 4.1 *Each channel status block consists of two 24-byte sequences. (Courtesy JAES)*

### *Consumer Interface*

The IEC serial format used in consumer digital audio equipment is derived from the AES/EBU standard. It is sometimes referred to as the SPDIF (Sony/Philips Digital Interface Format) standard. Generally, digital audio players and recorders should not be digitally interconnected if the connection mixes the professional format with the consumer format; improper data transmission or machine operation may result. The consumer format differs primarily in the definition of the channel status and user bits. The first bit of the channel status byte is set to 0, signifying a

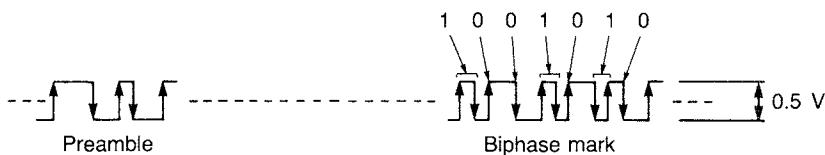


Figure 4.37 Data is modulated with biphase mark. (Courtesy JAES)

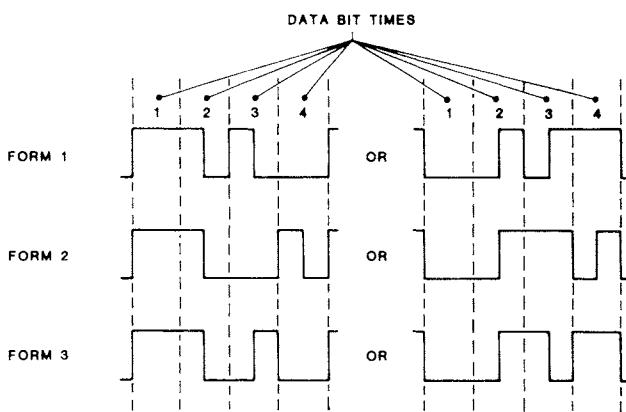


Figure 4.38 Three preamble forms are used. Form 1 denotes channel A, subframe, and block synchronization; Form 2 denotes channel A, otherwise; and Form 3 denotes channel B. (Courtesy JAES)

consumer interface. However, in the IEC/SPDIF version a different channel status specification is used, as shown in table 4.2. A general format defines a control code contained in the first 6 bits. A mode select is contained in the next 2 bits; presently only mode 00 is defined. In mode 00, there is a control code (as above), category code, simplified source field, sampling rate field, and sampling rate tolerance field. A category code of 100XXXX denotes a two-channel CD player. This sets the subframe structure as 16-bit words sampled at 44.1 kHz, and other parameters. Specifically, data is represented in two's complement notation with the MSB in position 27, the user bit channel may optionally contain the CD subcode, the V bit is optional, the control bits in the channel status are derived from Q subcode control bits, and the clock accuracy bits in channel status bits 28–29 are set according to source accuracy and use of variable speed.

The consumer interface also has provision to encode SCMS (Serial Copy Management System) data so that source data may be copied once, but

<b>General Specification</b>	<b>Mode 00 Specification</b>
<i>Control Word</i> (bits 0–5)	<i>Category Code</i> (bits 8–14)
bit 0 = consumer, 1 = professional	000XXXX = general format
1 0 = normal, 1 = digital data	100XXXX = 2-channel CD player
2 0 = copy prohibit, 1 = copy permit	010XXXX = 2-channel PCM
3 0 = pre-off, 1 = pre-on	110XXXX = 2-channel DAT
4   reserved	(X = don't care)
5 0 = 2-channel, 1 = 4-channel	<i>Generation Indicator</i> (bit 15)
<i>Mode</i> (bits 6,7)	0 = copy
bit 67	1 = original
00 = mode 00	(for CD player output)
01 } reserved	<i>Source Number</i> (bits 16–19)
10 }	0000 = undefined
11 }	1000 = source 1
<i>Program</i> (bit 15)	0100 = source 2
bit 15 0 = consumer or professional mode	.
1 = program transfer mode (if bit	.
0 = 0)	.
	1111 = source 15
	<i>Channel Number</i> (bits 20–23)
	0000 = undefined
	1000 = A (left channel)
	0100 = B (right channel)
	1100 = C
	.
	1111 = 0
	<i>Sampling Rate</i> (bits 24–27)
	0000 = 44.1 kHz
	0100 = 48 kHz
	1100 = 32 kHz
	.
	} reserved
	.
	1111
	<i>Clock Accuracy</i> (bits 28–29)
	00 = normal accuracy
	10 = high accuracy
	01 = variable speed

Table 4.2     *IEC/SPDIF channel status specification.*

second-generation copying is inhibited. SCMS encoding is contained in channels status bits 0, 2, and 15. Bit 0 must be set to 0, denoting a consumer interface. Bit 2 is a copyright bit; second-generation copying is prohibited when it is set to 0. Bit 15 is a generation indicator that denotes whether the data is original or copied. In the case of a CD player's output, when the program is original the bit is set to 0, and when it is copied data the bit is set to 1. For most other category codes (as denoted by bits 8

through 14), the bit's meaning is reversed: 0 denotes a copy, and 1 denotes original data. When a CD (with bits 2 and 15 set to 0) is copied to a digital recorder (such as DAT, DCC, or MD) through a consumer interface (with bit 0 set to 0), the data from the CD is treated as original data, and copying is permitted. However, the recorder will set bit 15 to 0 to mark the new recording as copied, thus preventing copying of the copy. If the CD's bit 15 is set to 1, no copying would be permitted. If bit 2 on the CD is set to 1, the copy will be infinitely copyable.

As noted, when interfacing CD data, provision is made to transmit the subcode data in the user bit channel. Subcode data is transmitted as if it were output from the disc, one subcode channel bit at a time, over ninety-eight frames. Each subcode field begins with a minimum of sixteen zeros, followed by a high start bit. Seven subcode bits (Q–W) follow. The next start bit and subcode bits may follow, or up to eight zeros may be inserted for timing purposes. The process repeats ninety-eight times, until the subcode block has been transmitted. The interface provides a user bit for each data sample, but because there are fewer subcode bits than audio samples ( $12 \times 98 = 1176$ ), the remaining bits are filled with zeros. The P subcode channel is not transmitted since it is merely a pickup location flag.

In the IEC/SPDIF version of the format used by consumer electronics manufacturers, video coaxial cable with phono plugs can be used to convey the data. Alternatively, many CD players feature fiber optic outputs with an optical connector to convey the digital data. Thanks to the transmission standard, digital audio devices from different manufacturers can connect for the transfer of digital audio information.

## FOR FURTHER READING

Adams, R. "Design and Implementation of an Audio 18-Bit Analog-to-Digital Converter Using Oversampling Techniques." *Journal of the Audio Engineering Society* 34(3):153–66, March, 1986.

Analog Devices. *Analog-Digital Conversion Handbook*. Englewood Cliffs, N.J.: Prentice-Hall, 1986.

Audio Engineering Society. "AES Recommended Practice for Digital Audio Engineering—Serial Transmission Format for Linearly Represented Digital Audio Data." *Journal of the Audio Engineering Society* 33(12):975–85, December, 1985.

\_\_\_\_\_. "AES Recommended Practice for Digital Audio Engineering—Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data." *Journal of the Audio Engineering Society* 40(3):147–165, March, 1992.

—. “AES Recommended Practice for Digital Audio Engineering—Format for the User Data Channel of the AES Digital Audio Interface.” *Journal of the Audio Engineering Society* 40(3):166–183, March, 1992.

Chamberlin, H. *Musical Applications of Microprocessors*. Rochelle Park, N.J.: Hayden, 1980.

Dijkmans, E. C., and P. J. A. Naus. “The Next Step Towards Ideal A/D and D/A Converters.” In *Proceedings of the AES Seventh International Conference—Audio in Digital Times*. K. C. Pohlmann, chair. New York: Audio Engineering Society, 1990.

Doi, T. “Recent Progress in Digital Audio Technology.” In Blessler, B., B. Locanthy, and T. G. Stockham, Jr., eds. *Collected Papers of the AES Premiere Conference—Digital Audio*. New York: Audio Engineering Society, 1983.

EBU (European Broadcasting Union). “Specification of the Digital Audio Interface.” *EBU Doc. Tech.*, 3250.

Fielder, L. D. “Evaluation of the Audible Distortion and Noise Produced by Digital Audio Converters.” *Journal of the Audio Engineering Society* 35(7/8):517–35, July/August, 1987.

Finger, Robert A., “AES3–1992: The Revised Two-Channel Digital Audio Interface.” *Journal of the Audio Engineering Society* 40(3):107–116, March, 1992.

Goedhart, D., R. J. Van de Plassche, and E. F. Stikvoort. “Digital-to-Analog Conversion in Playing a Compact Disc.” *Philips Technical Review* 40(6):174–79, 1982.

IEC (International Electrotechnical Commission). “Draft Standard for a Digital Audio Interface.” *IEC Report TC 84/WG11*, November 1, 1986.

Jung, W. G. “The Magnavox 16-Bit Series.” *Audio Magazine* 71(6):74–80, June, 1987.

Kloker, K. L., B. L. Lindsley, and C. D. Thompson. “VLSI Architectures for Digital Audio Signal Processing.” In *Proceedings of the AES Seventh International Conference—Audio in Digital Times*. K. C. Pohlmann, chair. New York: Audio Engineering Society, 1990.

Lagadec, R., and G. J. McNally. “Labels and Their Formatting in Digital Audio Recording and Transmission.” *AES Preprint No. 2003*, 74th Convention, 1983.

Matsuya, Y., K. Uchimura, A. Iwata, T. Kobayashi, M. Ishikawa, and T. Yoshitome. “A 16-Bit Oversampling A-to-D Conversion Technology Us-

ing Triple-Integration Noise Shaping.” *IEEE J. Solid-State Circuits* SC-22(6):921–29, December, 1987.

Matull, J. “IC’s for Compact Disc Audio Decoders.” *Electronic Components and Applications* 4(3):131–41, May, 1982.

\_\_\_\_\_. “IC’s for Compact Disc Audio Decoders.” *Signetics Linear Data and Applications Manual*, vol. 2, 1985.

Meyer, J. “Time Correction of Anti-Aliasing Filters Used in Digital Audio Systems.” *Journal of the Audio Engineering Society* 32(3):132–37, March, 1984.

Naus, P. J. A., E. C. Dijkmans, E. F. Stikvoort, A. J. McKnight, D. J. Holland, and W. Bradinal. “A CMOS Stereo 16-Bit D/A Converter for Digital Audio.” *IEEE Journal of Solid State Circuits* SC-22(3):390–95, June, 1987.

Nijhof, J. “An Integrated Approach to CD Players, Part 1: The Optical Pick-Up.” *Electronic Components and Applications* 6(4):209–15, 1984 (Philips TP172).

\_\_\_\_\_. “An Integrated Approach to CD Players, Part 2: The Decoding Electronics.” *Electronic Components and Applications* 6(4):216–22, 1984 (Philips TP172).

Pohlmann, K. C. *Principles of Digital Audio*. Indianapolis: Sams, 1989.

\_\_\_\_\_. “Low-Bit Conversion and Noise Shaping.” In K. C. Pohlmann, ed., *Advanced Digital Audio*. Carmel, Indiana: Sams, 1991.

Ranada, D. “CD Player Filters.” *Stereo Review* 50(7):50–52, July, 1985.

Schott, W. “Philips Oversampling System for Compact Disc Decoding.” *Audio Magazine* 68(4):32–35, April, 1984.

Schouwenaars, H. J., E. C. Dijkmans, B. M. J. Kup, and E. J. M. van Tuijl. “A Monolithic Dual 16-Bit D/A Converter.” *IEEE Journal of Solid State Circuits* SC-21(3):424–29, June, 1986.

Talambiras, R. P. “Limitations on the Dynamic Range of Digitized Audio.” In J. Strawn, ed. *Digital Audio Engineering: An Anthology*. Vol. 3. The Computer Music and Digital Audio Series. Madison, Wisc.: A-R Editions, Inc., 1985. (Originally published Los Altos, Calif.: William Kaufmann.)

Uchimura, K., T. Hayashi, T. Kimura, A. Iwata. “Oversampling A-to-D and D-to-A Converters with Multistage Noise Shaping Modulators.” *IEEE Trans. ASSP* 36:1899–1905, December, 1988.

- Van de Plassche, R. J. "Dynamic Element Matching for High-Accuracy Monolithic D/A Converters." *IEEE Journal of Solid State Circuits* SC-11(6):795–800, December, 1976.
- Van de Plassche, R. J., and E. C. Dijkmans. "A Monolithic 16-Bit D/A Conversion System for Digital Audio." In Blessing, B., B. Locanthi, and T. G. Stockham, Jr., eds. *Collected Papers of the AES Premiere Conference—Digital Audio*. New York: Audio Engineering Society, 1983.
- Van der Kamm, J. J. "A Digital 'Decimating' Filter for Analog-to-Digital Conversion of Hi-Fi Audio Signals." *Philips Technical Review* 42(6/7):230–38, April, 1986.
- Watkinson, J. R. "Inside CD, Part 2." *Hi Fi News and Record Review* 31(12):46–47, December, 1986.
- . "Inside CD. Part 5." *Hi Fi News and Record Review* 32(3):41–43, March, 1987.
- Yamaha. "Yamaha Hi-Bit Technology." Yamaha White Paper, June, 1987.

## *Chapter Five*

# Practical Concerns

### **INTRODUCTION**

All CD discs and players offer considerable advantages over other audio media. The specifications are great: flat response, low noise, low distortion over the entire audio frequency band, and absolute speed stability and accuracy. The operational features are great: small size, long playing time, timing and contents information, indexing, and programmability. The longevity is great: Discs are completely immune to damage from repeated playing and relatively impervious to handling damage. Still, as with any other product in the marketplace, there are good players and not-so-good players, good discs, and not-so-good discs. Remember—a smart consumer is a retailer's worst enemy.

### **DESIGN CONSIDERATIONS**

Evolution is the name of the technology game. While a company's marketing types are busy selling the public on the ultimate product, the engineers must be ready with its successor. When the compact disc was first introduced, a certain co-inventor touted it as the "last word" and "pure and perfect." Perfection implies the impossibility of improvement, and everything, especially everything technological, can be improved. Sure enough, CD hardware has evolved quite nicely. Of course, place that in perspective. Once a system is standardized, improvements are usually orders of magnitude less impressive than the original achievement. But a few adjustments can often be quite significant, especially in something as subtle as audio.

Manufacturers are taking advantage of several opportunities to improve CD players, not radically, but nicely. Let's check out some of these improvements—some borrowed from audiophile lore, others stolen from customized players, and the rest inspired from good old common sense engineering.

### ***Power Supply***

The power supply is probably the most unglamorous part of a CD player, mundanely contributing voltages, at prescribed current levels, to energize the system. Yet the power supply is a vital contributor to the player's audio performance. A player's servo control system, for instance, draws considerable current whenever the pickup moves around. This can occur normally—for example, in disc accessing—or under unusual circumstances—for example, when disc wobble (perhaps in the throes of acoustical vibration) stresses the servo. This power demand can translate into variations in the power delivered to other circuits and consequently affect their operating characteristics (rather like a brownout in a power grid). In short, fluctuations in the power supply modulate the rest of the player's circuitry. The distortion parameters of the output analog circuitry can change for the worse, or jitter might be introduced into the master clock, fouling up the digital circuitry. The solutions, increasingly endorsed by manufacturers, are high capacity power supplies which can meet demand under all conditions or separate power supplies for each part of the system: pickup and servo, digital processing, analog output, and visual display. Separate supplies are particularly nice because they permit isolation of grounds, reducing electrical interference between circuits. Because of servo stress, and for all kinds of reasons, carefully designed power supplies make a lot of sense.

### ***Optical Pickups***

Which is better, three-beam or one-beam pickups? In the final analysis, both designs perform the same function, but the engineering advantages may favor the one-beam design. The optics in a three-beam design are inherently more complicated; there are at least five components (objective, collimator, grating, cylindrical, and beam splitter) compared to the four components in a one-beam design (objective, collimator, wedge, and semitransparent mirror). In addition, the three-beam design requires several critical adjustments, including secondary beam and diffraction grating alignment, as opposed to the single adjustment of the horizontal position of the photodiode assembly in a one-beam design. Both designs require adjustment of drive current for the laser diode.

A three-beam design is also more complex from a mechanical viewpoint. A one-beam design can be mounted on a single pivoted arm which describes an arc across the disc surface; since there is only one moving part, wear is minimal and reliability is high. A three-beam design requires a linear sled mechanism because the tracking beams must stay in a fixed position relative to the pit track. In terms of traditional specifications,

such as frequency response and channel separation, there is no sonic difference between one-beam and three-beam designs. Ultimately, pickup performance must be evaluated on the basis of individual players.

Optical pickups have undergone extensive improvement. New types of lens assemblies have resulted in smaller and lighter pickups; this is important because a lower-mass pickup is better able to traverse the disc and respond to deviations in the pit track, as in the case of a disc with an eccentric center hole. Newer designs also require less calibration and are thus freer of misalignment, a problem with early-generation pickups.

Disc sleds have been greatly improved. They operate faster and smoother than their slow, bulky ancestors. Because sleds must carry the pickup across the disc surface, tracking and disc access are only as good as they are. Simple, linear motors are being used in many players, replacing rotary motors and their associated worm gears and reduction gear mechanisms. In a linear system, the motor's construction is flat underneath the disc surface, and the motor's torque is transferred directly to the pickup without any gears. The sled and its cargo, the pickup, may travel swiftly to access tracks quickly on a disc (some players can access any point on a disc in one second or less) yet provide precise tracking of the pit spiral. In addition, better sled suspension, ceramic material, and disc clamping help keep the pickup on track even when the room is rocking.

### *Servo Systems*

Servo systems must keep the laser pickup on track and in focus even under difficult conditions such as the presence of disc defects. When a disc defect passes by the laser, the RF (radio frequency) signal from the pickup momentarily drops in level. When this happens, the player loses its grasp on tracking and focusing information as well as on audio data. The error correction code springs into action to correct or conceal audible errors, and the servo tries to keep the pickup on track.

In recent designs, the drop in the RF signal is detected more quickly, and the pickup's position is frozen to prevent it from wandering too far off track. In addition, the error correction circuits are alerted to the trouble. Figure 5.1 shows an RF signal interrupted by errors and slow and fast detection circuits in action. Obviously, the faster circuit permits faster recovery from disc defects and more uniform tracking performance.

### *Opto-electronics*

Optical transfer devices have been introduced into CD technology. This solves a very basic problem for engineers: the incompatibility of analog and digital circuitry. The digital circuitry may cause interference in the

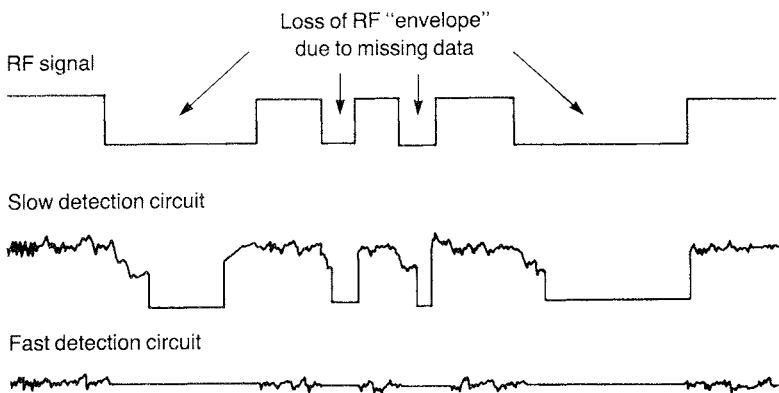


Figure 5.1 With fast detection circuits, dropouts in the RF signal may be sensed quickly, and the servo voltage held, until the RF signal is restored. (Courtesy Sony)

analog signal: For example, digital circuitry often generates radio frequency interference, resulting in noise in analog sections. The conflict is especially worrisome in a device where digital and analog circuits must coexist, as in a CD player.

Optical transfer devices help protect analog circuitry from digital interference. For example, some devices convert the electrical signals into light, which is focused and beamed through a short fiber optic cable to a receiver and then converted back into an electrical signal. Thus data is transferred from the digital circuitry to the analog, but electrical isolation is preserved. When separate power supplies with separate groundings are provided, the two types of circuitry can easily coexist. Optical transfer also helps iron out fluctuations introduced into the signal by power supply modulation.

With optical transfer, much of the spurious noise present in the digital signal (to which digital circuitry is largely impervious) is isolated. Hence the analog section has a much cleaner signal to work with. Will optical transfer make the analog output signal cleaner? Yes. And will it *sound* cleaner? Well, maybe. Much of the spurious noise lies in the MHz region, along with television, radio, and other supersonics. However, the noise can create components in the audio band, and optical transfer can help eliminate them.

Another role for light in CD player design is the use of fiber optics. Increasingly, players feature fiber optic outputs that convey digital audio data via modulated light instead of electricity. In this way, noise and interference encountered en route is minimized. Moreover, grounding

incompatibility between components is alleviated. Of course, your preamplifier must have both an optical input and D/A conversion circuitry.

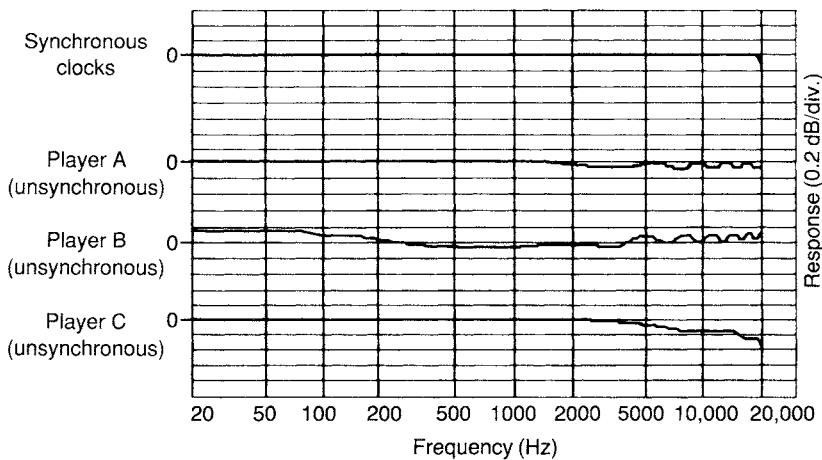
### ***Electronics***

Oversampling filters contain some extremely sophisticated bits of engineering, and the last word hasn't been said. One concern is pass-band ripple. Some studies indicate that even minute amplitude variations in frequency response cause other audible artifacts. High-resolution filters have been designed to hold ripple to less than  $\pm 0.0001$  dB across the audio band; some digital filters achieve this. The oversampling rate employed in digital filters has increased from two- or four-times to eight-times and beyond. This increase moves image spectra to a higher frequency, thereby decreasing the required steepness of the analog filter needed to remove them. While this is an asset, the oversampling rate alone is not a measure of quality. More important is the accuracy of the numerical computation inside the filter. Roundoff errors and other considerations ultimately determine the quality of the digital filter.

Many players exhibit high-frequency components accompanying the audio signal at the player's output jacks. These components amount to spurious noise, ranging from 24 kHz on up, beyond the presumed range of hearing. In many cases, the components are actually beat frequency intermodulation products caused when various digital clocks react with each other. Some players create them and let them pass through, and other players create them and then filter them out. Recently, however, more emphasis has been placed on preventing them in the first place.

When the player is designed with a single master clock, rather than with several unsynchronized clocks, beat frequencies that result in spurious supersonic components are usually reduced. A single master clock is employed for all the processing circuits, including the oversampling filter and D/A converter. As a result, the data stream from the EFM demodulator to the D/A converter is synchronous, preventing any internal beating.

What difference would a beat frequency of perhaps 88.2 kHz make? Although inaudible in itself, the component could intermodulate with audio frequencies, creating distortion products in the audio band. In addition, in-band measurements are often improved as well. For example, figure 5.2 shows the frequency response of four players—one with a single master clock and three without. The use of synchronous clocks clearly results in a flatter frequency response, because interaction between different clocks is reduced. In players A and B, fluctuations occur at high frequencies. Even worse, player C shows a high frequency roll-off. There is growing sentiment that a CD player should output the audio material present on the disc without any spurious supersonics.



**Figure 5.2** Frequency response can be affected by design of the player's clocks. Frequency response of a player with synchronous clocks is compared to asynchronous clocks in players A, B, and C. (Courtesy Sony)

### *Analog versus Digital Filters*

CD players may use two-times, four-times, eight-times, or even sixteen-times digital filters, in which the 44.1 kHz sampling rate is oversampled to 88.1 kHz, 352.8 kHz, or 705.6 kHz, respectively. Accordingly, the characteristic of the analog low-pass filter must vary. The lower the oversampling frequency, the closer to the audio band the image spectrum will be, and hence the sharper the analog filter required. Some designers feel that in oversampling designs of eight-times or more, the remaining image spectrum is at such a high frequency that no analog filtering is needed.

A digital filter design is thus an efficient method of accomplishing the task of anti-imaging without resorting to analog brickwall filters. This alleviates the problem of phase distortion introduced by analog filters. A digital filter can have almost perfect linear phase response.

In terms of overall performance, digital filters represent a great improvement over analog filters because a digital filter can be designed with considerable, stable precision, whereas an analog filter design is inherently limited. Other specific advantages of digital filtering are more subtle: The filter characteristic varies with the clock rate. Thus its parameters may be altered. For example, variable speed CD players could be easily designed for special applications where the data rate and the pitch of the music are varied.

In addition, the output circuits have an easier time dealing with the waveform because the successive changes in amplitude are smaller in an oversampled signal. More precisely, the slew rate, the rate of variation in the output waveform, is lower. This, along with less ringing and less overshoot, reduces intermodulation distortion. Finally, because digital filtering is a purely numerical process, the characteristics of the digital filtering, as opposed to analog filtering, cannot change with temperature, age, and so on.

On the other hand, a digital filter performs its calculations at high speed, which increases power consumption. In addition, the software design of the filter's numerical circuits is a somewhat esoteric task, so great care must be taken to provide for correct computation. Furthermore, because the number of samples has been increased by several times, the D/A converters must do their job several times faster. Thus higher quality, more expensive D/As are required. This cost may be offset by the lower cost of the digital filter chip relative to an analog filter.

In terms of sonic performance, does digital filtering do a better job than analog filtering? Do CD players with digital filters sound better than those with analog filters? After careful listening, most critics tend to favor digital filter systems. There are exceptions, however, and complete generalization would be a mistake. In practice, for a variety of reasons, virtually all CD players now employ oversampling filters.

So much for filtering problems in CD players. But what about digital audio recorders? What good is perfect phase linearity at playback when every disc contains phase nonlinearity introduced by the anti-alias brick-wall filters of the digital recorder? At first glance, the problem is insurmountable. The Nyquist theorem dictates that the input anti-aliasing filter be present before sampling occurs. In other words, because filtering must take place on the analog signal before digitization, the filter must be an analog design.

There is, however, a way to put the cart before the horse using oversampling techniques. In short, an oversampling digital input filter can be designed. In essence, such a filter samples at a very fast rate (following a very gentle analog low-pass filter) to greatly extend the Nyquist frequency. A computational process called decimation is used to lower the initial high sampling rate to 44.1 kHz. Digital filtering provides the anti-aliasing function to prevent aliasing between audio signals and the 44.1 kHz sampling rate. With this technique, at least in theory, perfect phase linearity throughout the digital audio chain is possible.

### ***One versus Two Converters***

A question over which circuit designers (or at least their company's accountants) disagree is the number of D/As to use in a player. Until the

all-digital system becomes commonplace, and CD players omit D/As altogether, we have the option of using either one or two D/As in a player.

At first glance, two D/As seem necessary. Since the compact disc is a stereo playback medium, two audio channels are reproduced. As figure 5.3a shows, the digital data stream can be demultiplexed to left and right channels and fed to two D/As, resulting in two audio outputs. For strictly budgetary reasons, however, many players employ only one D/A, as in figure 5.3b. The multiplexed data of the two audio channels is handled by one D/A (operating twice as fast as the dual D/As). Only after the D/A has performed its conversion are the two analog channels separated into left and right channels.

Although the results of each method are apparently identical, small differences are introduced. The single D/A design creates several anomalies in the signal. First, while the data bits from different channels are easily split apart in the dual D/A design, the interleaved segments of analog voltages produced from single D/A design are not so easily separated. A switch must be used to route these pieces of waveform (88,200 per second) to the correct output channel. The switch might cause distortion as it flips back and forth.

Second, the single D/A design introduces a slight time difference (11.34 microseconds) between the two audio channels. In a dual D/A design, the demultiplexed data appears at the two converters simultaneously, exactly as it appeared at the A/D converters of the digital recorder. The output analog channels are thus exactly in phase—that is, exactly matched in time. It is possible to compensate for the offset introduced by a single

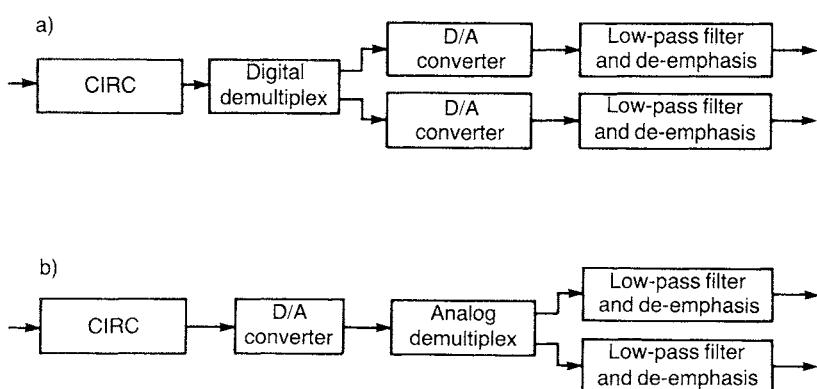


Figure 5.3     A CD player may employ two D/A converters (a) or only one multiplexed converter (b).

D/A. However, this is difficult to achieve uniformly over the entire frequency range.

Acoustically, the use of a single D/A converter and the resulting 11.34 microsecond interchannel difference can affect stereo imaging. In addition, the offset could affect downstream equipment. In stereo broadcasting, for example, cancellation could result in high-frequency roll-off of the L + R component and reinforcement of the L – R component. Dual D/As avoid these liabilities.

Unfortunately, not everything is sunny in dual D/A land; a pair of converter chips might mistrack because of fluctuations in their reference voltages or anomalies in the chips themselves. The result could again be phase error. Some manufacturers have solved both problems by offering dual D/A converters on a single chip, which is nearly an ideal solution. Nevertheless, the evidence weighs in favor of dual D/A converters, even with dual-chip designs.

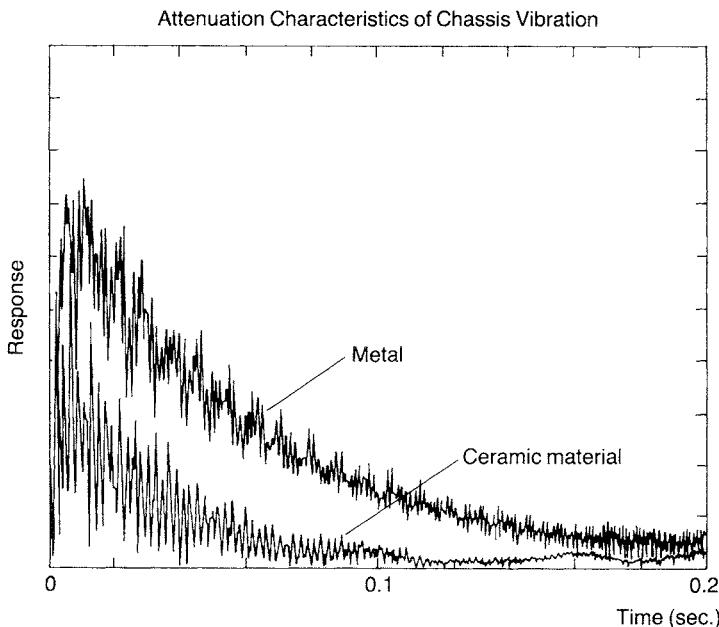
### ***Vibration***

Vibration is bad. Transmitted to the analog circuitry, it can deteriorate sound quality; for example, mechanical vibration can be converted into electrical signals by capacitors. As we have seen, a vibrating disc can stress the servo system and the power supply.

Two sources of vibration common to every CD player are the disc and the sound output itself. Internal motor and disc-generated vibration can be damped by using ceramic or fiberglass material in the disc drive. These materials act as poor conductors of vibration, creating friction as vibration tries to pass through. It is quickly attenuated. Figure 5.4 shows the damping characteristics of a metal part as opposed to a ceramic part. In addition, rubber bushings under the drive help isolate it from the chassis and the electrical components.

External vibration can be created acoustically from the pressure function of loudspeakers. The player's chassis forms the first line of defense. High quality player chassis are consequently more massive, with thick plastic or metal panels, and are better braced in order to reduce resonance. Some players employ a dual-layer construction in which an impact-deadening material is sandwiched between two metal plates of different thicknesses. Alternatively, calcium carbonate material, reinforced with fiberglass, has been used. These chassis are remarkable; when you rap on them with your knuckles, the sound dies, along with any unwanted vibrations. In addition, the location and number of welds and screws and the placement of absorbing materials applied to inside areas all help damp vibration.

As a last resort, rubber or plastic feet can be placed under a CD player to isolate it from structure-borne vibrations. This will not improve data as



**Figure 5.4** Ceramic materials help dampen vibration compared to metals. (Courtesy Sony)

read from a disc and will have no effect on the output signal under normal playing conditions. However, the isolation they provide can reduce signal distortion under extreme vibration conditions. This is not altogether surprising. Although a CD player's audio path is perhaps 90 percent digital, the final 10 percent is analog, and it has long been known that analog components are susceptible to vibration.

The problem is microphonics, in which electrical devices act as transducers, transforming the vibration signal into an electrical signal. Microphonics is a major problem in vacuum tubes and is also present to a much lesser degree in solid state devices such as capacitors and transistors. Once the vibration signal has entered the analog electrical signal path, it becomes part of the analog signal. As one might expect, the problem is most severe in high-gain devices, in which the signal (and the effect of the vibration) are boosted considerably.

For example, one CD player test provided the following results: Without vibration, THD + N distortion measured 0.0022 percent and with vibration measured 0.0044 percent. IM distortion similarly increased from 0.0032 percent to 0.0043 percent. Clearly, the vibration in the player affected the quality of the output analog signal. When isolation feet were

placed under the CD player, they helped decouple the player from the vibration source, and although distortion increased in the presence of vibration, it was not as high as without isolation. The decoupling effect was easily verified; by lifting the player an inch above the vibration surface, all microphonics stopped. In this test microphonics did not occur in the CD player until the sound pressure level exceeded 110 dB. A player with less mass could be susceptible at lower levels.

Vibration of the output section in a CD player should be avoided just as it should be avoided in any analog audio device such as a turntable, preamplifier, or power amplifier. While digital data would be affected only at the point of pickup mistracking, vibration can affect the analog output at relatively lower vibration levels. The efforts of manufacturers to minimize vibration in CD players are valid, and it appears that analog sections are more prone to problems than the disc transport itself. However, the effects of vibration are subtle. Generally, microphonics will occur only at high sound pressure levels where a slight increase in distortion might not be audible. Moreover, the added microphonic signal is correlated to the audio signal, thus making it even more difficult to hear.

Finally, it is important not to confuse player mass isolation with disc treatments such as sound rings—plastic bands slipped around the circumference of a disc. Tests with a CD player under vibration show that the increase in distortion in the analog signal is identical with or without disc rings. This is not unexpected because it is not vibration in the disc that causes microphonics, but vibration in the analog components. Disc treatments are discussed further at the end of this chapter.

### ***Other Improvements***

Capacitors used for the D/A chips can be susceptible to vibration. Capacitors with electrodes made of copper and dense film can be used for their antivibration characteristics. Heat sinks made of copper with a low *Q* factor also help control resonance.

The search for good capacitors continues. Some manufacturers have added ceramic powder to their electrolytic capacitors for better impulse response and faster and more accurate rise-time characteristics. Of course, all other components must be audio-grade, including polystyrene capacitors with copper foil plate, one percent tolerance metal film resistors, oxygen-free copper wire, and copper-plated chassis.

Electrical interference must be controlled both internally and externally. Every circuit designer knows the importance of shielding and grounding. One might also contemplate the question of the effect of RF noise transmitted by a CD player on other audio components, such as amplifiers. The proliferation of digital audio gear increases the problem of radio frequency interference.

Operational amplifiers still aren't perfect, or even close to perfect. Manufacturers still struggle to find distortion, noise, slew rate, and heat dispersion specifications fully adequate for audio.

Manufacturers have taken a close look at early-generation players and returned to the drawing board for some very worthwhile improvements. Are they drastic improvements? That depends on your perspective. If you liked CD playback from the beginning, you might hear a slight increase in the quality of the sound. If you had a problem with early-generation players, these improvements might very well satisfy you, in which case the result would be considerable. In any case, they are worthwhile because the improvements can be analytically verified, whether the problems were initially audible or not. And that is the essence of evolution in audio technology.

## SPECIFICATIONS

Perceiving nuances in the quality of different digital audio systems places great demands on even the best ears. Measured specifications are thus increasingly important in evaluating digital audio products. Most of us are lucky to hear 20 kHz, let alone a 0.1 dB deviation at that frequency. Specifications provide a standard of comparison. The CD player, the amplifier, the loudspeakers, the listening room acoustics, and personal subjectivity combine to produce an environment which is simply too unique to provide a meaningful reference. Thus specifications become an important part of any evaluation, for they permit us to probe both the limits of a player's sonics and the limits of our hearing in a standardized manner.

An understanding of specifications is crucial for the educated consumer. The numbers published by the manufacturer or quoted in a magazine test report are meaningful only if the basis for them is understood. Let's consider CD test procedures and some of the principal specifications for CD players. Specifications derived from analog technology are generally used; this is understandable because the analog output stages are often the limiting factor in many aspects of player performance.

### *Test Discs*

Any test needs a standard source. In the case of CD players, test discs are available from Philips (410 055-2/3, 410 055-2/5, 410 055-2/5A and SBC 429), PolyGram (DHF19-410741-2), Sony (CD3-YED-S7 and 48DG3), Denon (38C39-7147 and 33C39-7441), Technics (SH-CD001), Pierre Verany (PV788031/788032), CBS (CD-1), and other manufacturers, as well as organizations such as the Japan Audio Society (CD-1 YDDS-2), Hi-Fi

News (HFN0-03), and the Electronic Industries Association of Japan (EIAJ CD-1 YGDS 13). They contain diverse test signals and musical selections.

The CBS test disc CD-1 deserves special mention. It includes all of the tests specified by both the Electronic Industries Association (EIA) and the EIAJ, along with several additional tests devised by CBS. It contains tracks of low-level amplitude-swept and amplitude-stepped signals (from about -70 dB to -100 dB) recorded with dither. These permit measurement of linearity and distortion without the effects of quantization noise. The use of dithered test signals is crucial at low amplitudes because it would be otherwise impossible to separate the player's distortion from those in the test signal itself. The -100 dB signal demonstrates the ability of the CD system to accurately reproduce signals below the least significant bit when proper dither has been added. The contents of the CBS CD-1 test disc are shown in table 5.1

The EIAJ has issued a test standard in two reports (see the references at the end of this chapter). "Methods of Measurement for CD Players" contains the following thirteen tests: frequency response, signal-to-noise ratio, dynamic range, total harmonic distortion, channel separation,

FOR MEASURING CD PLAYER PERFORMANCE WITH E.I.A. STANDARD SIGNALS		
CBS CD-1 TEST DISC		
1 Reference, L & R, 0 dB, 1 kHz	10 19 997 Hz (Also used for Pitch Error)	16 Square Wave, L & R, 0 dB, 1 002.27 Hz
2 Left Separation, 0 dB, 1K, 125, 4K, 10K, 16K Hz	11 Sweep Frequency Response, 0 dB, 5Hz -22.05 kHz	17 Impulse & Polarity Test, 0 dB, L & R
3 Right Separation, 0 dB, 1K, 125, 4K, 10K, 16K Hz	12 De-emphasis Error, L & R, 1K, 125, 4K, 10K, 16K Hz	18 Linearity, 997 Hz, L & R, 0 dB, -1, -3, -6, -10, -20, -30, -39.99, -49.97, -59.94, -70.31, -80.77, -90.31 dB
4 Signal/Noise, L & R, Infinity Zero w/o em- phasis	13 Intermodulation Distortion (SMPTE, Twin Tone), L & R 60 Hz + 7 kHz, 11 kHz + 12 kHz	19 Linearity with Dither, 997 Hz, L & R *-70.31, -80.77, -90.31, -100 dB
5 Dynamic Range, L & R, 1kHz, -60 dB w/o em- phasis	14 Wow & Flutter, L & R, 0 dB, 3150 Hz	20 Fade to Noise with Dither, L & R *-60 dB, 500 Hz
6 Frequency Response, L & R, 0 dB, 4, 8, 17, 31 Hz	15 Access Time, L & R, 0 dB, 317 Hz	21 Monotonicity, L & R, *1 102.5 Hz, 10 LSB
7 61, 127, 251, 499 Hz		*Additional Test Tracks
8 997, 1999, 4001, 7993 Hz		
9 10 007, 12 503, 16 001, 17 989 Hz		

Table 5.1     Contents of CBS CD-1 test disc. (Courtesy CBS)

de-emphasis error, wow and flutter, intermodulation distortion, phase difference between channels, level difference between channels, output voltage, pitch error, and access time. In addition, it specifies a standard method for performing and documenting the tests. "Test Disk for CD Players" describes a standard test disc to be used to perform such tests.

The signals on test discs are computer-generated to guarantee accuracy. Standard specifications can be measured with their menus of signals. Of course, test gear is needed to decipher the results; a good test bench has an oscilloscope, a distortion and level meter, and a harmonic analyzer. (Indeed, a good test bench should also have a CD player and test disc. They provide a highly accurate set of reference signals useful in testing other audio equipment, room acoustics, etc.) As an alternative to separate pieces of measuring equipment, integrated, computer-based test sets offer high accuracy and fast operation. Equipped with a test disc, measuring gear, and the player under test, we can run some numbers. Many of the following procedures are more fully described in an applications note by Audio Precision, as listed in the references at the end of this chapter. A word of warning: Test discs contain maximum amplitude signals. Turn down your amplifier to avoid possible damage to the loudspeakers when playing them over a home system.

### ***Frequency Response***

Frequency response measures a signal's amplitude deviation over the audio bandwidth of 0 to 20 kHz; any amplitude deviation, either positive or negative, is measured in decibels (dB) relative to a 0 dB (maximum) amplitude at 1 kHz. The test can be performed by using either independent tones or a swept tone recorded at 0 dB. The two channels are tested separately. By definition, the digital signal should have a perfectly flat response, but the output analog circuitry usually interferes. A single channel measurement is shown in figure 5.5.

Frequency response measures one of the most audible aspects of any audio system. A slight difference in frequency response can account for diverse qualitative judgments, including image, depth, air, clarity, and warmth. Fortunately, the frequency response of digital audio equipment is one of its fortés. A good CD player will have an essentially flat response with a negative deviation of perhaps a few tenths of a decibel at the high frequencies.

### ***Channel Separation***

Channel separation measures the ability of the player to keep the stereo signals distinct from each other; again, the problem lies mainly in the analog output circuitry and usually grows more severe at high frequen-

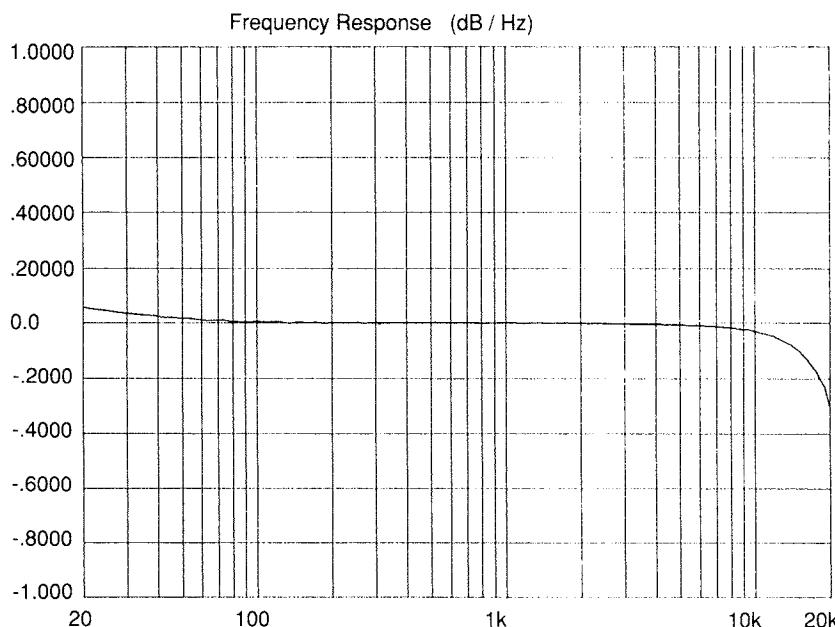


Figure 5.5 Frequency response measurement.

cies. For this test, 0 dB reference tones of various frequencies are played through one channel, and the crosstalk is measured on the otherwise quiet channel, as shown in figure 5.6. When a broadband measurement is taken, crosstalk of both the fundamental and its harmonic components are measured. Many manufacturers prefer to measure only crosstalk from the fundamental frequency under test. Thus a filter (a 22 kHz low-pass or a selected band-pass) is used to remove wideband signals which could decrease the measured value of separation.

Whatever method is employed, the crosstalk level is measured as an amplitude in dB. A figure of –100 dB at 1 kHz, for example, means that 1 kHz information on one channel will also appear on the other channel, but down 100 dB. A channel separation frequency response may thus be plotted. Curiously, channel separation is usually not reciprocal between two channels. A good player should have channel separation of more than 90 dB across the audio band.

#### **Total Harmonic Distortion Plus Noise**

Total harmonic distortion plus noise (THD + N) measures a device's ability to reproduce a signal without adding components of its own. Specif-

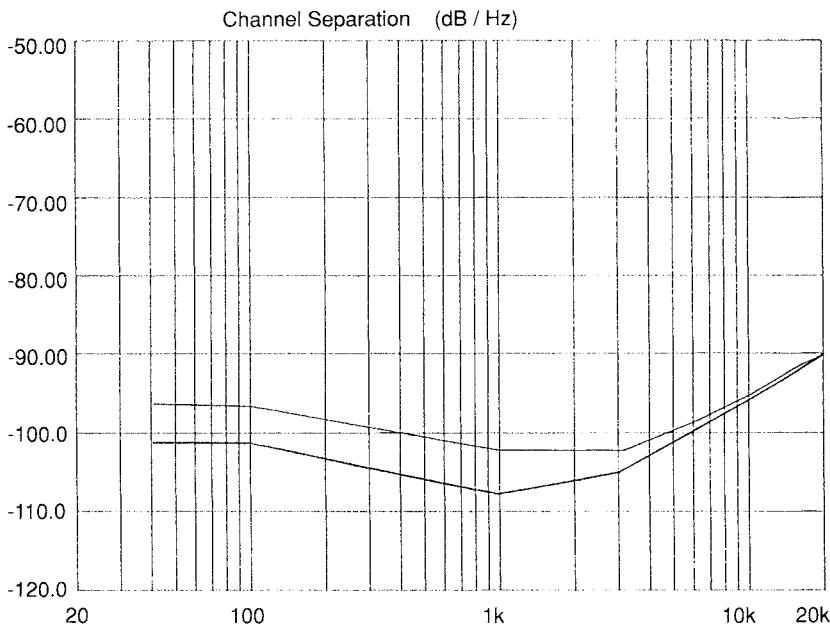


Figure 5.6     *Channel separation measurement.*

ically, distortion results if the device adds new harmonic information. For example, if a 50 Hz tone was applied, 50 Hz would be reproduced, as well as harmonics: 100 Hz second harmonic, 150 Hz third, 200 Hz fourth, etc. Devices such as D/A converters are often guilty of contributing harmonic distortion.

The THD + N of a device is measured by inputting a sine wave (a signal with no harmonics), then filtering out the sine wave at the output, and measuring the remaining harmonic signal as a percentage of the original signal. A THD + N analyzer accomplishes the process directly, as seen in figure 5.7. THD + N varies with respect to both amplitude and frequency. It must therefore be measured over a wide frequency range and at different levels. Unlike analog systems, in which THD + N increases with amplitude, a digital system's THD + N decreases with amplitude. CD player manufacturers wage war over THD + N measurements, often achieving levels below 0.002 percent.

The THD + N can be measured at frequencies recorded at 0 dB and lower amplitudes such as -24 or -30 dB. Some players show a rise in distortion at high signal amplitudes, due to distortion in the analog output stages, as well as bit weighting errors in the D/A converter which increase quantization error. Steps in the measured distortion over amplitude in-

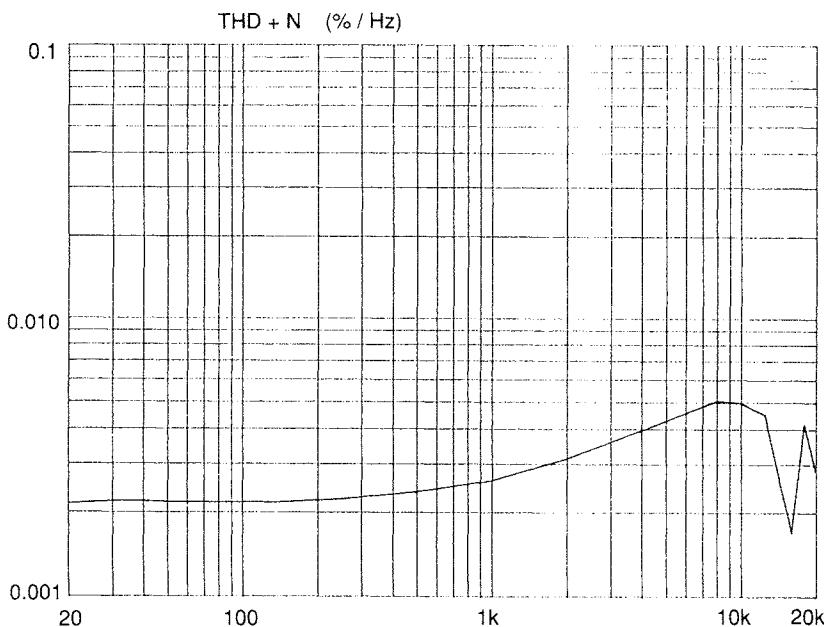


Figure 5.7 *Total harmonic distortion plus noise (THD + N) measurement.*

dicate the latter. A rise in THD + N at high frequencies is often observed, particularly in players with poor low-pass filters. This is usually not due to harmonic distortion but is caused by the beat frequency between the CD player's clock and the test signal. With a 44.1 kHz clock and 20 kHz test tone, for instance, the spurious signal will measure 24.1 kHz.

The EIAJ and most manufacturers insert a 22 kHz low-pass filter between the player's output and the distortion meter. Otherwise, spurious beat frequency components related to the player's sampling frequency add to the distortion measurement, whereas in fact they are not harmonic distortion. While the use of a filter is thus justified in THD + N measurements, it limits the use of the measurement; for example, a 24 kHz low-pass filter removes all legitimate harmonic components for frequencies under test greater than 12 kHz (24 kHz is the second harmonic of 12 kHz). Thus THD + N cannot be accurately measured for frequencies greater than 12 kHz. Thorough reviewers are careful to examine the unfiltered output signal with a spectrum analyzer to detect any components, harmonic or otherwise, which might degrade performance. Because of the low distortion levels to be measured, care must be taken in grounding and interconnecting the player to the test equipment.

### ***Quantization Noise***

Quantization noise can be evaluated by measuring distortion at very low signal levels. To measure quantization noise, a dithered test signal at various low-level amplitudes is used. THD + N is measured, thus removing the fundamental component of the test signal. In addition, low- and moderate-order harmonics may be removed with a high-pass filter (400 Hz) to prevent distortion products from obscuring the quantization noise measurement across the audio spectrum above 400 Hz. Figure 5.8 shows quantization noise in dB (y axis) against signal amplitude from maximum down to -90 dB (x axis).

### ***Intermodulation Distortion***

Intermodulation (IM) distortion measures the creation of new frequencies from a combination of any legitimate input frequencies. Some circuits generate a series of sum and difference frequencies related to two input frequencies. For example, if 50 Hz and 1,000 Hz are input, 950 Hz, 1,050 Hz, 1,100 Hz, and an entire series of frequencies might be added in. Since these frequencies are not harmonically related to the originals, they tend to be audibly disagreeable and might contribute to listener fatigue. To measure IM, twin sine waves most often recorded at 0 dB are input to the device and then filtered at the output; any remaining intermodulation signals are measured and expressed as a percentage of the original signal. Several standard frequency pairs can be used. A SMPTE IM distortion standard specifies 60 Hz and 7 kHz signals in a 4:1 amplitude ratio. A CCIF IM distortion standard specifies 11 and 12 kHz in a 1:1 ratio. In either case, the percentage of the beat frequency at the player's output constitutes the measurement. A low-pass filter is usually inserted as part of the measurement. A good player's SMPTE IM distortion should be 0.004 percent or less. A spectral analysis of the IM products may yield insight into their cause.

### ***Distortion Test Frequencies***

Distortion measurements in CD players point up a unique problem: Certain test frequencies must be avoided because they yield unrepresentative (and higher) distortion readings. Specifically, frequencies which are integral submultiples of the 44.1 kHz sampling frequency—they are thus synchronous with the clock frequency—will utilize a fewer number of quantization intervals than noninteger submultiples, resulting in higher distortion. The synchronous relationship does not test linearity of every state of the D/A converter because the points where the test waveform are sampled will repeat after a relatively small number of samples. A non-

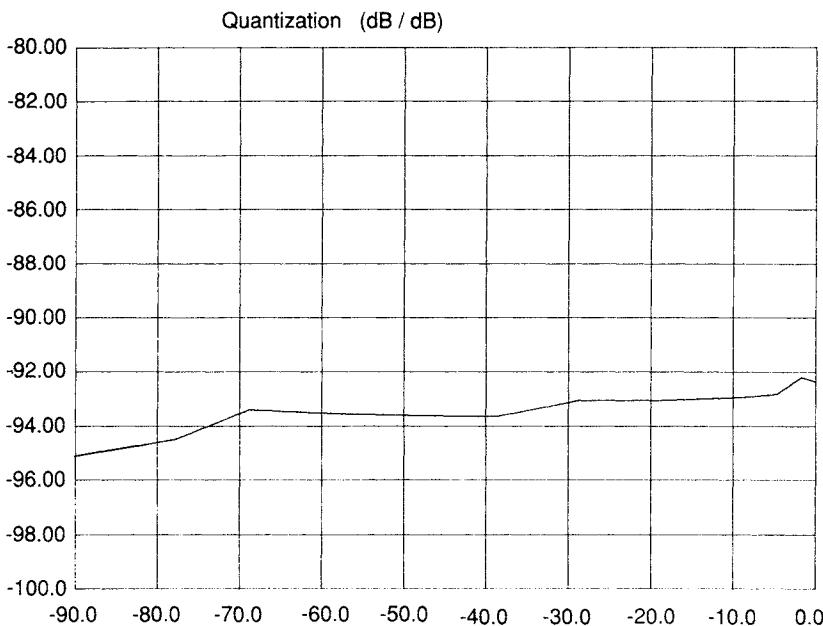


Figure 5.8    *Quantization noise (THD versus level) measurement.*

synchronous relationship is preferred because the waveform passes through many timing relationships, causing every converter state to be exercised. For example, a test frequency of 315 Hz (1/140 of 44.1 kHz) recorded at -20 dB would use 71 different quantization intervals in a period of 100 milliseconds in a 16-bit converter, whereas a 317 Hz signal of the same amplitude and over the same time period uses 3,282 intervals.

Thus the EIAJ and manufacturers recommend avoiding test frequencies that are integer submultiples. Hence new test frequencies have been standardized by the EIAJ. These are presented in table 5.2. Measurements of criteria other than distortion, such as frequency response, are not affected by the anomaly.

### **Signal-to-Noise Ratio**

Ideally, an audio device should have no output signal when there is no input signal. In practice, the S/N (signal-to-noise) ratio measures the noise floor of the player's output relative to the player's rated maximum output level (0 dB). The number of bits in a digital audio system determines the maximum signal-to-noise ratio possible, but the analog output circuitry is often the limiting factor. Any noise or hum in the output stages will de-

Normal Frequency (Hz)	Actual Frequency (Hz)	Difference (percent)
4	4	-
8	8	-
16	17	+ 6.250
32	31	- 3.125
63	61	- 3.175
125	127	+ 1.600
250	251	+ 0.400
500	499	- 0.200
1,000	997	- 0.300
2,000	1,999	- 0.050
4,000	4,001	+ 0.025
8,000	7,993	- 0.088
10,000	10,007	+ 0.070
12,500	12,503	+ 0.024
16,000	16,001	+ 0.006
18,000	17,989	- 0.061
20,000	19,997	- 0.015

Table 5.2 *Standardized non-integer submultiple test frequencies.*

termine the floor over which the signal may range. To make the test, a test disc track with 16 zeros per sample (digital silence) is played, and the noise which is measured is referenced to a maximum level of 0 dB. An S/N measurement with a digital silence track does not provide information on the player's digital circuitry since the D/A converter is not being exercised (and its output is muted in some designs). The digital silence track is thus useful for separating analog noise from digital noise sources. Although the D/A converter is quiescent, and thus contributes no noise due to conversion, low-frequency noise sources, such as the power supply, the analog circuitry, and poor grounding, as well as high-frequency noise sources, such as clock leakage, will limit the noise floor of the player. A spectral analysis of the noise signal is often revealing.

Once again, there are different measuring methods. Some manufacturers measure a broadband S/N ratio, while others employ a low-pass

filter to eliminate out-of-band components. (The EIAJ calls for a filter with cutoff at 4 Hz and 20 kHz and with an attenuation of at least 60 dB at 24.1 kHz.) An “A” weighted filter may be used to attenuate low-frequency components of the noise signal. A good player should have an S/N of greater than 110 dB, unweighted, at 1 kHz, as shown in figure 5.9. Careful attention to grounding and interconnections is required to measure such low levels.

### ***Dynamic Range***

Dynamic range is often confused with signal-to-noise ratio, but there is a difference. The EIAJ defines dynamic range as the level of total harmonic distortion plus 60 dB when a CD reproduces a 1 kHz signal recorded at -60 dB below maximum level. For example, a THD + N reading of -25 dB plus 60 dB yields a dynamic range measurement of 85 dB. The measurement is thus not equivalent to S/N ratio. The EIAJ specifies that an “A”-weighted filter and 22 kHz low-pass filter be used to measure THD + N. This produces a noise measurement as opposed to a distortion measurement, and the results can be compared to “A”-weighted noise measurements. However, it may be misleading to measure only harmonic components; for example, D/A nonlinearities may produce alias frequencies. A more rigorous test is to measure the total power of all harmonic, nonharmonic, and noise components relative to the fundamental. The EIA has proposed means of measuring dynamic range using 500 Hz dithered test signals recorded at -60 and -120 dB. The player’s output is passed through a one-third octave filter tuned to 500 Hz, and the output is measured in dB. The noise floor is determined by adding 3 dB to the output signal level from the -120 dB recorded level. This signal level is assumed to be the lowest signal that can be reliably distinguished from noise. Maximum signal level is thus 60 dB above the output signal level of -60 dB. Dynamic range is ascertained by subtracting the noise floor (-120 dB recorded + 3 dB) from the signal level (-60 dB recorded + 60 dB).

### ***D/A Linearity Measurements***

As we observed in Chapter Four, D/A converter linearity is a crucial concern in a digital audio system. No matter what word length is converted, the accuracy of the conversion and hence the fidelity of the audio signal hinge on the linearity of the converter. A linearity test measures the D/A converter’s ability to reproduce various signals, particularly low-amplitude signals, at the proper amplitude. Specifically, when a bit changes from 0 to 1 in a D/A converter, the analog output must increase exactly by a proportional amount. Any nonlinearity in D/A conversion will be apparent as a deviation from a nominal amplitude and will result in dis-

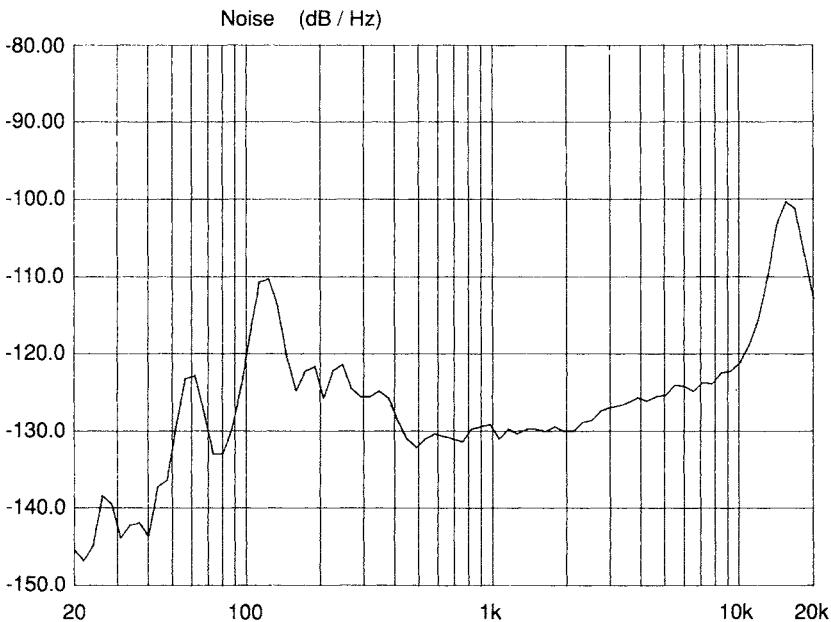


Figure 5.9     *Signal-to-noise ratio (unweighted) measurement.*

tortion of the audio signal. Multibit converters exhibit differential nonlinearity due to bit weighting errors; thermal or physical stress, aging, and temperature variations are also factors. Low-bit converters are more robust in this regard.

A linearity test measures the reproduced analog signals relative to their recorded amplitudes. The test is often performed with 500 Hz test tones, with descending test disc amplitudes of  $-1, -3, -6, -10, -20, -30, -40, -50, -60, -70, -80, -90$  dB, and  $-100$  dB, with the greatest scrutiny of levels below  $-60$  dB. Alternatively, a fade-to-zero test may be used, in which the tone fades from approximately  $-60$  dB to  $-120$  dB, where the signal disappears into system noise. It is important to note the difference between dithered and undithered test tones when making this test. For example, because of averaging, an undithered test tone of  $-90$  dB will be reproduced at  $-89$  dB through a perfect D/A. When a dithered test tone is used, the fundamental level, as obtained from a spectrum analyzer, should fall precisely at  $-90$  dB. Careful attention to grounding and interconnections is required to measure such low levels. In some cases, D/A converter nonlinearities can be uncovered by ear. This is done by playing a dithered fade-to-zero signal. Effects of nonlinearity are audible as harmonic distortion.

Using tests for D/A linearity, the results of poorly performing D/A converters can be evaluated. For example, a  $-80$  dB signal reproduced at  $-85$  dB would result in expansion of low-amplitude information. A  $-80$  dB signal reproduced as  $-76$  dB would result in compression of low-level information. A D/A converter with extreme nonlinearity problems might compress the signal through  $-80$  dB, then expand the signal at  $-90$  dB, reproducing low-level audio waveforms in a highly nonlinear fashion. Ambient information, for example, could be audibly affected. Figure 5.10 shows the output of a CD player in which the D/A output is quite linear through  $-90$  dB as measured with a dithered fade-to-zero linearity test. It would be safe to say that this nonlinearity would not be audible under most listening conditions.

To help equipment manufacturers use D/A converters to their best advantage, many multibit D/A chips provide a means of calibration. Some D/A chips offer calibration of the MSB, while others may offer calibration of the 4 most significant bits. As a practical tip, remember to let the D/A converter warm up for a half-hour or more before attempting any calibration; its linearity changes as the device reaches a stable operating temperature.

### ***Phase Measurements***

Interchannel phase difference measures the shift between identical frequencies recorded on two channels. The stereo signal must be recovered from the single data stream and reconstructed with the original simultaneous phase between the two channels. If a single D/A converter is used, and if compensation isn't applied, the channels will be out of phase with respect to each other as the converter delays one channel relative to the other. For example, a 44.1 kHz clocked player with a single D/A may show an 82 degree phase shift at 20 kHz (11.34 microseconds at the 44.1 kHz clock rate). Similarly, an 88.1 kHz clocked player may show a 41 degree shift, and a 176.4 kHz clocked player would show about a 20 degree phase shift. Time delay compensation used in single D/A players may improve this phase shift. Even dual D/As can contribute a slight error.

To measure any interchannel phase error, 20 kHz is output to each channel; any difference will be apparent from the misaligned waveforms, such as the ones in figure 5.11a. A good player should have negligible interchannel phase error. This is illustrated in figure 5.11b. Alternatively, time delay or phase error between channels can be measured through a Lissajous pattern. The output of one channel is connected to the horizontal (X) input of an oscilloscope, and the output of the other channel is connected to the vertical (Y) input. A trace angled at 45 degrees (from lower left to upper right) indicates the channels are in phase.

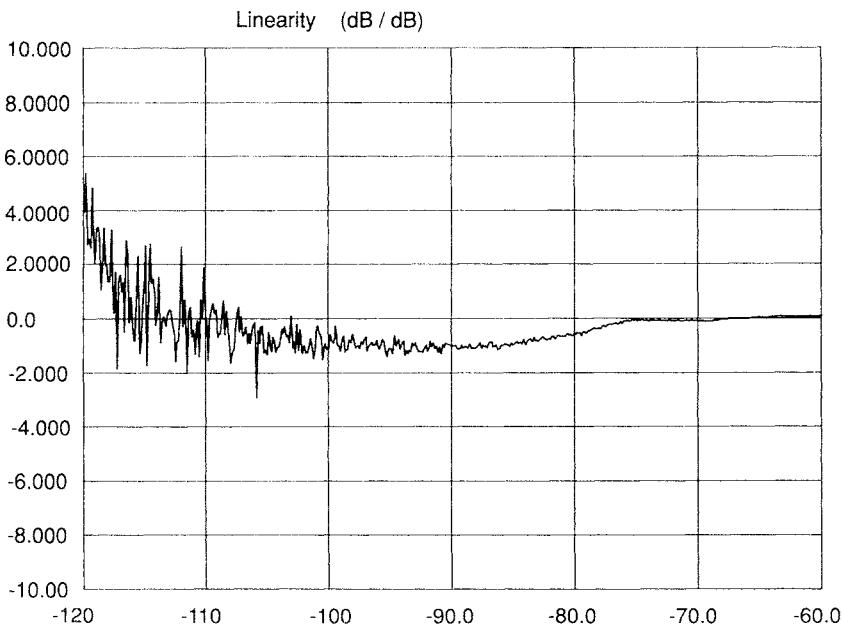


Figure 5.10     *Fade-to-zero linearity measurement.*

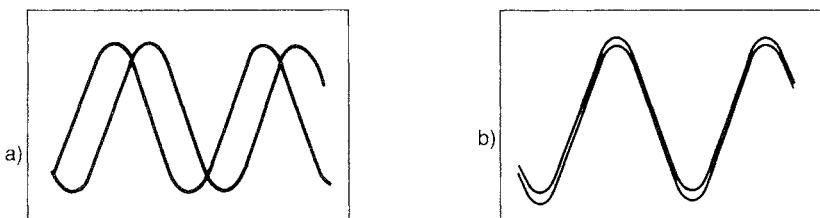


Figure 5.11     *Interchannel phase difference checked at 20 kHz: phase error (a); phase linearity (b).*

Phase linearity measures any phase shift between different frequencies. Digital filters and carefully compensated analog filters can reduce phase error to a few degrees or less. Although this is not a rigorous test, some test discs contain twin tones of 2 kHz and 20 kHz, each on a single channel output. The tones are encoded with simultaneous positive-going axis crossings. Any relative phase discrepancy can thus be observed on an oscilloscope. Simultaneous zero-crossing of the positive-going waveforms shows there is no phase discrepancy. An offset of the waveforms at that

point would indicate a discrepancy. A CD player with digital filtering might show phase linearity of less than 0.1 degrees at 20 kHz.

### *De-emphasis*

To reduce noise, some compact discs are recorded with pre-emphasis, a boosting of high frequencies as described in Chapter Three. During playback, a high frequency de-emphasis circuit should automatically restore a flat response. De-emphasis error measures any deviation occurring when the de-emphasis curve is not the correct reciprocal of the pre-emphasis curve. To accomplish the test, selected frequencies are output, and any de-emphasis deviation is measured in dB. The EIAJ proposes use of 125 Hz, 1 kHz, 4 kHz, 10 kHz, and 16 kHz.

### *Other Evaluations*

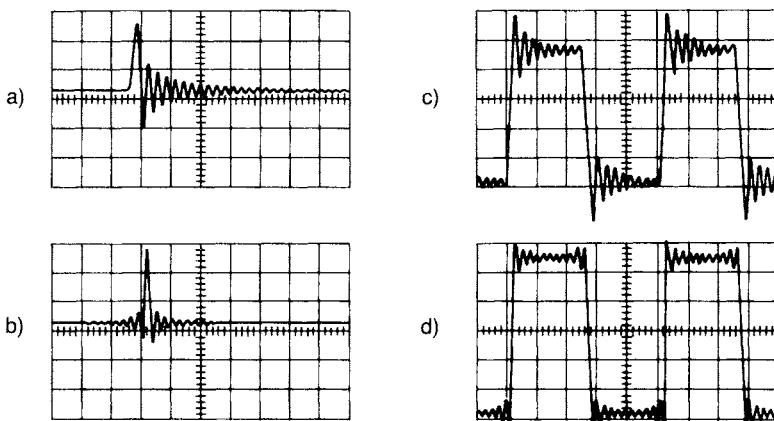
Any level imbalance between the two output channels is measured by comparing two 1 kHz signals recorded at 0 dB. The difference is expressed in dB. In addition, voltage readings of the maximum output values are often specified.

Although the wow and flutter in CD players should be negligible, some test discs supply a 3.15 kHz tone, the frequency usually used to measure speed fluctuation in analog equipment, to accomplish the measurement for CD players. The EIAJ specifies that the test be made for at least 5 but no more than 30 seconds. Typically, the weighted peak value of CD wow and flutter is below the  $\pm 0.001$  percent measurement threshold, and most manufacturers simply state, correctly, that it is below measurable limits.

Although an unlikely occurrence, any pitch error introduced by a CD player may be specified by measuring a 20 kHz output signal with a frequency counter and expressing the error as a percentage:  $(F_1 - F_0)/(F_0) \times 100$ , where  $F_0 = 20$  kHz.

The nature of the output filter can be examined through single pulses and square waves. A single pulse or a 1 kHz square wave will show the damped pattern of analog filtering or the symmetrical pattern of digital filtering as shown in figure 5.12. However, it is important not to misinterpret the ripple present on the waveform. A perfectly square square wave could only exist in the case of infinite bandwidth—that is, only if an infinite number of sine wave harmonics were present to completely compose the square wave. In the case of a CD player, we know that our bandwidth is limited to about 20 kHz. Thus the square wave will show ripples. This is normal operation and is described as the Gibbs phenomenon.

The single pulse response is also useful to uncover incorrect absolute polarity. For example, if a positive-going pulse recorded on the test disc results in an oscilloscope trace showing a negative signal, the player's absolute polarity is reversed.



**Figure 5.12** *Output filters can be examined through single pulse and square waves: analog filter impulse response (a), digital filter impulse response (b), analog filter 1 kHz square wave (c), digital filter 1 kHz square wave (d). (Courtesy Stereo Review)*

The EIAJ also specifies a test to measure access time, the time it takes for a pickup to move to a location on the disc. Short and long access times are specified. Short access time measures the time for the pickup to move to the next adjacent track and begin playing. Long access time measures the time for the pickup to move from the innermost track to the outermost track and begin playing. Both measurements are specified in seconds.

Test discs can be used to exercise the error correction/tracking ability of the player. The Philips 5A test disc contains three types of intentional defects on specified tracks: simulated fingerprint, calibrated black dots on the readout side, and calibrated interruptions on the information layer. The dimensions of the black dots are 300, 500, 600, and 800 micrometers, and the interruptions on the information layer are 400, 500, 600, 700, 800, and 900 micrometers. These obstacles help determine a player's limitations in error correction, concealment, and tracking. All CD players should be able to play through the fingerprint and the smallest of the other defects. In theory a player with fully implemented error correction strategy (not all have this) should correct any defect up to 2400 micrometers. In practice most players handle all the defects on the Philips test disc.

The more strenuous defects in the Verany 788032 disc may be used to determine the player's limits. This disc contains seven test conditions designed to evaluate the mechanical, digital, and processing performance of a CD player. The first series of tracks is recorded with linear velocity ranging from 1.2 to 1.4 meters/second. The second series of tracks combines maximum and minimum cutting velocities with maximum and min-

imum track pitches. One track in the test is recorded with a continuous velocity variation. The third series of tracks contains variations in track pitch ranging from 1.5 to 1.7 micrometers. The fourth series of tests can be used to measure the detection level of the RF signal from a pickup by simulating variations in cutting intensity. Asymmetry is created by varying the ratio of pit/land amplitude. The fifth series of tests evaluates a pickup's dropout performance. Nondata areas ranging from 0.05 millimeters to 4.0 millimeters are traced across the disc surface. The CD standard specifies correction of 0.2 millimeter dropouts, but the CIRC is theoretically capable of correcting dropouts of up to 2.47 millimeters. Interpolation of 8.5 millimeters dropouts is theoretically possible. The sixth series of tests combines dropouts with minimum track pitch. Dropouts ranging from 1 to 2.4 millimeters are recorded at a pitch of 1.5 micrometers. The seventh test can be used to evaluate performance in the face of successive dropouts. Some players may be able to correct a single dropout but fail to correct two dropouts. The test contains two successive dropouts, each ranging from 0.1 to 3 millimeters. Because many of these test conditions exceed the expectations of the CD standard, error correction and tracking performance may be critically evaluated.

### Caveats

The use of low-pass filters to limit measurements to in-band information has generated confusion. Some players output only in-band information with no spurious high-frequency components. Specifications such as THD + N can be measured directly. However, other players exhibit out-of-band frequency components. The rationale for bandlimited measurements is that supersonic components cannot be heard and should not therefore affect the measurement.

The difference between unfiltered and filtered measurements is illustrated in figure 5.13. The first oscilloscope trace of the unfiltered player under test shows a 20 kHz signal and an image frequency at 24.1 kHz (44.1 kHz–20 kHz) as well as images of those two frequencies at 64.1 kHz and 68.2 kHz. The second trace shows what happens when a 20 kHz seventh-order low-pass filter is put on the player's output; the troublesome images are largely attenuated, leading to much better measurements.

We should note that some critics are not so sure of the validity of filtered measurements. Since many players measure satisfactorily without a special brickwall filter, is it fair to use special measuring conditions for the benefit of other players? Moreover, it is possible that some listeners can hear out-of-band components or that the components might affect in-band signals. If a player has them, measurements should show that fact. Finally, unless the manufacturer wishes to supply the consumer with a special low-pass filter, it doesn't seem right to measure the player *with* one.

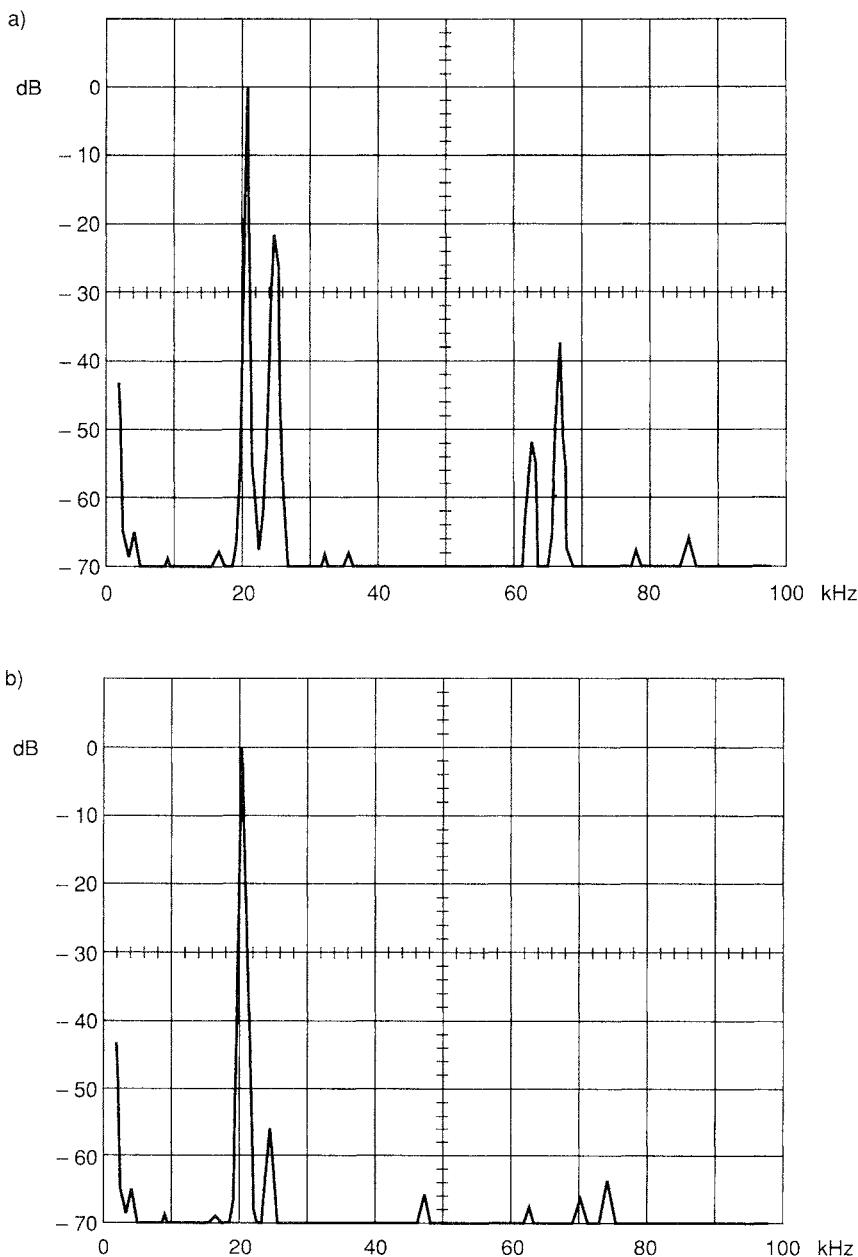


Figure 5.13     Effects of output filters used in making measurements: unfiltered spectra (a) and filtered spectra (b).

The bottom line is that measurements are valid only if they accurately portray performance. If some players have higher THD + N, then that should be made clear. If THD + N measurements are made with a brick-wall filter, then a spectrum analysis should also be supplied with the player to show the supersonic components. Honesty in specification is important, especially in a new technology such as digital audio. Only then can we expect evolution toward better performance.

A specification is meaningful only if we understand the conditions of measurement. Greater consumer awareness will lead the way to new specifications designed to clarify the performance of digital audio devices and to help supplement the pair of precision measuring devices you're naturally equipped with.

## BUYING A PLAYER

Are all CD players alike? The answer is definitely no. Some players are a lot more expensive than others, some players sound a lot better than others, and the correlation isn't always what you'd expect. Exactly how do you shop for the best buy? Let's consider a list of effective maneuvers for smart CD shopping—questions that will separate the great players from the merely good.

### *Basic Questions*

“Is this CD player free?” This question is often overlooked by many shoppers. I've found that it never hurts to ask. Moreover, if the answer is yes, you can skip the remaining questions.

Who is the manufacturer? This is a good overall measure of quality in CD players. After evaluating many players, one finds consistent quality among certain manufacturers, whereas other manufacturers may have hits and misses. Surprisingly, expense isn't always the reason. Do your homework and ascertain exactly which companies have the required expertise; read reviews in magazines and ask a lot of questions of knowledgeable people. It isn't always obvious.

Does the player have the features you want, or are you paying extra for gimmicks you'll never use? All players have the basic controls to start, stop, and move the optical pickup around, and many are programmable, but there are a lot of differences: remote control, multidisc playback, headphone output, digital audio output, and so forth. Maybe a portable player better suits your needs. However, at a given price, if the manufacturer added a feature, they probably sacrificed somewhere else. Some of the most important features include track number and index display, disc/track/elapsed/total time display, track and index programmability, disc

and A-B repeat, multispeed manual search, random track playback, headphone jack with volume control, digital output, remote control with volume control, multidisc capability, and disc memory library.

Some players offer dynamic range compression. The high dynamic range of CDs presents a problem in environments with a low dynamic range. A compression circuit reduces the music's dynamic range to help compensate for high ambient noise conditions. This feature is particularly useful in a car player or in a home player used to copy discs to tape for car listening.

Are the ergonomics well designed? Make sure the controls are easy to use and largely self-explanatory. Try out the controls and make sure the player does everything you expect it should. Similarly, verify that the display provides all the information you want. While you're at it, evaluate the player's responsiveness to your commands; some players have a considerable lag between your command and their response. In time, this might become very annoying.

Is the player's construction robust? Is the chassis made of metal or plastic? Do the disc drive and drawer operate smoothly and quickly, without a lot of racket? If you can manage a peek inside, are the circuit boards and components securely mounted? Does the disc drive appear to be solidly constructed?

### *Performance Questions*

Is the impact immunity good? When the sales people aren't looking, whomp the player a few times on the sides and top while it's playing. Good pickup servos are amazingly resistant to shock. Obviously, this test is even more crucial for a portable player.

How good is the error correction? Discs aren't indestructible, especially as they enter the harsh environments of car and portable players. Unfortunately, consistent specifications for error correction performance are not commonly available. Although error correction test discs are available, they are sometimes hard to find.

Even without a special test disc, it is possible to gauge a player's ability to play through errors without mistracking. You can make your own "torture disc" to test for tracking ability: take a strip of tape perhaps one millimeter wide and lay it radially across a disc data surface. A good player should play through it. Now try two strips and see if it plays. Keep loading them on until the player capitulates. While you're at it, try skipping around on the torture disc; some players are able to play straight through errors but fail to locate damaged tracks.

Such a stripped disc simulates errors but cannot be used to fully determine a player's ability to track real-life defects. A player might perform

poorly with the strips but do much better with scratches and dirt. Ideally, a badly damaged and soiled CD would make a better torture disc.

If you're caught without your torture disc, lay a few juicy fingerprints on any disc. This will stress a player with a real-life problem and may cause audible errors. Every player hiccups occasionally when it encounters a serious defect—evaluate its tenaciousness relative to other players. Repeat your impact immunity test with a torture disc; a pickup may have more trouble tracking a damaged disc.

Of course, even a normal disc contains errors, and the player's error correction and concealment circuitry must suppress their audibility. Correction strategies provide for absolutely correct data so that the error becomes nonexistent and inaudible. However, error concealment synthesizes data to the best of the circuit's ability. Because not all pickup/servo systems are equal, players' abilities to read disc data varies; the poorer the data quality, the more concealment must be used. Concealment circuits aren't equal, either. The bottom line is audible errors—on some players you hear them once in a while. It isn't a mute or a click, just badly concealed data, and the distortion may be audible for an instant. On good players, you don't hear that very often.

Are there vibration problems? Airborne and structureborne vibrations from loudspeakers (CD listeners tend to play back at louder levels, and with augmented bass response) may cause vibration in discs that are not tightly clamped to the spindle motor shaft. As the disc vibrates, the pickup servos are worked harder, allowing greater chance for mistracking, and the result may be concealment distortion at loud listening levels. Even if the pickup tracks properly, its constant correction movements might draw excessive current from the power supply, leading to distortion in the analog output circuitry.

Checking for this problem in stereo showrooms is fun indeed. If you crank up the amplifier level, the problem may occur, but it might not audibly manifest itself because amplifier or loudspeaker overload will mask the problem. You have to crank up the level and then listen with closed-ear headphones. This is guaranteed to drive the sales staff crazy. Listen closely. On some players you may hear a deterioration of sound quality at high listening levels. It's difficult to describe the effect, but you'll know it when you hear it, if you keep an open mind.

### ***Design Questions***

Does the player have separate power supplies for the servo, digital, and analog circuitry? This alleviates the problem created when the servo system draws power intended for other circuitry and also helps isolate one circuit from another, reducing crosstalk, in the form of noise and interference, between different player circuits.

Does the player use a one- or three-beam pickup design? Don't over-concern yourself with this one. There is little difference in performance between one-beam and three-beam designs, except on a one-to-one basis. The above tests determine the quality of the design. If it's a toss-up, opt for the one-beam type; it is inherently simpler, as we have seen, and thus more likely to stay aligned over time.

What is the digital filter's oversampling rate? In theory, there is no absolute reason why a higher oversampling rate should sound better than a lower rate. In practice, a higher oversampling rate moves the first image spectrum to a higher frequency, hence a lower-order analog filter can be used; this should improve sound. On the other hand, a higher oversampling rate demands better (faster) D/A converters and demands more critical design of the digital filter. A four-times oversampling rate places the image at 176.4 kHz, and a third-order analog filter can sufficiently attenuate the image. Moreover, such analog filters can be easily designed without audible artifacts. Thus generally a four-times rate provides all the benefits of oversampling, without any other complications. On the other hand, many critics are particularly enamored of the eight-times oversampling variety.

Does the player have dual D/A converters? When the output data is multiplexed through a single D/A converter, its analog output signal must be alternately switched between the stereo channels. As discussed, a time delay between left and right channels is created, resulting in an inter-channel phase difference. Dual D/A converters do not cause this phase error between channels. On the other hand, a pair of converter chips might not be evenly matched. Ideally both D/A converters should be placed on the same chip, thereby ensuring good collaboration. Otherwise, dual D/A converter chips are still better than a single multiplexed chip.

The D/A converter design is particularly important. As we have seen, low-level linearity is crucial for successful reproduction. In theory, an 18- or 20-bit converter is better than a 16-bit model, but only if its measured performance verifies the linearity of the converted bits. In an effort to address linearity problems in multibit converters, manufacturers have introduced low-bit D/A converters (sometimes referred to as 1-bit converters). These converters are discussed in detail in Chapter Four.

Questions concerning pickup design, oversampling rate, and D/A converter design ultimately must be decided on a case-by-case basis. Take all advice (including this book's) with a pinch of salt, and let your own ears decide.

### ***Specification Questions***

Do the specifications measure up? Although mid-priced players often have similar specifications, small differences begin showing up in high-priced

players. In addition, independently measured specifications often reveal weaknesses. Look for a frequency response deviation of less than 0.3 dB, unfiltered total harmonic distortion plus noise (THD + N) of 0.0025 percent at 1 kHz and 0 dB, channel separation greater than 90 dB at 1 kHz, signal-to-noise ratio of 110 dB (A-weighted), intermodulation (IM) distortion of 0.004 percent at 0 dB, near-perfect interchannel phase linearity, and low-level linearity error of less than 2 dB at -90 dB.

One specification problem to be aware of is supersonic artifacts (or at least those above 21 kHz). Beat frequency components caused by clock interactions and other high frequency signals can increase broadband distortion and decrease S/N values. When making measurements, most manufacturers insert brickwall filters to keep these spurious signals out of their published specifications. A good player design uses synchronous clocks so that beat frequencies are never generated.

### ***The Bottom Line***

Does the player sound good? This question is tougher to answer than you might think. In general, most CD players sound okay. Most showrooms have always been a tough place to listen to anything other than the sales pitch. With digital audio, it's even rougher. The trick is to concentrate on one aspect of sound quality at a time. Play a familiar disc. Listen for badly corrected errors. Listen for stable localization. Listen for high-volume disc instability. Listen for full ambience with "air" around the music. Listen for naturally warm sound. Listen for tight bass response, smooth treble response, and low background noise. Relax. Take your time. Try several discs—one disc might bias your decision according to the type and quality of recording you're hearing. Some factors are difficult to evaluate with a music program. You might want to invest in a test disc for more critical evaluation. For example, deviations in frequency response are often exposed when one listens to a noise signal.

When comparing players, compare only two at a time; otherwise it gets too confusing. Make sure they are playing at the same loudness level, or you'll tend to favor the louder player. Alternate between different loudspeakers, or else the loudspeaker's own coloration will bias your decision. Critical listening isn't easy, and it takes time. Sometimes it takes hours of comparison before a person can reliably choose the better player. As we all become more familiar with what we're listening for in digital audio, we'll get better at it.

A shopper typically checks the player's features, construction, measured specifications, and, finally, sonic quality. In reality, though, the first thing you should do is listen before your evaluation is influenced by knowledge of factors such as construction quality, the manufacturer's reputation, the price, and the specifications. While auditioning, you might even

cover up the unit so that its front panel styling won't influence your decisions. Whenever possible, make double-blind listening comparisons in which the players' identities aren't revealed until you have recorded your results.

In summary, we can list the important evaluation criteria, roughly ranked in order of importance: sonic quality, robust construction, error correction and concealment performance, specifications, oversampling filter, dual D/A converters, separate power supplies, immunity to air and structure-borne vibration, digital output, manufacturer, features, ergonomics, impact resistance, and vibration-damping chassis and parts.

## **PLAYER NOTES**

Unlike plants, animals, lawns, and fine woolens, CD players require little care. Nevertheless, it is important to treat your player with at least a little respect. It is, after all, the most sophisticated piece of audio electronics to reach the home. The laser optics, servo system for automatic focusing and tracking, control microprocessors, and output digital-to-analog circuitry are all high, high tech. On the other hand, they are amazingly robust, as digital electronics are wont to be. With a little consideration, your player will consistently and without complaint continue to deliver its 4.3218 million bits per second.

### ***Installation***

When installing a CD player, avoid heat sources, including radiators and even power amplifiers. Heat is circuitry's worst enemy: It causes premature aging and failure. Ventilation is circuitry's best friend. Make sure air can flow over, under, and around the back of your player. Avoid placing the player on rugs, sofas, or in stacks of equipment. Racks, bookcases, cabinets, and built-in installations are certainly convenient, but unless they are well engineered, they can spell trouble. Heat build-up can happen all too easily. Avoid racks unless they contain fans to move the air. Otherwise, the thermal stress will shorten the component's lifespan.

Also avoid areas subject to direct sunlight, excessive dust, cold, moisture, mechanical shock, or vibration. Keep the player away from sources of 60 Hz electrical hum, such as transformers or motors. Of course, AC-powered CD players should never be used near water.

Hooking up a CD player is an easy task. First, check for transport screws. Many players have one or more screws (or buttons) on the bottom panel that are used to lock the optical pickup in place to prevent damage during transport. The pickup must be unlocked before you attempt to play a disc.

The electrical connections are located on the player's back panel. Most players have two cords to deal with: AC power and audio. If you're the cautious type, check to make sure its power requirements correspond with that supplied by the country you're in (e.g., 120 volts, 60 Hz in the U.S.A.). The player's power requirements are printed on back or clearly marked on players with switchable power supplies.

An accessory that is highly recommended is an AC line noise filter and surge suppressor, sometimes called a spike filter. Invest your money in a high-quality unit, perhaps shopping for it in a computer store. It should be inserted in-line prior to all of your audio components. No, it won't improve frequency response or clean up error concealments, but next time your power company delivers a wad of bad voltage, it might save you from a big repair bill.

The player's analog audio outputs contain the left and right audio signals; most players use two phono connectors. In some cases, the level of the audio outputs can be varied from the front panel. Some players have a digital output that accepts a single phono plug which carries both audio channels to an external D/A converter. Some players also feature a digital audio optical connector. A fiber optic cable is used to convey data from the player. Often there is also a digital data output on/off switch. Digital data should be switched on only when you use this output; otherwise, it could interfere with the analog output.

The analog outputs should be connected to your preamplifier's CD inputs. A digital output should be connected to the input of an external D/A converter. The converter's analog outputs are then connected to a preamplifier. Of course, switch off system power before making or breaking any connections. Finally, take time to study the owner's manual accompanying your player; you'll probably uncover some subtlety you would otherwise have overlooked.

### ***Other Considerations***

The pickup's laser diode produces a strong electromagnetic field. This is harmless for human lifeforms but may affect magnetic media such as audio or video cassettes, floppy disks, or even some wristwatches. In other words, don't store cassettes on top of a CD player. The field is probably significant only when the player's shielding covers have been removed, but you never know until it's too late.

The digital processing circuits inside a CD player, with their clocks clocking in the megahertz range, make ideal radio stations, jamming channels used for legitimate radio and television broadcast. CD players must be tested and must comply with Federal Communications Commission guidelines for a Class B computing device; that is, they must reasonably limit the amount of radio frequency interference they generate. In gen-

eral, a player will not broadcast signals strong enough to bother your neighbors but might very well interfere with your own reception. Several measures can be taken to minimize the trouble: reorient the receiving radio or television antenna, relocate the player with respect to the receiver, move the player away from the receiver, or plug the player into a different AC outlet so that the player and receiver are on different branch circuits. Finally, of course, interference can be terminated by turning the CD player off.

Moisture condensation can cause temporary problems. If the player is brought in from a cold to a warm location, or is placed in a very damp room, moisture may condense on the lenses inside the pickup. Should this occur, the player may not operate. Fortunately, the cure is quite inexpensive: Remove the disc and leave the player turned on for about an hour. The heat generated will evaporate the moisture. Discs with condensation can be wiped dry (always wipe a disc with radial motions) and played immediately.

Although some components, such as preamplifiers and power amplifiers, may be left on—indeed, professional audio gear is specifically designed to be powered indefinitely—a consumer CD player is best turned off when not in use. This will extend the life of the player. On the other hand, some audiophiles believe that CD players require a thorough warmup (an hour or so) before their sound quality becomes optimal. To avoid the wait, they keep their player's power on. This may have merit. For example, D/A converter manufacturers recommend a one-half hour warmup before service technicians calibrate their devices.

The outer cabinet, front panel, and controls may be cleaned with a soft cloth lightly moistened with a mild detergent solution. Do not use any type of abrasive pad, scouring powder, or solvent such as alcohol or benzine; the latter can melt or discolor cabinet paint. If any liquid should spill into the player, unplug the unit and have it cleaned and checked by qualified personnel. Incidentally, soft drinks such as colas are particularly renowned for their ability to quickly corrode contacts and connectors.

The pickup's objective lens should be kept clean for best performance. However, never touch the lens; it is far too delicate. A little compressed air, like that used to clean a camera lens, is recommended. If absolutely necessary, an objective lens may be cleaned with lens-cleaning paper, wetted with a 50/50 mixed solvent of IPA (isopropyl alcohol) and Freon (Freon 113 CC1 F-CC1F). In the preventive maintenance department, except when inserting or removing a CD, keep the disc drawer closed to keep out dust.

If a player won't be used for an extended length of time, unplug the player from the wall outlet. This will protect the player from lightning strikes which could easily jump the player's on/off switch and cause havoc. As noted, an AC line filter is always recommended for any audio gear. When installing an audio system, route the power supply cords so that

they are not likely to be stepped on or pinched by doors. In addition, try to distance audio cables away from power cords.

If your player has one or more transport screws, be sure to secure them prior to any transporting. Their job is to secure the sled or swing arm housing the laser pickup. Without the screws, in some models, the sled or arm is free to move about when the player is moved. Carelessness in this matter could result in a misaligned or broken laser pickup. You may well find the cost of replacing a pickup is more than the price of the player itself.

### **Repairs**

Depending on your own handyman abilities, a trip to the service shop might be required in the following situations: the power supply cord or plug has been damaged; objects have fallen into the player; liquid has been spilled into the player; the player has been exposed to rain; the player has been dropped or the cabinet damaged; the player does not appear to operate normally or exhibits a marked change in performance. Conversely, an unnecessary trip to the repair shop, wasted money, and incredible embarrassment can be saved by a little common sense troubleshooting.

Unless you are qualified, home repairs on broken CD players are not recommended. Be warned: There is very little inside that is user-friendly, nothing that is user-useful, but numerous things that are user-dangerous. In general, the pickup and electrical circuitry are quite sophisticated and will defy most casual attempts at calibration or repair. For the qualified, or the foolish, the player's service manual is prerequisite to attempting any major repair.

Care must be taken when working near a pickup's laser beam. The laser beam diverges as it leaves the pickup, so after the light passes through the focus point (a few millimeters from the objective lens), the light's intensity rapidly decreases with the inverse square law to an inconsequential intensity. However, the intensity at the focused spot may reach a value of  $1.3 \times 10^4 \text{ W/cm}^2$ . This could damage the human eye. It would be hard to get your eye close enough to the objective lens to run any risk. Even so, to be safe, do not look at the laser light through the objective lens directly, nor through a lens or mirror. Most players have a safety interlock to defeat the laser. A disc must be loaded, and the play button must be pressed before the laser turns on. Note that, unlike videodisc players, the laser beam in a CD player is invisible to the eye.

The laser diode could be damaged by a surge current or an electrostatic discharge. For safe handling, grounding of the human body and service equipment is strongly recommended.

In general, a CD player should last a long time. The electronics are long-lived, particularly because they operate at fairly low temperatures. The laser diode will eventually burn out, but life expectancy is between

ten and twenty-five thousand hours of operation. When a laser burns out, the entire pickup assembly is typically replaced. Perhaps the most troublesome aspect of player longevity is the alignment of the laser pickup. Misalignment would result in mistracking or increased error concealment. Rough handling or simple drift of the servo circuits (from age and temperature) can also cause tracking and focusing problems. That would mean a trip to the repair shop.

Aside from the normal considerations of reasonable handling and the use of transport screws when shipping, there is no preventive maintenance required. Even a yearly checkup is probably not worthwhile. CD players truly exemplify the rule that if it isn't broken, don't fix it.

## **DISC NOTES**

Despite manufacturers' assertions that compact discs sound perfect and are indestructible, it's worth looking at imperfections in discs as well as disc mortality. Specifically, the highly complex process of disc manufacturing inevitably introduces errors into discs. For a consumer, it is important to be able to identify such errors, and, if they are excessive, demand an exchange or money back from the record store. In addition, in spite of the remarkable robustness of discs, certain precautions should be observed in handling them. It's important to minimize the chances for damage after it's too late to get your money back.

### ***Compact Disc Defects***

The manufacture of CDs, as we shall see in Chapter Seven, is a highly complex operation, and there are numerous opportunities for defects to creep into the product. Theoretically, the quality control personnel at the factory reject all substandard work, but in practice some clunkers slip through. Problems can occur anywhere in the manufacturing chain. The master tape or the master disc itself could be defective. If it is, every CD made from that master would contain the same defect. For example, there might be a dropout on the tape, or a dust particle on the master disc, which could cause an error in all the production discs. For that reason, quality control is very rigorous at this stage, and errors here are unusual. It's a little more likely that defects could appear in the molds used to make the CDs. These metal replicas of the original glass master disc might have a flaw, and every disc made from that mold (perhaps 10,000 CDs per mold) would duplicate the error.

An injection molding machine forms both the disc substrate itself as well as the billions of pits on the substrate's surface. If the resulting disc is warped, optically distorted, or has impurities such as contamination

or air bubbles trapped inside, the laser might not be able to see the pits properly.

Any badly formed pits, perhaps caused by temperature inconsistencies in the mold, would translate into bad data. It is particularly important that the pit walls be clearly defined, because the transitions from pit to substrate surface (and vice versa) signal binary 1s. Dust is particularly troublesome to the naked discs that are still warm; when they leave the mold they hold a high static charge which attracts dust particles. A dust particle would disrupt large amounts of data.

Dust particles are problematic to the metallizing process. Any dust particles on the disc prevent the metal molecules from adhering to the disc substrate properly, creating a pinhole in the metal layer. The reflection of the laser beam will be disrupted by the pinhole, and the data in the pits will be lost. Pinholes occur all too often. Check your discs for yourself. Hold them up to a bright light, and you'll sometimes see several (or many) holes in the aluminum layer, letting the light shine through. On some discs, the printed label conceals the pinholes. Other metallizing problems, such as the pebbly "orange peel" surface, can reflect the laser beam unevenly. Fortunately, error correction can negate most of these disc defects.

The final production steps are not especially prone to errors. The top protection layer is applied as a fast-drying plastic liquid; as long as it spreads evenly and covers the disc's entire surface, there should be no problem. The printed label is even easier; problems typically yield only cosmetic defects. Perhaps the biggest problem at this stage is mislabeling discs—a not uncommon occurrence.

In some manufacturing methods, the center hole is molded during the injection molding process (the hole is smoothly shaped), while in other methods, the center hole is punched out after the top layer is applied (the hole is often rough). Either way (more likely in punching), the center hole could be off center or misshapen, resulting in an unplayable disc.

At the end of the manufacturing process, discs are checked by both automated and human checkers. While bad discs (perhaps 2 percent) are rejected, random errors on individual discs can pass by unnoticed. If defects are detected before metallizing, the rugged polycarbonate scrap is sometimes sold to companies which make football helmets.

### ***Evaluation Criteria***

The complexity of CD manufacturing makes it an imperfect art, and the reproduction system has been designed with this in mind. As we have seen, all discs contain error correction encoding, and all players use error correction decoding to correct many errors. Error concealment methods, by which players use cunning strategies to hide flaws, can minimize larger

problems. In short, although a disc may contain thousands of small errors, few or none of them are typically audible. An especially bad disc, however, played on a player with less than adequate error concealment ability, can result in audible errors.

It is wise to check a few quality criteria (if possible, before you buy). Although they are audibly insignificant, look for cosmetic defects such as a poorly printed label or rough edges. Any defect in the data area that is visible might cause an audible error. Check the disc for pinholes in the metal layer by holding it up to the light. A few are permissible; many are not. Use a magnifying glass to look for particles, scratches, bumps, or roughness on the reflective surface or on the substrate.

Load the disc in your player and check to see that it initializes and plays all tracks properly without skipping. If the player gets stuck, skips, or emits bursts of clicks, the disc is obviously defective. If a disc plays improperly, and no scratches are visible, the problem may be a warped disc. This can upset the laser readout considerably. To check for warpage, lay the disc on a flat surface, such as glass, and press down in the center. Turn the disc over and press again. If you detect movement of over one millimeter (about  $\frac{1}{32}$  inch), the disc is defective.

Finally, listen to the disc carefully. Errors will most likely be subtle. Preferably through headphones, listen for errors such as soft, quick clicks and pops and poorly concealed errors that occur as short bursts of distortion. How often do such audible disc errors occur? Very rarely. Depending on your perception, perhaps one out of 100 discs might require exchange.

An almost random variable in diagnosing bad discs is the player you're using. There are sometimes surprising irregularities between players' error correction performance and the type of error. A player might be a champion at handling a badly metallized disc yet mistrack on a simple scratch. The point is that pickup and error correction circuit designs vary among manufacturers and have different performance results. Until types of errors are classified and standardized, it will be difficult for consumers to consistently evaluate a player's abilities.

It is ironic that, because of the CD's fidelity, it is often very difficult to differentiate between extraneous noises present in the original recording and disc problems. For example, in orchestral recordings you'll often hear squeaking chairs, respiratory and intestinal problems, singing conductors, air conditioning rumble, randomly slamming doors, and even outside traffic noise—things obviously beyond your control.

To help advise consumers of the technical origins of a disc, many record labels have adopted a code similar to one proposed by SPARS (Society of Professional Audio Recording Services). Its three letters describe the type of tape recorder technology used at the three primary stages of a recording: session recording, mix down and editing, and mastering. For example, a CD labeled ADD would indicate that it had been originally

recorded on an analog tape recorder, mixed down to a digital recorder, and mastered to a digital recorder. Since all CDs are derived from a digital master, the code's third letter is always a D. Thus a CD labeled with an AAD code is purely an analog recording released on CD. Of course, there is no guarantee that a DDD recording will sound any better, nor does the code provide information on any analog signal processing devices or mixing consoles used, or transfers that may have taken place during an "all-digital" DDD recording. In fact, because of the code's incompleteness, SPARS has withdrawn support of the code's use.

### ***Disc Care***

Of course, aggressively abusive treatment will ruin even the most perfectly made disc. Although impervious to conditions that annihilate other audio media, a CD is best treated with the kind of respect that its high technology deserves. The following guidelines for disc handling should guarantee that your CDs will still sound just as good for your grandchildren. Of course, by that time they will be terribly obsolete, interesting only as curiosities.

1. Handle the disc by its edge or center hole. To keep the disc clean, avoid touching the data surface.
2. Do not stick paper or tape on the label surface; this probably won't hurt the disc but might jam the disc drive.
3. Because the CD's data surface is embedded directly beneath the label, protected by only a thin layer, this top surface is more vulnerable to catastrophic damage. A scratch that penetrates the label will directly destroy a considerable amount of data, whereas a scratch on the substrate's bottom can often be overlooked by the optical pickup.
4. Do not write anything on the label surface with a ballpoint pen. This certainly will hurt the disc. Furthermore, do not use a marking pen; the solvent could penetrate the top layer and deteriorate the metallization layer. However, the polycarbonate substrate of the disc is chemically impervious to almost everything.
5. Do not expose the disc to high temperature or humidity for an extended time—for example, in a car parked in direct sunlight. The considerable rise in temperature could warp the disc. Polycarbonate begins to soften around 100 degrees centigrade.
6. After subjecting the disc to extreme cold (which is not recommended), give it half an hour to return to room temperature before you play it.

7. If you must, clean a disc with a soft, moistened cloth. If possible, pat it dry. Otherwise, wipe the disc from the center out (like spokes in a wheel), and never wipe across the disc as you would an LP. This will prevent scratches from paralleling a pit track and impeding error correction. A scratch across the pit tracks is easier to correct. Modest fingerprints or smudges will not overly affect the quality of reproduction; their scrupulous removal might cause more permanent harm (in the form of scratches, for example) than good. To reiterate, preventive cleaning is unnecessary and potentially harmful. Clean only when it is required.
8. Do not use solvents such as benzine, thinner, volatile solvents, commercially available cleaners, or antistatic sprays intended for analog discs. These could irreparably damage the CD's surface.
9. After cleaning a disc, return it to its case. To prevent warpage, make sure the disc is not subjected to stress.

Although largely unnecessary in this author's opinion, a commercial cleaning solution might be useful in cases of particularly dirty discs (the surfaces, not the lyrics). A liquid cleaning fluid, a soft cleaning pad, and a soft brush should do the job. Avoid aerosol sprays; the resulting evaporation off the disc rapidly cools it, possibly leading to minute surface cracks.

There are also numerous disc cleaning machines on the market. Although most are suitable (at least for the incurably lazy), some are potentially dangerous to your discs. They wipe the disc with either a circular motion or straight across, potentially creating microscopic scratches along the pit track. Make sure any disc cleaning machine wipes the disc only radially.

If a disc substrate is badly scratched and all else fails, you might try removing the scratch: apply a little polishing compound, such as the presoftened white variety, with a soft, dampened cloth. Never use a rubbing compound. Rub the damaged area with a light motion until you restore the plastic's shine. A little water on the cloth prevents the compound from drying. Remove any residue when you have finished. If that doesn't help, nothing will. If the scratch is on the upper label side, abandon all hope immediately and panic.

#### *A Final Word about Placebos*

Considerable controversy has raged in the debate over CD treatments such as stabilizers and fluids. Some people swear by the treatments' mag-

ical sonic qualities, while others swear at them. CD tweakers advocate such treatments. They perceive the need to enhance the sonic performance of CDs through arcane techniques—for example, by applying green ink to disc edges or placing flexible plastic rings around discs. Hence the commercial availability of CD panaceas—remedies believed to correct the faults of CD and bring it up to the audiophile level.

On the other hand, more analytical listeners point to the remarkably impressive specifications delivered by even modestly priced players and argue that the CD format is in little need of tweaking. Even if there were a small improvement, it would probably be inaudible. Furthermore, analytical listeners tend to be skeptical regarding any unorthodox approach to audio. They argue that techniques such as painting disc edges green and applying wax to a CD surface are only placebos at best.

Conventional wisdom would support the analytical listeners. Methods that fall outside normal bounds of measurement are usually suspect. But when it comes to a subjective pastime such as listening to music, many other variables must be considered. Indeed, even the placebo effect may play an important role. As employed in medicine, a placebo is an inert substance often used in a controlled medical experiment; for example, half of the subjects may be given a new drug, and the other half given a placebo such as sugar pills. When a placebo is administered, a sick patient may show a response even though the placebo itself should have no medical effect on the illness. In other words, the placebo seems to affect people more through mental powers rather than physical. To assess the physical effect of a new drug, it must therefore be measured against the mental effect of taking a placebo.

Would the placebo effect explain why some listeners are persuaded that sound is improved by CD treatments even though others deny any change? Or are various audiophile treatments valid and simply unrecognized by the establishment? Of course, these very issues divide the audio world and have been debated for years. There is no question that CD player design has undergone much-needed technical improvement over the last decade, just as any technical design undergoes evolution. But what about tweak techniques such as disc fluids and rings? Are they panaceas, or placebos?

To help answer that question, this author employed a number of treatments and evaluated their effects in both the digital and analog domains. Loosely, the treatments may be divided into two groups: disc fluid treatments and disc mass treatments. Five disc fluid treatments were tested: Two were audio products while three were products not intended for audio use. They are car wax, a vinyl protectant, and a green ink pen. Although the latter three products make no sonic claims on their packaging, the first two do. For example, the packaging for one product states, “improves dynamics and transient response, improves bass articulation, im-

proves low-level detail, improves image specificity and depth, improves timbral accuracy, decreases harshness and noise."

A single disc mass treatment was tested: plastic rings that are applied around the edges of a CD disc. According to the packaging, the product "stabilizes rotating CD for maximum sonic performance. . . . Improves sound! . . . Less vibration allows the CD player's laser to read data with better accuracy, resulting in a truer reproduction of sound for clearer, more natural sounding music."

The test designed to evaluate these treatments was twofold, with measurements of both digital and analog data output from a CD player. The reference CD player employed a conservative design, with four-times oversampling and dual 16-bit D/A converters. The laser pickup was a single-beam swing arm design on a die-cast frame. The player provided both digital and analog audio outputs.

For the digital domain test, the coaxial digital output from the CD player was connected to a sound processor employing a DSP chip, which in turn employed a personal computer as its host. Audio data output from the player was captured by the computer and analyzed using the DSP system. This hook-up permits examination of audio data after it leaves the laser pickup but before it is processed by the player's output conversion circuitry. Treatment that affects the reading of data from the disc itself should be evident. This test is especially appealing because it provides a direct view of the data stream, free of any conversion and other processing that alters the signal.

To test the treatments, an untreated disc was played and a portion of the disc data was digitally captured. The same disc was then treated, and the same portion of data captured. The two files were synchronized exactly and compared on a bit-by-bit basis. If any bit in a treated file differed from its corresponding bit in the untreated file, this test would identify that bit.

The results of this test are easy to describe. In the case of all the disc fluid and mass treatments, every bit in the treated files was identical to the corresponding bit in the untreated files. In other words, as played back on this player under normal conditions, the treatments did not have any effect on the digital data output from the player.

These results strongly contradict the theories behind several of these treatments and the claims made by them. For example, the theory often ascribed to green ink is that stray light inside the polycarbonate disc substrate can diminish the pickup's ability to correctly identify pit edges and hence the data they represent. The green ink is said to absorb stray light inside the substrate, thus correcting this deficiency. In this test, the data output from the laser pickup was identical for both inked and uninked discs. Similarly, the hypothesis behind one fluid treatment is that the surface treatment provides an "optical impedance match" between disc and

air. Again, the test shows that data is unaffected. The test thus tends to disprove the validity of these treatments.

The second part of the test required evaluation of the output analog signal to determine whether disc treatments provided improvement in the analog domain. The analog outputs from the CD player were applied to an analyzer where conventional parameters such as THD + N and IM distortion could be measured in the analog domain. Measurements showed that neither disc rings nor disc fluid treatments had any effect on the quality of the output analog signal. This is logical because these treatments are applied to the disc itself, and, because they do not affect the digital data from the disc, there is no reason to believe they will affect the signal after it is converted to analog. As described previously in this chapter, extreme player vibration will stress the player's servo system and generate microphonics in the analog output sections; however, tests showed that these CD treatments did not affect this condition either.

Certainly, tests such as this are neither exhaustive nor conclusive. However, CD disc treatments do not appear to affect the readout of the digital data or subsequently converted analog waveforms. Empirical studies such as these demonstrate that CD treatments are indeed merely placebos, and the only benefit they provide is of the psychological variety.

## FOR FURTHER READING

Audio Precision. "Application Note 1: Compact Disc Player Testing with the Audio Precision System One, Revision 1." Beaverton, Oreg.: Audio Precision, December, 1987.

Dijkmans, E. C. and P. J. A. Naus. "The Next Step towards Ideal A/D and D/A Converters." In *Proceedings of the AES Seventh International Conference—Audio in Digital Times*. K. C. Pohlmann, chair. New York: Audio Engineering Society, 1990.

EIAJ (Electronic Industries Association of Japan). "Methods of Measurement for CD Players." *EIAJ Document CP-307*.

\_\_\_\_\_. "Test Disk for CD Players." *EIAJ Document CP-308*.

FCC (Federal Communications Commission). "How to Identify and Resolve Radio-TV Interference Problems." Washington, D.C.: US Government Printing Office, Stock Number 004-000-00345-4.

Feldman, L. "A New CD Test Standard." *Audio Magazine* 69(12):54–58, December, 1985.

- Finger, R. "On the Use of Computer-Generated Dithered Test Signals." *Journal of the Audio Engineering Society* 35(6):434–45, June, 1987.
- Ford, H. "CD Test Discs." *Studio Sound* 29(1):62–63, January, 1987.
- Halbert, J. M., and R. A. Belcher. "Selection of Test Signals for DSP-Based Testing of Digital Audio Systems." *Journal of the Audio Engineering Society* 34(7/8):546–55, July/August, 1986.
- Halbert, J. M., and M. A. Shill. "An 18-Bit Digital-to-Analog Converter for High-Performance Digital Audio Applications." *Journal of the Audio Engineering Society* 36(6):469–80, June, 1988.
- Lenk, J. D. *Complete Guide to Compact Disc (CD) Player Troubleshooting and Repair*. Englewood Cliffs, N.J.: Prentice-Hall, 1986.
- Sony. *Compact Disc Technology Update*. Sony White Paper, March, 1988.
- Sweeney, D. *Demystifying Compact Discs: A Guide to Digital Audio*. Blue Ridge Summit, Penn.: Tab, 1986.
- Takai, M., and S. Matsuoka. "Real Time Measurement of Jitter in CD Players." *AES Preprint No. 2414*, 81st Convention, 1986.

## *Chapter Six*

# Diverse Disc Formats

### **INTRODUCTION**

The versatility of the compact disc has quickly become apparent to manufacturers and users alike. Since music is an extremely storage-hungry phenomenon, the CD must have a tremendous storage capacity. To take further advantage of that capacity, other CD formats have been devised. CD-ROM was developed, incorporating nonaudio data, such as data base and software data. CD-I and DVI are specific applications of CD-ROM. Their data storage opportunities include audio-visual information stored in a user-interactive manner. CD-V merges the compact disc audio format with the Laservision video format. CD-R is a write-once format allowing users to store their own data. In addition, recordable/erasable CD formats have been developed as well as incompatible formats such as the Mini Disc. CD + G/M features storage for graphics and other nonmusical data. A smaller diameter compact disc, the CD-3 is used for applications requiring shorter playing time.

### **CD-ROM**

The information storage and retrieval industry is characterized by some mind-boggling statistics: the Library of Congress has 83 million items on its 535 miles of shelves and 6 million microfilms. The library houses 21 million books, 400,000 newspapers, and 10 million movies, photographs, and other artifacts. The National Archives has about 6 billion pages of images, 16 million photographs, and 5 million maps and charts in its collection. The Internal Revenue Service has about 2 billion completed tax forms on file. In 1985, 18 million American office workers each generated enough paperwork to fill 4 file cabinets.

Clearly, the need for large-capacity, fast-retrieval storage is critical. The compact disc offers several advantages over paper or magnetic storage. Foremost is its high data density. An audio CD contains a large amount of data, deceptively consolidated into a few tracks of music. The CD-ROM standard is derived from the audio CD standard but defines a format for general data storage. For example, a CD-ROM telephone book for the United States would be feasible. Assuming 100 million listed telephones, and 40 bytes of information associated with each one, the total of 4 Gbytes could be stored on about seven discs—a stack about  $\frac{1}{2}$  inch high. Moreover, any telephone number could be accessed in a few seconds. This kind of storage is the intent of the Compact Disc-Read Only Memory (CD-ROM) format.

### *Applications*

CD-ROM is the logical extension of the compact disc format toward the much broader application of information storage in general. Rather than being used to store music, the CD format can be treated as a read-only memory system used for any kind of program material. It is intended mainly as a medium for professional databases and mass storage for computer-related applications. In short, CD-ROM is a cost-effective way of distributing large amounts of information, especially information not requiring frequent updating.

The amount of data stored on a CD-ROM is, of course, equal to the amount of information stored on an audio CD, but a ROM underscores just how information-hungry a music recording is. For example, if you were to download all the information from one fully loaded CD-ROM at 300 baud (for example, transmitting it to a friend by means of a modem) for 24 hours a day, it would take 184 days to transmit the contents.

Think of CD-ROM as electronic paper (and a lot of it); anything publishable is a candidate for CD-ROM. However, a CD-ROM is much more efficient than paper. For example, the Grolier Academic Encyclopedia is available on a CD-ROM, but the text occupies only about 60 megabytes, or one-tenth of the disc storage space. A comprehensive index, occupying another 60 megabytes, greatly facilitates data access. You enter a subject and a listing of entries comes up on the screen. These entries are far more detailed than those in a conventional encyclopedia because the system's catalog can access many entries, comparing your request with the contents of the entire encyclopedia. A phrase like "freedom of speech" might have only a few citations in the index of a conventional encyclopedia, but in a CD-ROM-based system every mention of that phrase in the entire encyclopedia would be displayed. You can scroll to any entry and read the full text. The words you searched for are highlighted in the text.

In short, CD-ROM forms the basis for a new electronic publishing medium applicable to book publishing, dictionaries, technical manuals, business catalogs, computer storage, expert systems, artificial intelligence, and so on. It represents an entirely new technology of information dissemination.

### ***Data Format***

Although a CD-ROM disc is visually identical to a music CD, it employs a modified data format; a CD-ROM disc automatically identifies itself as differing from an audio CD (through the Q subcode channel). Data in a music CD is derived from the 44.1 kHz sampled signal, with 16-bit quantization. The 16 bits are divided into higher and lower 8-bit bytes. These data bytes (before EFM modulation) are grouped into frames of 24 bytes, and parity and a synchronization word are added. While satisfactory for music applications, a frame is too short for numerical applications, and there is no provision for addressing. A solution was derived from the subcode field in which subcode from 98 frames is summed. In CD-ROM, ninety-eight 24-byte frames are summed, thus the effective size for the data area becomes 2352 (24 bytes  $\times$  98 frames), a workable area. This data format is shown in figure 6.1.

This 98-frame area is called a sector, and it forms the basic CD-ROM data unit. This block is sufficiently long to handle numerical data as one sector without altering the audio CD format. Audio data is replaced by user data and system data that occupy 2048 bytes. The remaining 304 bytes are used for synchronization, headers, mode selection, and extended error detection and correction. Because there are 75 blocks output per second, a 74-minute equivalent disc could hold a maximum of 333,000 blocks, each with 2048 bytes of user data, for a total of 681,984,000 bytes. A 60-minute disc would hold 270,000 blocks, or 552,960,000 bytes.

In practice, a storage capacity of approximately 550 Mbytes is often utilized. In this way the outer 5 millimeters of the disc data area (equivalent to 14 minutes of music) are not used. This area is the hardest to manufacture and the hardest for users to keep clean. In any case, this is a large storage area, approximately equivalent to 1500 half-megabyte floppy disks, 275,000 pages of alphanumerics, 18,000 pieces of computer graphics, or 3,600 still video pictures.

With CD-ROM, PCM data is recorded with EFM, as in audio CDs. The first 12 bytes from the 2352-byte block are used as a synchronization word. The next 4 bytes form a header field used for time and address flags. Specifically, the header contains three address bytes and a mode byte. The address bytes store location as time: One address byte holds minutes (74 minutes maximum), the second byte holds seconds (0 to 59), and the third holds the block number within the second (75 blocks per second).

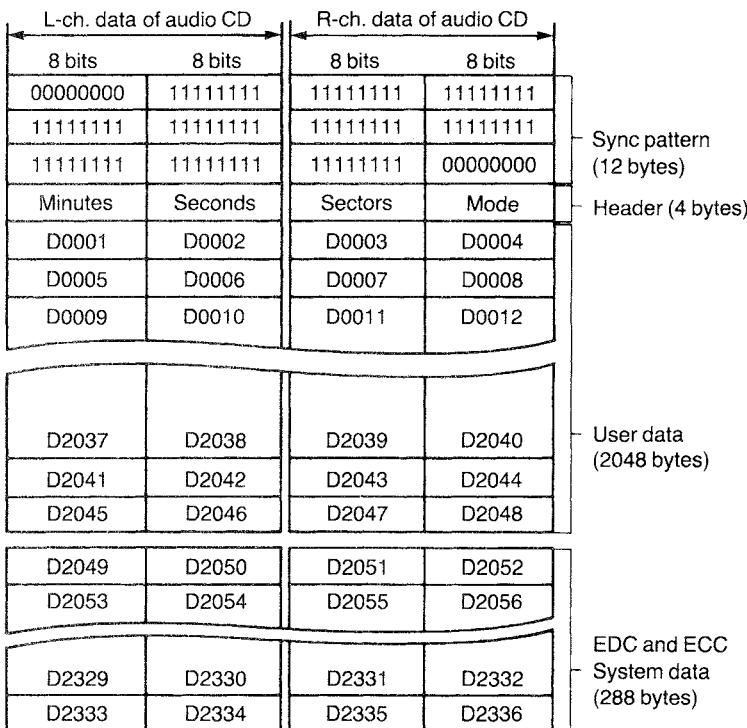


Figure 6.1     *CD-ROM data sectors are formed from ninety-eight consecutive CD frames. (Courtesy Nippon Columbia)*

For example, an address of 59-20-45 denotes the 45th block in the 20th second of the 59th minute on the disc. Although this repeats information already found in the Q subcode channel, it speeds and provides greater accuracy for searching.

The mode byte indicates the nature of the user data; for example, computer data is distinguished from audio-visual data. Figure 6.2 shows the three modes that are used. The Mode 0 format is used for null data. The Mode 1 format specifies that 2048 bytes of each block are devoted to user data. The Mode 2 format allows for 2336 bytes of user data. The difference between Mode 1 and 2 is that Mode 1 uses 288 bytes to provide an additional layer of error detection and correction (EDC/ECC). This is in addition to the basic error correction used on all compact discs, ensuring a level of data integrity essential for information that does not degrade gracefully, such as text and database binary code. Mode 2 trades this benefit of data security for maximum data rate by making the additional 288 bytes per sector available for user data. The basic error correction encoding is suitable for gracefully degrading data such as video and audio.

Mode 0

Sync (12)	Header (4)			All zero (null data) (2336)			
	Block address (3)	Mode	Min. (1)	Sec. (1)	Block (1)	0 (1)	

Mode 1 (153.6 kbytes/sec.)

Sync (12)	Header (4)			User data (2048)	Auxiliary data (288)			
	Block address (3)	Mode	Min. (1)	Sec. (1)	Block (1)	EDC (4)	Space (8)	ECC (276)
						P-parity (172)	Q-parity (104)	

Mode 2 (175.2 kbytes/sec.)

Sync (12)	Header (4)			All user data (2336)			
	Block address (3)	Mode	Min. (1)	Sec. (1)	Block (1)	2 (1)	
						Scrambled area	
						2,352 bytes/1 CD block	

Figure 6.2 Three modes are defined in the CD-ROM format. Mode 1 provides for extended error detection and correction. Mode 2 permits more user data. Mode 0 is a null mode. (Courtesy Nippon Columbia)

Because of extended error correction, Mode 1 will have the most numerous applications; each block contains exactly 2 kbytes ( $2 \times 1024$ ) of user data. The 2048-byte area is where the actual CD-ROM data resides.

### Error Correction

As noted, the remaining 288 bytes, called system data, provide for a higher degree of error detection and correction (EDC/ECC). This is a prerequisite to successful numerical data storage, which is more demanding than audio data. This system data forms an extended error correction code (EDC and layered ECC) existing independently and in addition to the cross interleave Reed-Solomon error correction code already encoded for each CD frame. The error rate is improved over that of an audio CD. Theoretically, CD-ROM bit error rates will be between  $10^{-16}$  and  $10^{-17}$ , or one uncorrectable bit in every  $10^{16}$  or  $10^{17}$  bits. The extended code may be processed either by hardware circuitry contained in the CD-ROM player or by a host computer.

The CD-ROM standard recommends that burst errors in the RF signal due to local defects should include almost no C2 uncorrectable errors. Two methods may be used to define the required disc quality: (1) With a

simple C1 and C2 error correction decoder, no more than one symbol error may occur in a data block measured at the input of the C2 decoder; (2) the number of successive C1-uncorrectable blocks should be less than seven. A block is called C1-uncorrectable if more than one symbol is erroneous. A symbol is called erroneous if one or more bits are erroneous.

CD-ROM production begins with a master data tape or tapes, recorded in computer magnetic tape format. The data is entered into a pre-mastering minicomputer where the data is placed into the CD-ROM format (header, error correction, subcode, and so forth are generated). Following pre-mastering, actual disc production is identical to that of audio CDs, albeit with greater quality control to ensure accuracy of data.

### ***CD-ROM Players***

While CD-ROM uses a data format similar to that of music CDs, the players are not compatible. A CD-ROM player contains laser optics, modulation, and error correction, but D/A conversion and audio output sections are replaced (or supplemented) with a computer interface to output the ROM data to a host computer, such as the one shown in figure 6.3. Data is transferred to the host computer in blocks of 2 kbytes. Using the header field, the host computer verifies the sector address in minutes, seconds, and blocks. The disc table of contents (TOC) can be transferred to the CD-ROM player and used to locate a requested track.

The average data transfer rate is 75 sectors per second, or 153.6 kbytes per second (Mode 1) and 175.2 kbytes per second (Mode 2) for a constant linear velocity (CLV) of 1.2 meters per second. Access time for any point on a disc is less than a second.

Because they are not tied to one specific operating system or data processor, CD-ROM devices can be interfaced with all existing and future computer systems. CD-ROM is limited only by the capabilities of the operating system and microprocessor in the host computer.

A good example of the portability of the CD-ROM format, in a literal sense, is the Data Discman product introduced by Sony Corporation. This portable player incorporates a LCD screen, QWERTY and control soft keys, disc drive, and firmware search engine in a hand held package. It accepts 3-inch CD-ROM discs enclosed in a protective caddy, each holding the equivalent of 100,000 pages of printed text, 32,000 graphic images, or a combination of both. Six standardized search and retrieval methods let the user locate text or graphical information by entering the first or last few letters of a search word, or by entering several keywords. You can also select topics from text or graphic menus. A cross-reference "hot-link" system lets the user jump to and from screens of additional text or graphic information. The Data Discman may be powered from batteries or AC line voltage, and video and data outputs permit connection to large

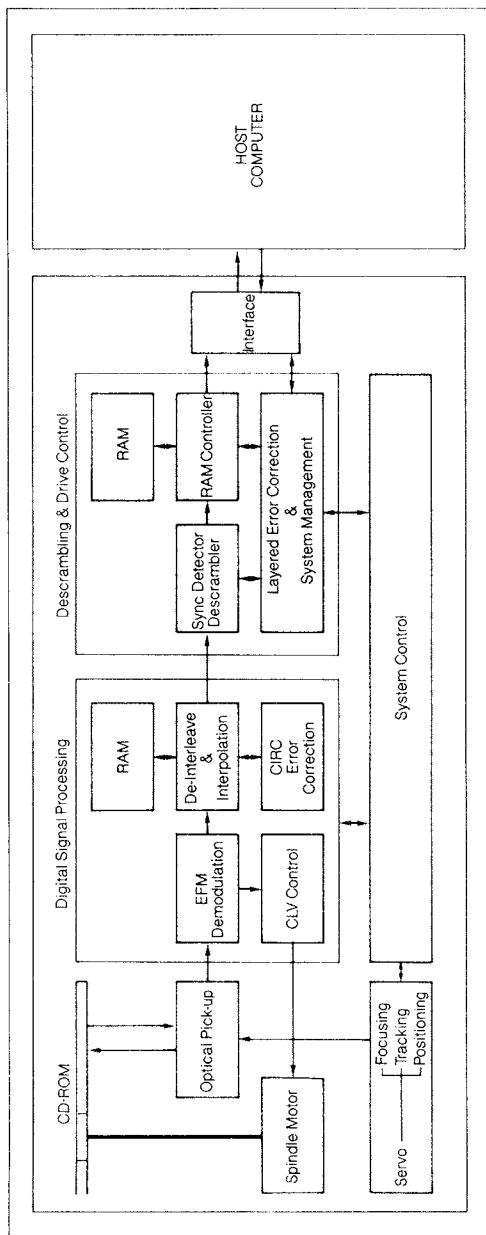


Figure 6.3 *Block diagram of a CD-ROM drive and interface to host computer.*  
(Courtesy Sony)

screen televisions and personal computers. Examples of disc titles include bibles, telephone books, airline schedules, language translators, dictionaries, restaurant guides, novels, and games. This electronic book is a good example of the power embodied in the CD-ROM format, and truly lets you slip an entire encyclopedia into your shirt pocket. However even this forward-looking consumer product remembers its origins—the Data Discman can also play CD-3 digital audio discs.

### ***The CD-ROM Standard***

It is important to note that the CD-ROM standard as specified in the *Yellow Book* available to licensees, unlike the audio CD standard, does *not* link CD-ROM to any specific application. The format is thus transparent. It does *not* define the type of information that is stored in the user areas of Mode 1 and Mode 2. Hence the user data may be text, video, software, graphics, audio, and so forth. Nor is the encoding principle for various types of information defined. The means for encoding or compressing text, graphics, audio, or video are not specified. Furthermore, the layout of information on the disc is not defined. The standard does not indicate where and how to store the directory, how to identify the beginning or end of a file, or how to open a file. The CD-ROM standard has been enhanced to include the CD-ROM XA (extended architecture) format. It incorporates some of the audio and video technology of the CD-I format to form a bridge between CD-ROM and CD-I. CD-ROM XA Audio contains CD-I's B- and C-level audio standards, as well as the CD-DA format. CD-ROM XA Screen Presentation contains CD-I's 640 × 480 pixel resolution CLUT, single-plane graphics, and ISO 8859/1 character set as well as run-length encoding and four VGA video modes. A maximum of 256 colors are supported. As with any CD-ROM disc, CD-ROM XA is not dependent on a specific operating system.

Because the CD-ROM standard does not provide a logical file structure for CD-ROM media, software developers have implemented their own file formats as well as extensions to common operating system environments such as MS-DOS, VMS, etc. One group of vendors, the High Sierra Group, developed a proposal for a standard CD-ROM logical file structure. This has been codified by the International Standards Organization (ISO) as standard ISO/DIS 9660, "Volume and File Structure of CD-ROM for Information Interchange." All information on CD-ROMs that follow this standard can be made readable as standard files in all operating system environments. In addition, the logical file structure proposed by the High Sierra Group is forward compatible with CD-I equipment, so that properly coded CD-ROM discs operate on CD-I systems. On the other hand, a CD-ROM system is not capable of processing information on a CD-I disc, unless the CD-ROM system uses the processor,

operating system, and audio and video encoding as specified in the CD-I format.

## CD-I

Compact Disc-Interactive (CD-I) is a special application of the CD-ROM format. Rather than store specific data on a CD-ROM, or music on an audio CD, CD-I permits storage of a simultaneous combination of audio, video, graphics, text, and data, all functioning in a defined, interactive format. CD-I is thus a multimedia extension of digital audio found on CD-DA (audio) discs. Because CD-I players reproduce conventional CD-DA discs, CD-I is also an upscale CD-DA system.

The CD-I format defines both hardware and software standards, much like the audio CD format. Although CD-ROM can also store text, graphics, and so forth, CD-I defines a special integration of such functions. Because CD-I is an interactive medium, its information may be accessed through a dialogue procedure. The system can present the user with alternatives that can be used to steer the user to the desired information.

The specification for the CD-I standard is detailed in the *Green Book* available to all CD-I licensees. CD-I is a specific application of CD-ROM, with rigidly defined implementations as in the audio CD format. The CD-I standard defines the following:

1. How various types of information, such as video, audio, text, executable code, and graphics, will be distinguished and identified on the media.
2. How each type of information will be encoded, including specification for several formats of high-, medium-, or low-resolution video, audio, graphics, data reduction, and so forth.
3. How the logical layout of files will be handled on the disc.
4. How hardware will read discs and decode the contained information.

CD-I is specified in detail because it is targeted for the mass consumer market as a user-friendly, standardized system. It is thus important for all present and future CD-I discs to be interchangeable and to play in all present and future CD-I players.

### *Applications*

CD-I systems are single-medium systems, containing both the operational program and the information itself, so that no secondary media such as floppy disks are required for boot-up. In addition, CD-I is intended to ride

piggyback on other electronics products, such as stereo playback systems and television, for multimedia presentation. CD-I uses include automotive applications such as maps, navigation, real-time animation, and diagnostics. Education and training applications include do-it-yourself formats, home learning, interactive training, reference books, albums, and talking books. Entertainment applications include music plus text, notes and pictures, action games, strategy games, adventure games, and activity simulation. Creative leisure applications include drawing and painting, filming, and composing. Work applications include document processing and information retrieval and analysis.

A single CD-I disc might contain over 17 hours of audio, 7,000 color pictures, 150,000 typed pages, or any combination of the three. Thus CD-I applications are diverse. For example, a CD-I dictionary might contain a word and its definition, as well as spoken pronunciation, pictures, and cross referencing to synonyms, antonyms, word relationships, origins, and translations into foreign languages.

A single CD-I disc could contain a biography of a composer, providing textual information, pictures, recorded examples of the composer's music played while the score is displayed, and a complete catalog of musical works as well as available recordings. Another application is the "teach-yourself" or "how-to-do-it" field; the CD-I's ability to convey text, pictures, and diagrams combined with sound make it ideally suited. For example, the sound of a motorcycle engine in various stages of tuning could be reproduced on the disc. A hobbyist's book on ornithology could reproduce bird calls. Tourists could obtain a multimedia preview of popular vacation spots. Music lovers could define the parameters of reproduced music using elastic music programs. Motion pictures could be provided with a labyrinth of plot deviations, differing each time the movie is viewed, steered at the discretion of the reader.

### ***Data Format***

These applications require flexibility in audio and video quality levels. Thus audio fidelity, or video resolution and color coding, can be selected. Longer playing times are achieved by dividing a disc into channels, the number of which would depend on the quality of the audio and video programs. In addition, real-time requirements for interactivity necessitated the adoption of executable object code as the coding method for application software. Multimedia applications also call for physical interleaving of the three basic types of data (audio, video, and text/binary) to ensure synchronized presentation of different data types. Finally, the disc must provide efficient storage for both data that degrades gracefully (audio and video) and data that does not (text/binary). The former thus re-

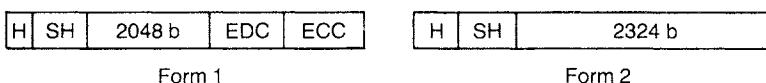
quires maximum bandwidth, while the latter requires extended error detection and correction.

The CD-I data format may be considered a subset of the CD-ROM data format. A CD-I system is designed to read CD-I discs, compatible CD-ROM discs, and CD-DA (audio) discs. The CD-I format provides for this three-way compatibility. It uses a subheader and two physical formats (referred to as forms), as shown in figure 6.4. The subheader (SH in the figure) is used for real-time physical interleaving of data, while the two forms define two levels of data integrity.

Form 1 is intended for text, computer data, and highly compressed visual data; EDC/ECC (extended error detection and correction) is used. The extended error code (CIRC encoding) used is the same as for CD-ROM Mode 1. In Form 1, user data occupies 2048 bytes, and 280 bytes are reserved for extended error detection and correction codes. Form 2 is intended for real-time audio and video; EDC/ECC is omitted. In Form 2, user data occupies 2324 bytes. Data is recorded to disc with EFM modulation.

To differentiate one kind of CD-I data from the other, a subheader is placed immediately after the regular 16-byte CD-ROM sync/address/mode header. Depending on the kind of data in that block, the data is directed to the appropriate circuitry for reproduction, display, or processing. In Form 1, the 8-byte space usually found between the header and the data area is omitted to compensate for the 8 bytes used by the subheader. This yields 2048 bytes. Otherwise, in Form 2, 2324 user bytes are available, which is eight less than in CD-ROM Mode 2.

The compatibility hierarchy of the CD family is summarized in figure 6.5. When only a physical format compatibility is required, then specific CD-ROM discs may be associated with specific systems, as shown at the top of the figure. At a second level, with professional text-oriented contents, a bridge exists between CD-ROM and CD-I, based on the Mode 1 format. For the consumer marketplace, a further level of compatibility is required so that all CD-I discs are playable on all CD-I systems. However, some discs may contain data not recognizable by a lower level (base case) player. In addition, audio CD discs may be played on CD-I systems. To summarize, CD-I discs are based on the CD-ROM Mode 2 format.



Form 1

Form 2

Figure 6.4     *The CD-I format uses two forms—for extended error detection and correction or for greater storage capacity.*

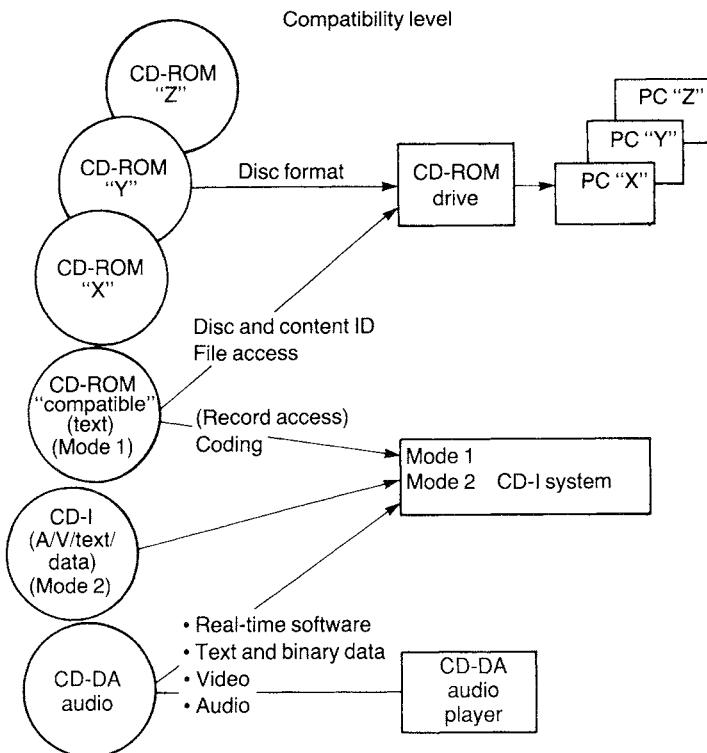


Figure 6.5 A compatibility hierarchy exists between CD-DA, CD-ROM, and CD-I. (Courtesy Philips)

CD-ROM/CD-I compatible discs are based on the CD-ROM Mode 1 format as specified by the High Sierra Group. The CD-I system is capable of reading audio CD discs.

The CD-I format calls for a total storage capacity of approximately 650 Mbytes. Because a CD-I disc is recorded with constant linear velocity, a constant readout rate of 75 sectors (or frames) per second is achieved. This results in a data transfer rate of 150 kbytes per second for Form 1 (and Mode 1) and 170.2 kbytes per second for Form 2. In both cases, this represents user data, excluding headers, and so forth.

### ***CD-I Audio***

The CD-I format offers five levels of audio quality, to be selected according to the need for fidelity. These levels are shown in table 6.1. To allow for extended video information, data reduction techniques may be

Level	Encoding	Sampling Rate	Word Length	Signal-to-Noise Ratio
CD-DA A	PCM	44.1 kHz	16	98 dB
	ADPCM	37.8 kHz	8	90 dB
B	ADPCM	37.8 kHz	4	60 dB
C Information	ADPCM phonetic	18.9 kHz —	4 —	60 dB —

Level	Band-width	Number of Channels	Data for 1 Second of Sound	Typical Playing Time
CD-DA A	20 kHz	1 stereo	171.1 kbytes	1 hour
	17 kHz	2 stereo	85.1 kbytes	2 hours
		4 mono	42.5 kbytes	4 hours
B	17 kHz	4 stereo	42.5 kbytes	4 hours
		8 mono	21.3 kbytes	8 hours
	8.5 kHz	8 stereo	21.3 kbytes	8 hours
C Information		16 mono	10.6 kbytes	16 hours
	—	—	—	10,000 hrs

Table 6.1     *The CD-I audio format provides for five audio quality levels.*

used to decrease the storage space required to encode the audio program. A CD-DA mode uses linear 16-bit PCM with 44.1 kHz sampling rate, yielding an audio bandwidth of 20 kHz. The quality duplicates that of stereo audio CDs. Alternatively, 8- or 4-bit adaptive delta pulse code modulation (ADPCM) is used for three levels of varying fidelity. The "A" Hi-Fi audio level uses 8-bit ADPCM with a sampling rate of 37.8 kHz, yielding an audio bandwidth of 18 kHz. The quality is similar to that of vinyl LP records. The "B" Mid-Fi audio level uses 4-bit ADPCM with a 37.8 kHz sampling rate, yielding an audio bandwidth of 18 kHz. The quality is similar to that of an FM broadcast. The "C" Speech audio level uses 4-bit ADPCM with an 18.9 kHz sampling rate, yielding an audio bandwidth of 9 kHz. The quality is similar to that of an AM broadcast. The CD-I format also supports an "information audio level" in which phonetic information is encoded. Its quality is similar to that of a telephone. Because the data is not fully decoded by hardware, but uses both the microprocessor and the ADPCM audio decoder, there is a slight delay, during which the phonetic information is synthesized for output.

An audio program may be divided into multiple channels (each with a maximum of 74 minutes of playing time) stored in parallel across the disc playing surface. The 16-bit PCM channel is stereo; the three ADPCM modes can be recorded in either stereo or monaural. The number of channels increases as the fidelity level decreases. Of course, there are always twice as many monaural channels as stereo channels. Based on a CD-DA stereo program length of 1 hour, the "A" Hi-Fi level would permit two stereo or four monaural channels of 4 hours; the "B" Mid-Fi level would permit four stereo or eight monaural channels of 8 hours, and the "C" Speech level permits eight stereo or sixteen monaural channels of 16 hours. The Information audio level permits storage of 10,000 hours of speech database. The parallel audio channels may store simultaneous data (e.g., a story in different languages) or sequentially played data. With the latter, there is a 1- to 3-second pause when switching from the end of one channel to the beginning of the next channel.

The audio information on a disc might share its storage capacity with video or other information. Given the different quality levels, 1 hour of audio leaves greater or lesser storage space for video information. Figure 6.6 illustrates the ratios. Similarly, the amount of video information depends on its quality level. The CD-I standard also incorporates the MPEG (Motion Pictures Expert Group) extension for digital audio (and full-motion digital video). The MPEG extension is discussed in more detail later in this section.

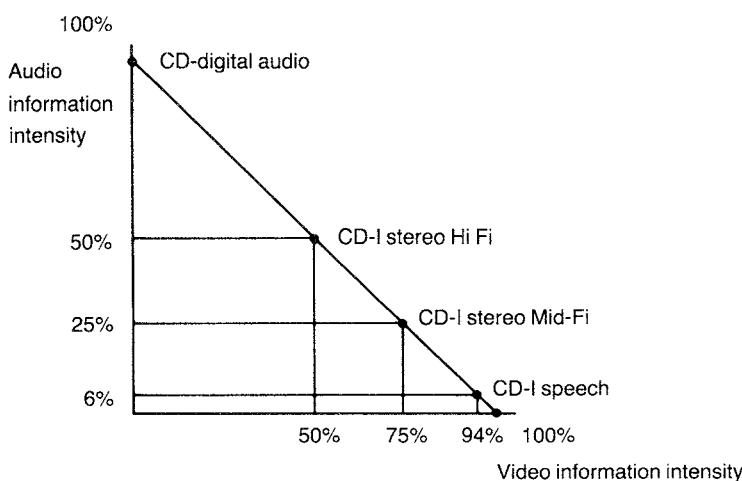


Figure 6.6 *Disc storage space dictates a trade-off between video and audio information capacity. (Courtesy Philips)*

### The CD-I ADPCM Audio Format

ADPCM combines elements of pulse code modulation (PCM) encoding and delta modulation (DM) encoding. Because of its ability to store digital data with fewer bits, ADPCM is extremely efficient. On the other hand, ADPCM requires additional processing beyond regular PCM for both encoding and decoding. With PCM, the audio signal is sampled, and the amplitude is quantized and coded into binary pulses; the code permits the determination of the quantized value. With delta modulation the audio signal is sampled, but the size of each quantized staircase step is kept constant. Furthermore, the staircase can move by only one quantization level at a time. Thus the quantized signal must either increase by one step size or decrease by one step size at each sample time. Delta modulation thus encodes the difference from one sample to the next, as shown in figure 6.7.

Unfortunately, delta modulation faces several problems. The accuracy of the encoding is directly dependent on the size of the positive- and negative-going steps—the smaller the steps, the better the approximation to the audio signal. However, when the signal is changing rapidly, small steps cannot keep up, as shown in figure 6.8. The problem can be solved by increasing the sampling rate, but this detracts from the efficiency of the scheme.

One variation on delta modulation is adaptive delta modulation (ADM). With ADM the quantization step size is varied according to the behavior of the audio signal, illustrated in figure 6.9. Large steps are used for fast-changing waveforms so slew-rate distortion is reduced. Similarly, small steps are used to maintain resolution for slowly changing wave-

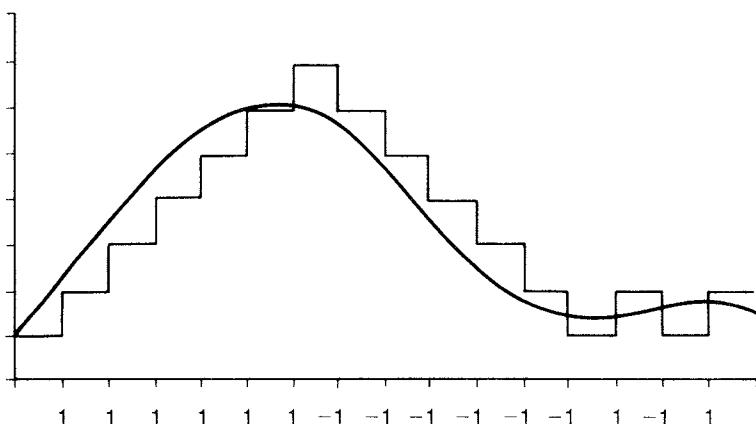


Figure 6.7     *Delta modulation encodes the difference from one sample to the next.*

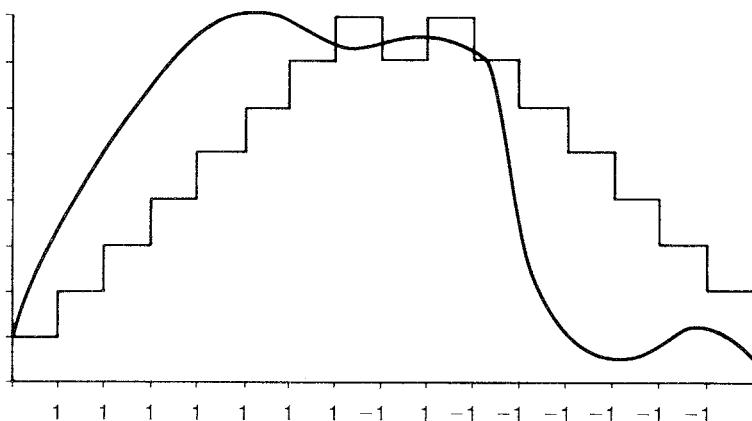


Figure 6.8     *A delta modulation system may not be able to respond to a rapidly changing audio amplitude. The result is slope overload distortion.*

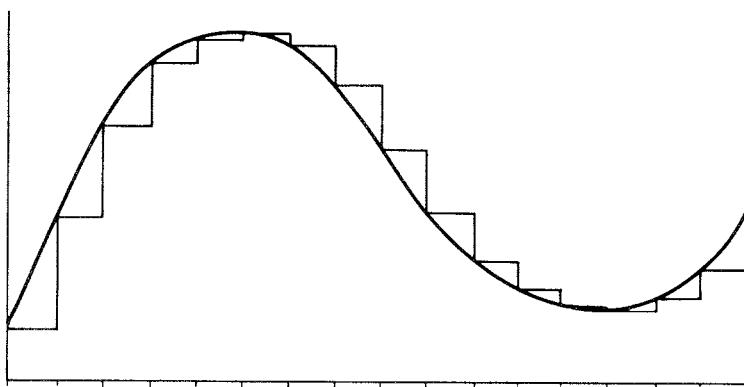


Figure 6.9     *With adaptive delta modulation, step size can be varied.*

forms. Various algorithms may be used to select the proper step size. For example, alternating data polarity would indicate good tracking, and the increment size would be reduced for even greater accuracy. When continuous polarity is present, the waveform's amplitude is clearly changing quickly, so larger steps are selected. However, certain conditions, such as simultaneous high frequencies and high amplitudes, demand large steps which increase quantization noise. This can produce modulation noise in which the noise floor becomes a function of the signal itself.

Adaptive delta pulse code modulation (ADPCM) combines the difference signal of DM with the binary code of PCM; instead of a 1-bit

correction signal, a greater number of bits are used. A PCM word is recorded or transmitted and used to reconstruct the audio staircase. The ADPCM used in the CD-I format uses 4- or 8-bit words, depending on the level of sound quality required. With a 4-bit system, correction information is available at one of sixteen levels, which provides better resolution. Since each of the quantization levels is assigned a step-size scale factor, step sizes may be adapted with greater accuracy, because more step-size information is available. The scale factors are based on the statistics of the signal itself; for example, scale factors for an ADPCM circuit designed to process speech would be selected differently from those for a system designed to process music.

The ADPCM system employed in the CD-I format may be considered a bit rate reduction technique. This technique is based on the fact that the amplitude and spectrum distribution of audio signals tend to be concentrated in a specific region. The scale factors and other coding elements in the CD-I ADPCM format take advantage of these statistical properties of audio. During encoding, the original audio data has a sampling rate of 44.1 kHz. So a sampling rate converter is used to reduce the sampling rate to 37.8 kHz or 18.9 kHz, depending on the CD-I audio quality level to be used. Word length is then reduced from the original 16 bits to either 8 or 4 bits per sample. A bit rate reduction (BRR) encoder accomplishes this. According to fluctuations in the frequency distribution of the audio signal, it selects different delta PCM modes. It also employs companding to increase signal dynamic range. Range and filter information, along with other residual data, are added at the block encoder. A subheader is added. Audio data is multiplexed with video or computer data, a sync word and header are added, and error detection/correction and EFM encoding are performed during disc mastering.

Decoding is accomplished in the CD-I player. EFM demodulation and error detection/correction decoding are effected, and headers are removed. Interpolation is performed if error flags are present, and ADPCM audio data is block-decoded and expanded to linear 16-bit form by the BRR decoder. D/A conversion, low-pass filtering, and de-emphasis complete the audio processing.

At a similar sampling rate ADPCM performs competitively with regular PCM. If the signal were always near its maximum frequency, ADPCM's performance would be similar to PCM's. With audio signals, however, this is almost never the case. Since the instantaneous audio frequency is relatively low, the signal is changing more slowly. Amplitude changes are therefore smaller, so ADPCM's quantization error is less than that for PCM. Thus in theory, ADPCM can achieve better signal resolution for the same number of quantization bits.

As implemented in the CD-I format, ADPCM's fidelity is lower than the PCM signal on an audio CD, but it is suitable for many CD-I appli-

cations, particularly when speech composes the audio program. ADPCM is well suited for the CD-I format because relatively fewer bits are needed to achieve good performance.

### ***CD-I Video***

A CD-I disc can store video material compatible with both the NTSC (America and Japan) and PAL/SECAM (Europe) television standards. CD-I discs may be manufactured to provide either 525 scanning lines (NTSC) or 625 scanning lines (PAL). Because CD-I visual information is coded digitally and because the digital coding is the same for all television standards, either type disc is playable on any standard television. When a standard disc is played on a corresponding standard television (525 through 525, or 625 through 625), negligible visual distortion is produced. When a disc is played through a television adhering to another standard (525 through 625, or 625 through 525), approximately 7 percent visual distortion results. Generally, the eye is able to perceive 5 percent distortion in a side-by-side comparison or 10 percent distortion otherwise. Alternatively, a CD-I disc may be produced with a 525/625 "compatible" video format, playable on either a 525 or 625 television with 3.6 percent distortion.

The CD-I format calls for varying quality levels for resolution and pixel coding. Three standards of video resolution are supported: Normal Resolution of  $360 \times 240$  pixels (NTSC) and  $384 \times 280$  pixels (PAL); Double Resolution (twice the horizontal resolution) with  $720 \times 240$  pixels (NTSC) and  $768 \times 280$  pixels (PAL); and High Resolution (twice the horizontal and vertical resolution) with  $720 \times 480$  pixels (NTSC) and  $768 \times 560$  pixels (PAL). Normal Resolution corresponds to the best resolution achievable with present television receivers, while Double and High Resolution represent the quality of future enhanced or digital receivers. Pictures are generally noninterlaced, although interlacing can be used.

Image coding provides for three picture qualities: studio picture quality and two graphics qualities. Natural (photographic) pictures use DYUV (Delta YUV) coding in which Y represents luminance and U and V are chroma. Using statistical properties of pictures, a data reduction of 3 over RGB (red-green-blue) coding is achieved. Pictures occupy about 325 kbytes per picture without interlacing (650 kbytes with interlacing). All natural pictures are reduced with DYUV coding to about 85 kbytes (NTSC) and 105 kbytes (PAL). At a data rate of 170.2 kbytes per second in Form 2, one full-frame natural picture is transferred in about 0.5 seconds.

The first graphics mode is designed for end-user manipulation applications; it is based on RGB 5:5:5 coding and supports 16-bit/32,768 colors, with 5 bits for each of the R, G, and B components and 1 bit for transparency. A 16-bit RGB graphic would occupy about 170 kbytes per NTSC picture and 210 kbytes for PAL.

The second graphics mode is based on CLUT (color lookup table) graphics which permit 4-bit/16, 7-bit/128, or 8-bit/256 color full screen animation. Up to eight colors may be redefined per scan line. Because 256 levels may be defined for each R, G, and B element, the total number of colors available for a single image is  $256 \times 256 \times 256$ , or about 16 million. The CLUT 8-bit mode requires 108 kbytes per picture. Taking advantage of the statistical properties of pictures, a data reduction of 2 over RGB is achieved. With reduction, the interleaving of sound and CLUT picture results in the ability to provide full-screen animation in Form 2. Both 7-and 8-bit CLUT pictures (like DYUV) occupy about 85 kbytes in NTSC and 105 kbytes in PAL.

Run-length CLUT coding may be used for cartoon-like animation or average resolution pictures. Taking advantage of the statistical properties of pictures, it achieves a data reduction of between 8 to 12 over RGB. Either 3-bit/8 or 7-bit/128 color coding may be used. A drawing requires only 10 to 20 kbytes of data.

CD-I developers have also proposed a high resolution extension to the CD-I video standard which yields high picture quality without the need for large data storage or long loading time. The system is called QHY (Quantized High-resolution Y). It is a new method of coding the digital data on the disc to produce more distinct images on the screen. The edge of an image defines its cleanliness. QHY uses pixel manipulation on image edges to remove redundant data from a high resolution picture with only a slight increase in storage data and loading time over normal resolution.

### ***Full-Motion MPEG Video***

The CD-I format also includes an extension to the base case which permits full-motion video (FMV) to be displayed on a television screen. This extension is also referred to as FSFM (full-screen, full-motion) video. Using this extension, a CD-I disc can hold over 60 minutes of digital video with a digital audio stereo soundtrack. To achieve this, the CD-I format adopted the MPEG (Moving Picture Experts Group) international standard for coded representation of moving pictures, associated audio, and their combination when used for storage and retrieval on digital storage media. MPEG is a working committee sanctioned by the ISO and IEC standards organizations. Interestingly, CD-I was the first large volume application of MPEG anywhere, beating the personal computer introduction of MPEG.

The MPEG audio and video coding algorithms allow use of video sequences coded with a variety of CD-I picture formats, as well as a variety of CD-I audio formats. Although FMV images normally fill the screen, they can also be presented in smaller rectangular windows of various sizes, shapes, and positions. The images appear in a video plane that can be mixed with the two standard CD-I image planes, allowing FMV data to

be easily incorporated into CD-I applications. For the screen display, the FMV plane and the normal CD-I planes are selected by the application as required. Switching within pictures is possible down to the pixel level.

The FMV format supports video source material with 25 or 30 frames/second, as well as a 24 frames/second film rate. Linearly encoded real-time video requires a high data delivery rate. For example, the data delivery rate for standard quality pictures as specified by the CCIR 601 standard is 165 Mbits/second. The achievement of FMV lies in reducing such data rates to within the 1.41 Mbits/second data delivery capacity of CD-I. Using the reduction supplied by the MPEG video coding algorithms, pictures coded at the CCIR 601 rate of 165 Mbits/second can be reduced to 1.2 Mbits/second, a reduction ratio of 140:1. Even higher data reduction ratios are available for the lower picture qualities that are acceptable in many applications.

Similarly, the MPEG audio algorithms can reduce the CD-DA data rate of 1.41 Mbits/second to approximately 0.2 Mbits/second, a ratio of 7:1. Using the reduction algorithms in combination, full-motion video with stereo audio can be output from the CD-I disc. The audio coding algorithms permit a variety of coding rates and sound quality levels. Depending on applications, three layers of coding with increasing complexity can be employed. In all three layers, the time domain input audio signal is converted into a frequency domain representation. In Layers I and II a filter bank divides the signal into 32 subbands which are quantized and adaptively blockwise coded with the help of a psychoacoustic model. Layer II achieves greater reduction through more precise quantization and removal of redundancy in the scale factors. Layer III achieves still greater reduction through a hybrid filter bank in which each of the 32 subbands are further split by a linear transform that operates on 18 subband samples in each subband. These higher resolution frequency lines are nonuniformly quantized and are coded with adaptive segmentation and entropy coding as well as a psychoacoustic model for greater coding efficiency.

Stereo bit rates range from 64 kbytes/second to 448 kbytes/second, and the standard also supports mono audio coding of 32 kbytes/second. In addition, in the stereo modes, stereophonic irrelevance and redundancy can be optionally exploited to further reduce the bit rate. Audio bit rates below 256 kbytes/second are useful for applications requiring more than two audio channels while maintaining full screen motion video. Rates above 256 kbytes/second are useful for applications requiring higher audio quality, with partial screen video images. In either case, the bit allocation is dynamically adaptable according to need. Designers claim that the 256 kbytes/second rate yields an audio quality that is subjectively indistinguishable from CD-DA playback.

The MPEG video algorithm uses a modified DCT (discrete cosine transform) technique to compare sequential frames and detect spatial differences between moving images. Because static image information need not

be redundantly stored, the algorithm can achieve the substantial reduction in data needed to represent a sequence of frames.

Information from the DCT can be coded with Huffman code processing to further reduce the bit rate. In addition, the algorithm exploits the visual characteristics and limitations of human vision to place coding impairments in frequencies and regions where they are perceptually minimal, thus relatively enhancing the perceived picture quality. A rectangular spatial format is used, with a maximum picture area of 352 pixels by 288 lines, but the picture area can be used flexibly. For example, a low and wide area of 768 pixels by 132 lines, or a high and narrow area of 176 pixels by 576 lines, could be coded. During playback, the entire decoded picture or other defined parts of it may be displayed, with the window's size and shape under program control.

The video quality of the MPEG standard used in CD-I is similar to the quality derived from the VHS format. In addition to the MPEG1 standard used in CD-I, the ISO/IEC MPEG Group has called for two other standards, MPEG2 for digital television and computer-quality images and MPEG3 for HDTV-quality images. These standards may require data rates of 5 to 10 Mbits/second and 60 Mbits/second respectively. Data reduction is described in the context of the Photo CD and mini disc formats later in this chapter.

During playback, a demultiplexer in the player distributes the data from the CD-I audio and video sections, while FMV data is directed to the separate FMV audio and video decoders. The FMV decoder distinguishes five play modes: play forward, freeze, single step forward, slow motion forward, and scan. During the scan mode, a sequence may be scanned in forward or reverse directions; each scanned picture is displayed briefly as a still picture. The audio decoder distinguishes two modes: play forward and mute.

In a base case model, FMV is not incorporated in a CD-I player. In other models, FMV may be added by plugging an optional upgrade cartridge into the player. In other models, FMV is an integral part of the CD-I player. As with all CD-I applications, discs using FMV are compatible regardless of their country of origin; the MPEG decoder chip permits full-motion video to be shown at either 30 (NTSC) or 25 (PAL/SECAM) frames/second. All discs using FMV are playable in the NTSC (525 lines) and PAL/SECAM (625 lines) environments. With the addition of full-screen, full-motion video technology, CD-I becomes a complete interactive multimedia system for home entertainment as well as educational and industrial applications, with worldwide universality.

### ***CD-I Text***

Text encoding can be performed either by a bit-map process or by character encoding via systems software. The bit-map process requires 5 bytes

per character, resulting in about 120 million characters per disc (if only 16 colors are used in an  $8 \times 10$  matrix). These characters cannot be manipulated under program control. Up to four overlaying planes are defined, with both transparency and translucency for all except the background plane. One plane is reserved for external video and another for the cursor.

Character-encoded text can either be system text or application text in standardized form, using 1 byte per character, yielding over 600 million characters on a disc. The ISO 8859/1 standard character set is available in the CD-I format. Application text can be encoded with 2 bytes per character, specifying factors such as color, font type, and character size, resulting in over 300 million characters per disc. In both cases, it is possible to manipulate text via software—for example, copying text from disc to external storage.

In the normal mode, text is limited to 40 characters on 20 lines. The high resolution mode allows 80 characters on 40 lines. Compatibility is maintained between the two text formats. A wide range of visual effects are defined, including cuts, scrolls, overlays, dissolves, fades, and so forth.

### ***Interactivity***

The CD-I format permits a high degree of real-time interactivity. The extended data physical format allows for real-time switching between sectors. Thus it is possible to have sectors with different information content and with a different physical format. Sectors may be interleaved at the file level, channel level, or at the coding information type level. File level interleaving allows interspersing of nonrelated information of two or more files in a constant sector stream, thus better utilizing the disc area. Channel interleaving allows real-time selection when one is playing a disc. It is possible to switch between different types of information according to the desired application.

The compact disc real-time operating system (CD-RTOS) is based on the OS-9 operating system, a multitasking operating system suited for real-time capability and synchronizing diverse data. Synchronization is achieved with synchronization control data at the beginning of a real-time record. This indicates how the audio data is synchronized to the video data. The CD-RTOS is composed of four major blocks: Libraries to assure that specialized user functions, math, I/O, and synchronization are available; a CD-RTOS kernel derived from OS-9; managers to define virtual device levels and provide software support; and drivers to interface between virtual and hardware systems.

### ***Implementation***

CD-I is not a peripheral but a self-contained system. To ensure universal disc/drive compatibility, dedicated hardware and interfaces are specified.

A CD-I player contains a CD-ROM disc drive as the system's input, decoder chips for text, graphics, video, and audio, and microprocessor controllers. A CD-I player could be interfaced to a television and stereo. Although it could also be interfaced to a personal computer, it would not be cost-effective, and such interfacing would miss the medium's intent. A CD-ROM drive alone would make a better computer peripheral.

Because the CD-I format recognizes 16-bit PCM data (the first audio quality level), a CD-I player can play regular CD-DA discs. An audio CD player cannot play CD-I discs. As specified in the High Sierra Group proposal, CD-ROM Mode 1 discs can operate on CD-I systems. On the other hand, a microcomputer with a CD-ROM drive is not always capable of processing the information on a CD-I. Universal players capable of playing any kind of CD may appear.

The CD-I disc layout and file structure specify a number of criteria. The first track (track 0) of a CD-I disc is a mandatory CD-I track, complying both physically and logically with the CD-Audio standard, according to which a maximum of ninety-nine tracks may be encoded on a disc. If CD-Audio tracks are also present on a CD-I disc, then the first CD-I track must contain, at the very least, the disc label information. It is recommended that all CD-I information be placed in the first track and that the second and subsequent tracks be used as CD-DA tracks. In this way, if a CD-I disc is played on a CD-DA player, the CD-I track can be skipped.

The beginning of a CD-I track always contains a disc label that contains the file structure volume descriptor, boot record, and terminator record, which are required to initialize the CD-I player. The disc label also contains directory information that specifies the disc type and format and identifies the disc as being individual or part of a collection.

The electrical elements of CD-I player can be considered as five subsystems: central processing unit (CPU), drive interface and control, audio signal processor, video signal processor, and user peripheral controller. These subsystems and associated elements are shown in figure 6.10. The CPU must direct the flow of data through the player and interpret CD-I commands. To accommodate the high data rate, a dedicated CPU is required; the proprietary MC68340 was developed for CD-I applications. It is a derivative of the 32-bit 68020 microprocessor, optimized for embedded control, delivering a computation rate of over 5 MIPS (million instructions per second). In addition, this CPU is designed to be fully static and incorporates low-power modes—an important consideration for portable CD-I players.

The CD drive interface and control system must control the drive mechanism and employ data recognition to direct data to the appropriate area. Audio data and video data are directed to two independent DSP engines, the audio signal processor, and video signal processor. The former han-

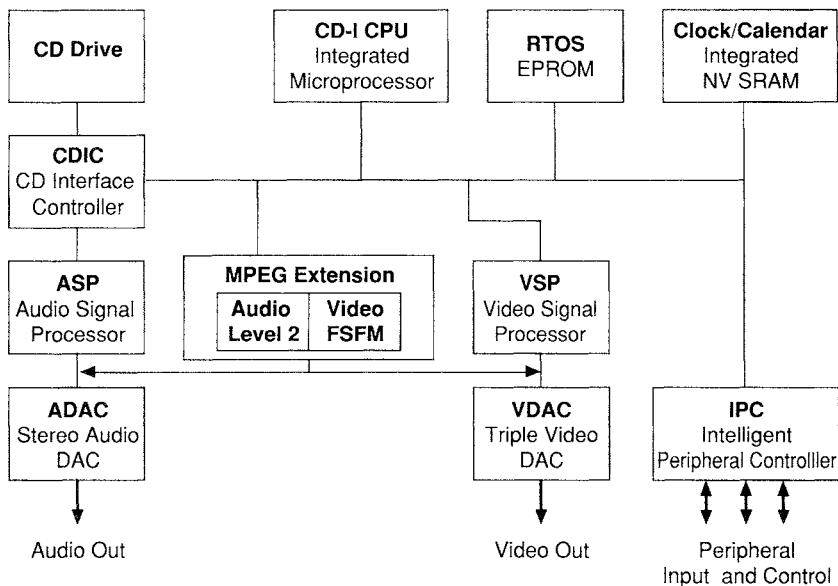


Figure 6.10 *Block diagram showing the electrical subsystems in a CD-I player with the MPEG extension. (Courtesy Motorola)*

handles PCM and ADPCM decoding, while the latter accommodates the wide variety of CD-I video formats. The video processor simultaneously decodes two channels of video in normal or high resolution mode and creates the interleaving of four video planes—a key feature in interactive software. The video processor's output is a digital RGB signal. In players equipped with the MPEG extension, MPEG audio and video data are handled separately by MPEG decoding chips. For example, the MPEG audio data is decoded with a 24-bit DSP56000 chip.

Audio and video data are sent to D/A converters so the information may be reproduced through consumer analog stereos and televisions. The D/A converter provides resolution of 16 bits or more, while the video converter provides 8-bit resolution. The user peripheral controller manages a variety of user controls such as remote control, joystick, and keyboard. A complete MPEG CD-I player contains over 2 million transistors excluding RAM; however, manufacturers have already envisioned a single-chip CD-I player using 0.35-micron CMOS technology on a 7-millimeter square chip, with a chip cost of \$10.

The scope of the CD-I specification is obviously considerable. The system's designers anticipated that a number of diverse players with different performance features and levels would be developed. Therefore, to ensure basic compatibility, the CD-I specification provides a minimum set

of requirements for CD-I systems. The basic system must be able to read the following: CD-DA tracks; CD-ROM Mode 1 tracks (as defined by the High Sierra Group Level 2 format); and CD-I tracks. The basic system must be capable of performing the following functions: decoding CD-DA and CD-I audio information; interfacing with a single TV with normal resolution and any coding, or high resolution with a 4-bit pixel depth and CLUT coding; decoding DYUV; decoding RGB 5:5:5; decoding 8-bit CLUT; decoding one-dimensional run length; handling special effects (e.g., cut, scroll) and multiplane overlays; performing equivalent to or better than a 68000 processor operating at 8 MHz without MMU (10 MHz with MMU) and OS-9-based CD-RTOS with a 16-bit data bus; carrying two DMA channels; storing at least two banks of 512 kbytes each in system RAM and having a direct RAM expansion slot; and providing at least 8 kbytes of nonvolatile memory.

The CD-I format thus presents considerable opportunities for hardware, software, and publishing industries to provide consumers with new forms of interactive entertainment and education. The rapidly emerging CD-I industry has already seen the introduction of numerous titles that take advantage of the video, graphics, audio, and text capabilities supported by this new standard.

## PHOTO CD

The Photo CD system, which was jointly developed by Eastman Kodak Company and Philips, offers a unique CD application to those more interested in photography than audio. This system uses digital imaging technology to store, manipulate, and distribute a consumer's photographic pictures; in particular, it allows existing photographic cameras to capture film images for playback on television screens through the CD-I format. In addition, high-quality prints can be made directly from a Photo CD using thermal printers, ink jet printers, or color copiers. Photo CD images offer sixteen times the resolution capable of being displayed on conventional NTSC or PAL/SECAM televisions. In addition, a Photo CD player will also play back regular audio CDs.

Pictures start with the consumer's 35 millimeter camera; images are processed with standard photographic techniques, and any slide or negative, black and white or color film can be used. These images are then scanned by a Photo CD authoring system to digitize the image. These images are further processed by an image data manager, which uses automatic algorithms to correct for exposure and color balance. These images are written to a Photo CD using the system's CD-R recorder. Up to 100 photographs can be stored on a 120-millimeter diameter CD. In addition, the system generates an index print of all images on the disc, with

a postage-stamp sized print and disc index number. The Photo CD system can also accommodate 120 roll or sheet film formats. Because there is more image data, fewer images can be stored on a disc. However, images can be played on a standard Photo CD player. In addition, hardcopy photographs could be scanned into the Photo CD format.

A CD-I or Photo CD player and television set are used to convert the digital signal into an analog signal for viewing. Users are able to view, zoom in and out, and pan across images stored on disc. Some players offer a digital port for connection to a personal computer. Because the images are compatible with the CD-ROM XA format (Mode 2 Form 1 sector structure), they could be digitally entered into a computer system to perform digital image processing. For example the system could adjust colors, remove unwanted blemishes, add text, rotate images, reduce exposure errors, crop and enlarge images, and create special effects. In addition, Photo CD images could be transmitted via modem or network.

Photo CD images conform to the *Orange Book II* standard and in particular to the CD-ROM XA standard, which is identical to the CD-I Mode 2 Form 1 format. Because images use only limited CD-ROM XA capabilities, images may be read with CD-ROM drives that support a Mode 2 data transfer.

The Photo CD authoring system scans images at 2048 pixels per inch; that is, it reads 2048 scan lines across the short dimension of a 35 millimeter frame, sampling 3072 pixels on each line. This results in a 3:2 aspect ratio, the same used in the 35 millimeter format. For each of the 6.3 million pixels, the scanner measures red, green, and blue colors with 12-bit quantization. This RGB data then undergoes processing with the Photo-YCC color-encoding scheme: A nonlinear transformation is applied to the RGB image data. The resulting values are encoded such that each color pixel is separated into one 8-bit luma component and two 8-bit chroma components. Those intimately familiar with video coding will want to know that a 20 percent reference gray chip is recorded with the following values: Luma = 79, Chroma1 = 156, and Chroma2 = 137. The resulting dynamic range is 350:1. The luma and chroma encoding equations used in the Photo-YCC format follow the CCIR Recommendation 601-1 standards document, and the values of R, G, and B, and nonlinear transformation coding, follow the CCIR 709 video standards document.

A full resolution image with 2048 lines by 3072 pixels by 24-bit coding would require 18 Mbytes of storage. Because the data transfer rate is 153.6 kbytes per second, it would take two minutes to access an image. Thus the Photo CD system employs data reduction and decomposition of the image signal to increase efficiency of storage and decrease data transfer time. During authoring, high resolution image files are reduced to about 25 percent of the original scanned data, depending on the actual image. This quadruples the number of images that can be stored and likewise de-

creases transfer time. Greater reduction ratios would relatively increase transfer time, begin to degrade image quality, and increase hardware cost.

The Photo CD system also employs chroma subsampling; file sizes can be reduced without significant visual loss by taking advantage of limitations in human visual perception. In each image component except 4Base (described below) the two chroma channels are two-times subsampled in both the vertical and horizontal directions. In other words, only one chroma value from each of the channels is present for every four luma values. Because of the nature of the 4Base component, it contains only luma values. With chroma subsampling, a high resolution file is reduced to 9 Mbytes.

High resolution data is stored in a hierarchical form; rather than repeat data already stored in lower-resolution files, only additional data is stored. Thus the data stored in a low-resolution file is omitted from a medium-resolution file, which in turn is omitted from a high-resolution file. This method is known as data decomposition. Three low-resolution image types (Base/16, Base/4, and Base) are stored uncompressed. However, high-resolution image types (4Base and 16Base) are formed by successively interpolating and correcting high resolution residuals (differences from pixels at the previous level of resolution) relative to the Base file. The Base image consists of 512 lines of 768 pixels; it is used to produce images on consumer televisions. The 4Base file consists of 1024 lines by 1536 pixels, and the 16Base file consists of 2048 lines by 3072 pixels; they are used to produce higher-resolution images. The Base/4 file consists of 256 lines by 384 pixels, and the Base/16 file consists of 128 lines by 192 pixels; they are used to produce the indexing photographs as well as picture-in-picture and other applications. Optionally, 64Base and 256Base files could be created for larger film formats. Decomposition permits both rapid image access (30 to 40 seconds to view a Base file) and ability to efficiently store 16Base files—for example, for 20 × 30 inch photographic enlargements.

To display a 4Base image at a resolution of 1024 lines by 1536 pixels, a 512-line by 768-pixel luma Base image is interpolated to form a 1024 by 1536 pixel image. A 1024-line by 1536-pixel residual is then decompressed from its Huffman-encoded form, and its elements are added to each corresponding pixel to form a detailed image. Likewise a 16Base image is recomposed by processing its luma data as well as two chroma channels. The recomposition schemes for the luma and chroma channels are shown in figures 6.11a and b.

Each image is stored in a region called an Image Pac; Image Pacs are stored sequentially, starting with the first image written to disc. For most images an Image Pac contains about 4.5 Mbytes of data. However, depending on the type of film used, and the image itself, it may range from 3 to 6 Mbytes. The structure of an Image Pac (without interleaving for greater clarity) is shown in figure 6.12. The IPA (Image Pac Attribute)

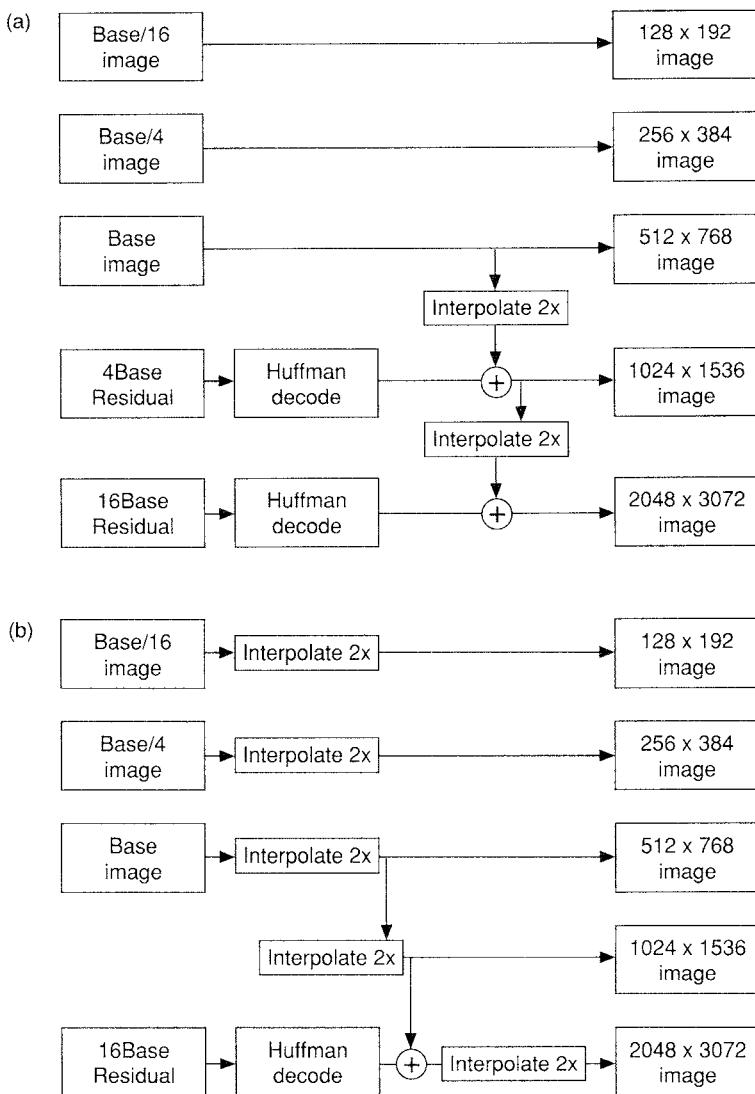


Figure 6.11 *Recomposition algorithm for luma channels (a) and chroma channels (b) as employed in the Photo CD system. (Courtesy Eastman Kodak)*

field is two sectors long; it contains photofinishing and other source information. The Base/16, Base/4, and Base fields contain uncompressed image data. The 4Base and 16Base fields store images in Huffman-encoded form. The IPE (Image Pac Extensions) field is presently unde-

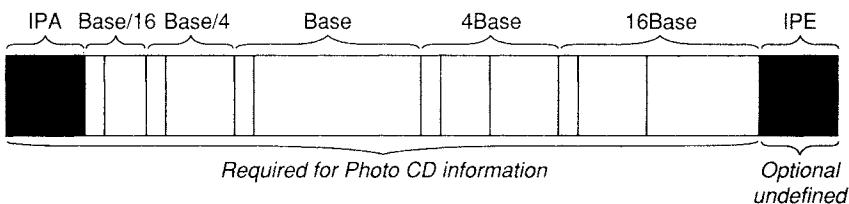


Figure 6.12 *Image Pac data structure used in the Photo CD system. (Courtesy Eastman Kodak)*

fined. Contained within image fields are MRSs (Microcontroller Readable Sectors), which permit low-cost players to read the images. The Base/16, Base/4, and Base data fields are internally defined differently from 4Base and 16Base fields. Every Photo CD also contains an Overview Pac with a copy of the Image Pacs of all the Base/16 images currently on the disc.

Images on a disc are organized as a file system with a hierarchical directory structure; this makes them more accessible to computer systems. All data on a Photo CD is contained in a directory called PHOTO\_CD, which in turn contains directories for the Overview PAC called OVERVIEW.PCD, an information file, called INFO.PCD, for applications programs, and a directory that contains the Image Pacs, called IMAGES. The names of individual Image Pacs include numbers that correspond to the index print numbers.

Photo CD discs themselves conform to the *Orange Book II* standard and are identical to those used in other CD-R applications using, for example, a gold reflective layer and dye recording layer. In addition, because Photo CDs conform to the *Orange Book II* provision that permits additional recording to a disc, a consumer can return a disc to a photo-finisher along with newly exposed film to have additional images recorded to a disc. Pacs initially recorded on a disc are structured as a file using the ISO 9660:1988 Logical Volume and File Structure of CD-ROM for Information Interchange standard. Subsequently recorded Pacs use a CD-WO Volume and File Structures format. All Pacs are addressed through the sector-addressing method used by CD-ROM discs and defined by the ISO/IEC 10149:1989 standard.

To locate images, the player uses the Index Table; it contains disc identification information, pointers to Image Pacs, and other information. Computer operating systems must use data structures defined in the ISO 9660:1988 standard.

Because the CD-I data format is a universal standard, discs are compatible with any Photo CD player. Photo CD players themselves conform to local television standards such as NTSC, PAL/SECAM, etc. In the

future, HDTV Photo CD players will make use of higher resolution data already stored on existing discs. In fact, the 35 millimeter version of Photo CD provides three to four times the resolution required in any proposed HDTV standard.

The Photo CD system thus unifies film, computer, and television technologies. Although many consumers will certainly find the Photo CD system appealing because of high image quality, ability to view photographs on television, and easy storage and access, it is difficult to predict overall success. Because there are 350 million 35 millimeter cameras in use generating 60 billion images annually, if 1 percent of these images were transferred to Photo CD discs, it would result in an annual output of 600 million digital images on CD. The consequent demand for associated products and services could be considerable.

## **CDTV**

The CDTV format is similar to the CD-I format. Developed by Commodore Electronics, it employs the compact disc standard as a basis for multimedia presentations of audio and video including images, graphics, text, and animation. Although CDTV discs are incompatible with CD-I discs, a CDTV player will also play back CD-DA and CD + G/M discs. Applications include reference works, hobby titles, educational discs, and games. A CDTV player connects to a consumer television and stereo system for playback. The CDTV system uses simple point and click operation with a wireless remote controller.

A CDTV player contains a Sony/Philips type CD-ROM drive (Mode 1/Mode 2 standard) supporting CD-ROM, CD-DA, and CD + G/M commands and the ISO-9660 standard. Data readout from the disc is 153.6 kbytes/second in Mode 1 and 175.2 kbytes/second in Mode 2, with 2 Mbytes/second of burst data output. Average disc access time is 0.5 second, and data capacity is 540 Mbytes.

The video display provides 400 lines/vertical frequency of 60 Hz, with a graphics co-processor with beam-synced draw, fill, and move modes. A maximum video memory of 1 Mbyte can be accommodated, as can a palette of 4096 colors. The CDTV system features four standard graphics modes:  $320 \times 200$  noninterlaced with 32 colors,  $640 \times 200$  noninterlaced with 16 colors,  $320 \times 400$  interlaced with 32 colors, and  $640 \times 400$  interlaced with 16 colors. Video outputs include analog RGB, digital RGB (DB-23 connector), composite video NTSC, component video Y-C (S connector for S-VHS and Hi-8), and RF modulated (F connector).

The audio section encompasses that of a conventional CD audio player, with eight-times oversampling, dual D/A converters, and other hardware

features. Maximum audio playing time on a CDTV disc is 28 hours—with AM radio quality sound. The sample rate is variable from the CD-DA rate of 44.1 kHz to 6 kHz.

The CDTV system is based on the Motorola 68000 microprocessor family and also includes three proprietary Amiga multimedia chips which enhance performance by taking over tasks such as audio and video processing, DMA, and graphics display. The memory complement includes 1 Mbyte of RAM, 2 kbytes of nonvolatile RAM reserved for system use, and 512 kbytes of ROM. Players provide industry-standard I/O ports including Centronics parallel interface, RS-232 serial interface, external floppy disk drive interface (Amiga compatible), hardwired connection for keyboard, mouse, or joystick, two audio outputs, and MIDI in and out ports. There is also a port for a 256 kbyte RAM/ROM personal memory card that allows users to save applications, as well as an internal video expansion slot for optional genlock or RF board, and a 15-pin edge connector DMA slot for SCSI, LAN, or other connection.

## DVI

Digital Video Interactive (DVI) is an all-digital optical disc format capable of reproducing full-motion, full-screen video, computer-generated video graphics, and digital audio via a CD-ROM drive. A DVI disc is, in fact, a CD-ROM disc containing DVI-specific data. This data specification permits motion video to be combined with foreground video objects, text, dynamic graphics, and multitrack audio—all under the user's control. Data reduction techniques permit the system to utilize disc storage space more effectively and output relatively large amounts of data more quickly. Although data on DVI discs is formatted to CD-ROM specifications and can be played on a CD-ROM drive (other digital media may be used as well), special DVI decoding technology is also required. The DVI format is thus incompatible with other interactive compact disc formats, such as CD-I and other CD-ROM implementations.

### *Data Reduction*

Working in the David Sarnoff Research Center in Princeton, New Jersey, GE and RCA Laboratories attacked the essential problems of storing video information as CD: too much data in too little space and output at a rate that is too slow. A screen of analog video measuring 512 by 400 pixels requires 600 kbytes of data per frame for digital storage. With the NTSC standard, 30 frames are displayed each second; this works out to about 30

seconds of digital video per CD. Furthermore, because the data readout rate of the CD is only about 150 kbytes per second (5,120 bytes each one-thirtieth of a second), it is not fast enough to show video at real-time speed; it would take over an hour to show that 30 seconds of digital video. In other words, the normal disc playing time limits the transfer rate.

Clearly, storage of raw video data is not the answer. Consequently, developers searched for a method of reducing data prior to storage and restoring it during playback. They conceived the idea of performing non-real-time reduction of video frames, with custom VLSI chips for real-time restoration. Reduction is accomplished before mastering. Since it is done once and need not be done in real time, extensive processing may be used.

Each frame is digitized to 512 horizontal pixels by 480 lines vertical by 24 bits of color. Next, each frame is filtered and subsampled to reduce resolution to 256 by 200 by 24 bits. Filter choices to remove high-frequency data affect quality.

The basis for video data reduction is the fact that one video image is usually related to the previous image. Therefore, rather than storing each image in its entirety, only the difference between successive images is needed. Using a pyramid delta encoding scheme, a video frame is reduced to only 4500 bytes, or about 1100 4-byte words per frame (it is well known that a picture is worth a thousand words). However, the first frame of a new scene requires about 15 kbytes. Required capacity of 4500 bytes per frame represents a reduction to 2.5 percent of the original size. Finally, data is formatted to CD-ROM specifications, and video and audio data are merged. The result is an hour of full-screen digital video with audio storage and a data rate supporting 30 frames per second.

### **DVI Playback**

Upon playback, the reduced data must be restored in real time each time the CD is played. Two VLSI chips form the VDP (video display processor). The first chip, VDP1, is the pixel processor that runs the decoding algorithm and is designed to operate at a rate of 12.5 MIPS (million instructions per second). Its instruction set contains several proprietary video/graphic instructions which permit simultaneous operations. This increases processing power and allows DVI to combine full-motion video with video overlays, text, and dynamic graphics all in the same video frame. A new program can be loaded in 120 microseconds (two scan lines) for rapid updates, a requirement of interactivity.

A second chip, VDP2, handles output display processing. It determines the resolution modes and pixel formats. Its resolution ranges from 256 to 768 pixels horizontally and up to 512 pixels vertically. The VDP2

pixel formats use 8, 16, or 24 bits per pixel; with 24 bits, any one of 16 million colors is available in each pixel.

Multitrack, multichannel digital audio is available on DVI. Up to four tracks are dynamically mixable down to two output channels. Audio encoding utilizes ADPCM, the same method used in CD-I technology, to reduce audio data to about 500 bytes per frame, or less. As is the case with CD-I, DVI offers various audio bandwidth and dynamic ranges. The standard audio data rate is 128 kbytes/second (4 bytes/sample  $\times$  32K samples/second) for an audio bandwidth of 16 kHz. Other rates provide audio bandwidths of 4 kHz and 8 kHz per track. Stereo audio is also available.

One DVI disc could hold a full-motion, full-screen video program ( $256 \times 240$ ) with a playing time of about 74 minutes. With an hour of video, the remaining disc space could be used for ADPCM audio, perhaps using its 4-bit mode. Of course, a higher audio quality level could be traded for less video playing time or reduced video quality. Storage capacity increases proportionately as the size of the video display decreases; up to 2 hours of half-screen video or 4 hours of quarter-screen video may be stored.

A DVI disc could also store still images. With data reduction techniques, a disc could contain 7,000 high resolution ( $768 \times 480$ ) stills, 10,000 medium resolution ( $512 \times 480$ ) stills, or 40,000 low resolution ( $256 \times 240$ ) stills.

### ***Implementation***

DVI is designed with a computer bus interface compatible with several computer architectures, including the IBM PC/AT. The first commercial incarnation of DVI consisted of two plug-in boards designed to fit into an IBM PC/AT. As with CD-ROM and CD-I products, a microprocessor-based hardware player (with disc drive) is required. The result is a personal computer which integrates software programs, text, graphics, sound, and full-motion video. It is the latter which differentiates DVI from CD-ROM systems; without DVI's data reduction methods, video storage is limited in playing time and quality. Of course, the FMV option available in the CD-I standard presents direct competition to DVI.

The video signal on DVI is full-screen and full-motion, but the data reduction methods do have drawbacks. Not all scenes are encoded equally well. Particularly busy scenes suffer picture deterioration. On the other hand, simpler or more slowly moving scenes are very good. In other words, like all technologies, DVI has its limitations and its applications.

DVI thus opens up several opportunities for product developers. For example, the educational market may use DVI as an affordable interactive video system (video tape is cumbersome and video discs are expensive). Training systems must provide a high degree of realism in their

simulations; often the poor quality of graphics-based pictures is inadequate for applications such as medical, aviation, and maintenance training. DVI could remedy this. In another application, many retail stores are turning to in-store marketing with video systems which allow buyers to configure their purchases. DVI could be used by customers in clothing stores and automotive showrooms to check out their order before purchase.

Like CD-ROM, DVI is a format standard for placing any kind of digital data on a compact disc. However, the data format used in DVI is incompatible with other CD specifications, such as the data formats and hardware used in CD-I. Because DVI specifies its own hardware requirements, it stands apart from the rest of the CD family. Recognizing that any format's success hinges on its universality and its acceptance by the manufacturing community, developers have invited participation in developing DVI technology. DVI is intended to be neither a competing format nor a new standard. Rather, it is hoped that DVI chips can be incorporated into existing CD formats, in which case this data algorithm technology could gain wide acceptance.

## **CD-V**

Compact Disc-Video (CD-V) is a combination of audio and video technology—specifically, a combination of the compact disc and Laservision (LV) systems. From this merger comes a new medium, the CDV-Single, and a new hardware product, the CD-V “combi-player,” which is a combination CD and LV player. It plays 12-centimeter video “CDV-Singles,” along with conventional full-length audio-only CDs, and all 20- and 30-centimeter LV video discs as well. The audio portion of a combi-player is identical to that of an audio CD player, and the audio quality of the CD-V duplicates that of an audio CD player. CD-V thus retains full compatibility with existing CD-DA standards.

However, because of incompatible analog television standards, both NTSC and PAL/SECAM players and discs must be produced. Thus CD-V violates the universality of the compact disc—the unique advantage of playability and interchangeability between any discs and any players worldwide. In CD-I, because video is stored digitally, compatibility exists. Because CD-V video information is stored and reproduced in analog form, a CD-V player, as well as CD-V discs, must be configured for one standard or another. In other words, the video portion of a PAL/SECAM CD-V disc is unplayable on a NTSC player and vice versa. In addition, the audio portion of the video program is incompatible because it is integrated into the video signal. The audio-only CD tracks on a CD-V disc are playable on any PAL/SECAM or NTSC or CD-DA player.

### **CDV-Single**

CD-V introduced an entirely new software format, the CDV-Single. The CDV-Single appears similar to an audio CD, although the substrate and top lacquer are gold-tinted to distinguish it visually. The disc is the size of an audio CD but contains approximately 5 minutes of full-motion NTSC video (encoded as an analog signal) with a CD digital audio soundtrack, as well as an extra 20 minutes of digital audio only. A PAL/SECAM standard disc encodes approximately 6 minutes of video material. CD-Audio standards are used for digital audio, including a scanning velocity of 1.2 to 1.4 meters/second and a track pitch of 1.6 micrometers. Because the audio portion of the CDV-Single disc is placed on the innermost portion of the disc (from 50 to 74 millimeters), the audio portion may be played on a regular CD player. Similarly, the CD-V player may play an audio CD. The video signal on the outer portion (from 78 to 116 millimeters) is read only by the CD-V player. The placement of digital audio and video programs on a CDV-Single is shown in figure 6.13.

The video information on the outer portion is encoded as an analog signal according to the LV format. In the NTSC standard, disc speed ranges from 2700 to 1815 rpm, yielding a scanning velocity of 11 to 12 meters/second. In the PAL/SECAM standard, disc speed varies from 2250 to 1512 rpm, with a scanning velocity of 9.2 to 10.2 meters/second. The track pitch is 1.6 micrometers in either case. All CD-V specifications are defined in the *Blue Book* standards document.

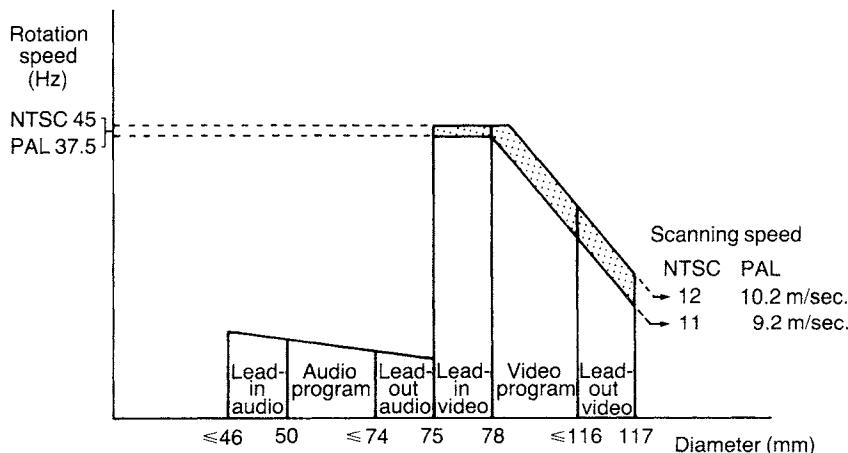


Figure 6.13 Disc layout for a CDV-Single. Digital audio data is contained in the inner area, and analog video data is contained in the outer area.

The player must sense the presence of either digital audio-only or video signal and adjust its operating parameters accordingly. Typically, the player's pickup would read the audio lead-in area, skip to the video portion to begin playback, and then return to the audio-only portion.

### ***CDV-EP/LP***

The 20- and 30-centimeter diameter discs are denoted as CDV-EP (extended play) and CDV-LP (long play), respectively. As with the smaller CDV-Singles, they contain audio conforming to the CD-DA standard and video conforming to the LV standard. Analog audio may be used in lieu of digital audio. The CDV-EP plays two 20-minute sides, and the CDV-LP plays two 60-minute sides. Obviously these discs cannot be played on an audio CD player. Scanning speed in the NTSC format is 10.1 to 11.4 meters/second. In the PAL/SECAM format it is 8.4 to 9.5 meters/second. Because of their long playing times, these discs often contain material such as concert videos, classical music videos, and feature films.

In the NTSC format, the audio portion of LV discs is encoded with two FM carriers placed at 2.3 and 2.8 MHz, which is below the FM picture signal. In the PAL/SECAM format, the FM carriers are placed at 648 and 1066 kHz. In either case, the digital audio signal in CD-V discs is placed in the 1.75 MHz band below the FM picture signal carriers. NTSC discs may therefore be encoded with both analog and digital sound. Digital audio NTSC CD-V discs are playable on existing LV players with analog audio circuitry. An NTSC CD-V player will play existing LV discs. However, in the PAL/SECAM format, the digital audio signal replaces the analog FM carriers. Thus a PAL/SECAM digital audio CD-V disc may not have analog audio as well. A PAL/SECAM CD-V player is able to reproduce both digital audio LVs and analog audio LVs. Older LV players cannot reproduce the audio portion of digital audio CD-V discs.

CD-V players use audio circuitry identical to that of CD-DA players, including, for example, 16-bit D/A converters and four-times oversampling digital filtering. Their features encompass those of both CD and LV players, such as programmability, picture search, fast/slow motion, stills, and speed increase/decrease. Although primarily using constant linear velocity (CLV) scanning, CD-V also provides for constant angular velocity (CAV) scanning.

From a manufacturing standpoint, CD-V discs must be mastered on a dual-speed lathe using video from a 1-inch C-type machine and audio from a 3/4-inch U-matic machine synchronized with time code. The high bandwidth of 10 MHz (versus 1.5 MHz for CD-DA) necessitates manufacturing procedures closer to those of video discs (as opposed to those for audio CDs); still, CD-V discs may be manufactured at audio CD plants.

## CD-R

In its earliest incarnations, the CD was a playback-only medium. However, developments in optical technology soon led to a writable CD format known as CD-R (Compact Disc-Recordable) and also as CD-WO (Compact Disc-Write Once). A write-once optical disc is exactly that. Data is written permanently, can be read many times, but cannot be erased, in much the same way that a photograph is a permanent record. Such systems are sometimes also referred to as WORMs, for *Write Once, Read Many*.

A write-once optical disc may be implemented in a variety of ways. In some systems a process called ablation uses a laser writer to burn holes in a thin film. Other mechanisms include lasers which cause bubbles or blisters to form in the medium. Many manufacturers have selected tellurium and tellurium alloys for the thin film because of their low melting points and high sensitivity. In other systems, an irreversible phase change is used to provide a change in the index of reflectivity of the medium at the point where the laser strikes. Some systems using this method employ a thin metallic recording layer which changes its physical property from amorphous to crystalline when it is thermally heated (to approximately 170 degrees centigrade) by the writing laser. The phase transition alters the reflectivity of the recording layer at written spots, thus allowing laser reading of the data. In some systems a layer of antimony-selenium is employed, evaporated onto a hard plastic disc substrate. An alternative to metal film is the use of polymer dye binder bi-layer media. Colored dyes in a plastic medium over a reflective material are written with infrared light and read with red light. Unlike metal film discs, polymer dye technology can use a flexible substrate.

Philips and Sony, co-inventors of the CD, released a tentative CD-R standard in February 1988. The CD-R format encompasses both CD-DA and CD-ROM applications, and a disc may contain both user-recorded and prerecorded material. To ensure compatibility among users, the format uses a pre-grooved disc. As with other write-once systems, a variety of materials may be used. The recording layer may undergo a decrease or increase in reflectivity during recording. In CD-R, to simplify the servo system design, the high-to-low reflectivity transition materials are used in a narrow pre-groove system, and low-to-high transition materials are used in an alternative wide pre-groove disc. The disc also contains CLV clocking information (as a radial 0.03 micrometer groove wobble of 22.05 kHz) and timecode (as a  $\pm 1$  kHz frequency modulation of the groove wobble) over the entire disc surface.

A CD-R disc contains a prerecorded lead-in area holding the TOC (table of contents). A pre-grooved program area holds user-recorded

(after recording) information; the area consists of an optional prerecorded user table of contents (UTOC) area and recordable user data area. There is also a prerecorded lead-out area. Any prerecorded data, including the optional prerecorded area, would be stored as CD pits. All user data is stored as a reflectivity change. An area called the PCA (power calibration area), located prior to the lead-in area, allows the recorder to automatically make test recordings to determine optimum laser power for data recording. A PMA (program memory area) contains interim TOC data of recorded tracks. Figure 6.14a shows the layout of a CD-R disc without prerecorded tracks, and figure 6.14b shows the layout of a disc with prerecorded tracks.

The prerecorded TOC in the lead-in area contains the start position of the optional prerecorded CD tracks, the position of the UTOC area, the start position of the recordable user data area, the position of the lead-out area, and information on the physical parameters of the recording layer, such as required laser recording power and the polarity of the recorded information (reflectivity increase or decrease). The start position of the user recordable area is also stored in the UTOC.

Both CD-DA and CD-ROM data can be recorded in one track. Tracks in the user recording area can be created at will through pointers in the UTOC. A single track numbering system is used for all prerecorded and user recorded tracks. A disc can contain a maximum of 99 tracks of any duration. A disc may be recorded entirely at once or discontinuously over a period of time, with new tracks starting at the end of previously recorded ones. In terms of disc dimensions, scanning velocity, encoding, error correction, subcode, and data structure, CD-R is identical to audio CDs.

### ***CD-R Applications***

A variety of CD-R recorders are available. All of these recorders contain an audio data encoder and disc recorder. However, some employ an external personal computer to generate subcode data, whereas others are self-contained. In addition, some models permit stop and start recording of tracks, whereas others require the disc to be recorded without pause. In either case, they produce write-once discs that are *Red Book*-compatible and playable on CD players adhering to the *Red Book* standard. In addition to CD-DA discs, these recorders can often produce discs that are compatible with the CD-ROM or CD-I formats. Unlike CD mastering lathes used in pressing plants, these recorders are real-world devices, free of any special requirements such as a clean room environment.

An encoder circuit accepts an input PCM signal and performs CIRC error correction encoding, EFM modulating, and other coding and directs the data stream to the recorder. The encoder's input can be prefaced with

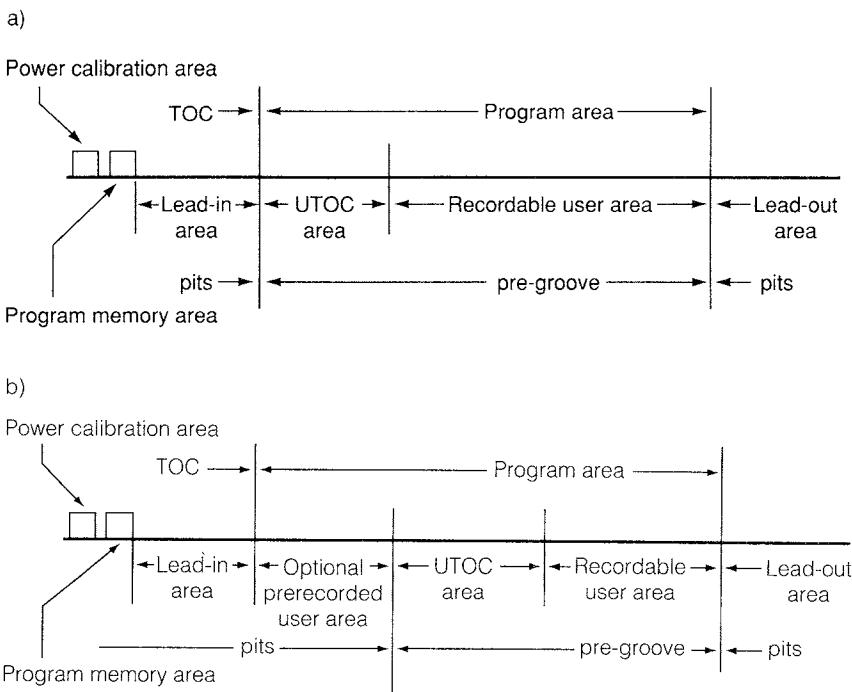


Figure 6.14     *The CD-R format provides disc layouts for discs without prerecorded tracks (a) and for discs with prerecorded tracks (b).*

a number of digital or analog interfaces, depending on the application. As a result, anything from a  $\frac{1}{2}$ -inch analog tape to a  $\frac{3}{4}$ -inch digital tape can be used as a source. The encoder may also contain a built-in synchronizer to read SMPTE timecode. The recorder section accepts audio data and records up to 74 minutes in real time. Recorders are often designed to record both 8- and 12-centimeter discs. Using a 12-centimeter disc, in CD-ROM Mode 1 and CD-I Form 1 applications, storage capability is 650 Mbytes. Storage capability in CD-I Form 2 is 710 Mbytes. An 8-centimeter disc stores 19 minutes, 170 Mbytes, and 190 Mbytes respectively. Generally, recording laser power is between 4 and 8 mW. The block error rate is typically less than 0.03, as measured at the C1 decoder. Recorders sometimes employ a MIDI control bus to allow multiple disc recording. Depending on hardware specifics, multiple recorders can be connected to a single encoder for simultaneous recording.

Some CD-R recorders combine a complete read/write unit, converters, encoder, decoder, and subcode generator in a chassis the size of a CD player. They are self-contained and do not require interfacing to a per-

sonal computer. They can be clocked directly from the digital input signal. A table of contents with track numbers and running times is generated automatically. A complete block diagram is shown in figure 6.15. As one would expect, audio specifications for these types of recorders are top-notch: Frequency response deviation is less than 0.2 dB, THD + N is less than 0.008 percent, channel separation is greater than 80 dB, and phase error is less than 3 degrees. Of course, sampling frequency is fixed at 44.1 kHz. In addition to audio data, a complete subcode table is written in the disc table of contents, and appropriate flags are placed across the playing surface.

Write-once media is manufactured similarly to conventional playback-only discs. As with regular CDs, they employ a polycarbonate substrate, a reflective layer, and a protective top layer. Sandwiched between the substrate and reflective layer, however, is a recording layer composed of an organic dye. Together with the reflective layer (gold is often used) it provides a typical in-groove reflectivity of 70 percent or more—this meets the *Red Book* standard and ensures playability on diversely designed CD players. These discs have precisely the same dimensions as playback-only CDs. Track pitch is 1.6 microns. Unlike regular CDs, a pre-grooved spiral track is used to guide the recording laser along the spiral track; this greatly simplifies recorder hardware design and ensures disc compatibility. Shelf life of the media is said to be 10 years or more at 25 degrees centigrade and 65 percent relative humidity. Life acceleration tests indicate that life expectancy may be decades longer. With proper storage conditions, a 50-year life expectancy can be assumed. However, the dye used in these discs is said to be vulnerable to sunlight. Thus discs should not be exposed to bright sun over a long period. Error rate is less than 20 per second—well below the CD's CIRC tolerance of 220.

The recording mechanism itself may be described as heat-mode memory. The recording layer is actually a photo-absorption surface in that it absorbs energy from the recording laser as heat. Temperature at the focused spot rises above 250 degrees centigrade. This causes the substrate layer to expand into the absorption layer and mix with the dye materials there. Together, the polymer mixed with dye decomposed from heat acts to form pits in the substrate. These pits create the change in reflectivity required by standard CD player pickups. The result is an eye pattern and modulation amplitude essentially identical to that of conventional CDs. This is the case with either single- or three-beam pickup designs. In other words, a CD player cannot distinguish between a mass-replicated CD and a CD-R disc. In theory, a problem would occur only if the CD player failed to meet the CD standard for minimum reflectivity of 70 percent. However, not all hardware manufacturers follow the standard exactly. Thus compatibility problems may exist, as they do with some mass-replicated CDs in some players.

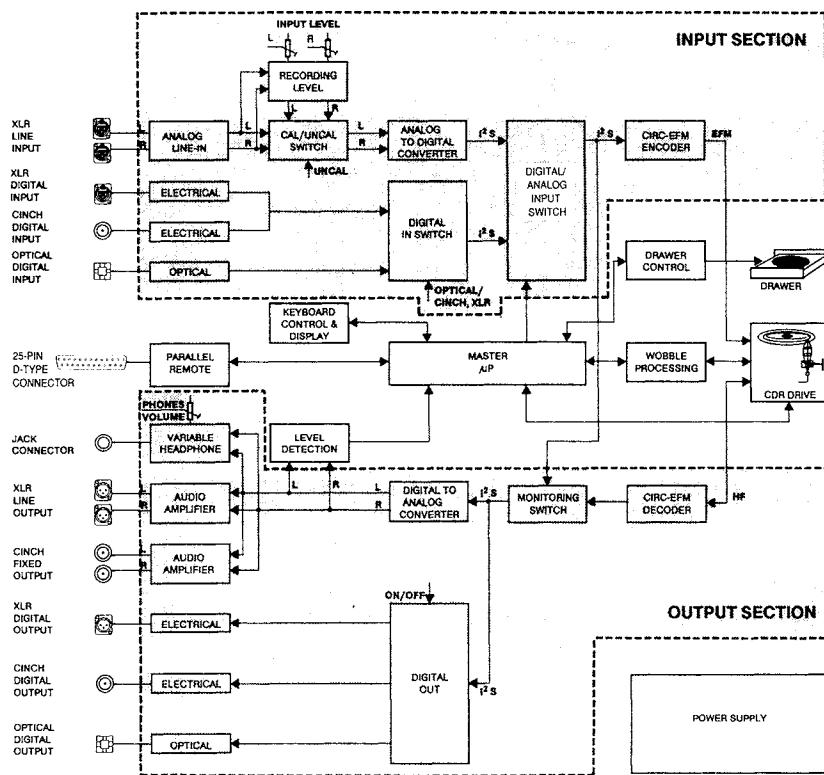


Figure 6.15 Block diagram of a CD-R recorder. (Courtesy Studer)

In some models, PQ subcode editing and generation is performed with a PC software package. The encoder interfaces to a host computer; an IBM PC/AT is commonly used. The PC generates the PQ subcode required by any CD to establish the table of contents (TOC) at the start of a disc and marks tracks and timings throughout the duration of the disc. The PQ editing software package is used to mark start and end points of tracks, set index points, and set intervals between tracks. TOC data can be checked and edited and printed out to form a cue sheet. SMPTE time code can be generated via software and used to synchronize audio and subcode data at the beginning of a recording.

In CD recorders adhering to the *Orange Book I* standard, a disc must be recorded in one pass; start-stop recording is not permitted. In record-

ers adhering to the *Orange Book II* standard, recording may be stopped and started. In many players, tracks may be recorded at different times and replayed, but because the disc lacks the final *Red Book* TOC, it can be played only on a CD-R recorder. When the entire disc is recorded the interim TOC data is transferred to a *Red Book* specification table of contents, and the disc may also be played in any CD audio player. The PMA area located at the inner portion of the disc before the lead-in area accommodates this. It contains the interim TOC record of the recorded tracks.

The ability to add tracks to an already-recorded disc is implemented using the hybrid disc format. A hybrid disc has multiple program areas called sessions. A distributed table of contents structure is used in which each TOC contains a pointer to the start of the next session, as shown in figure 6.16. To locate tracks, the player uses an index table; it contains disc identification information, pointers, and other information. When a session is added to a disc, new directory structures replace previous ones and describe the cumulative files on the disc. Some recorders exceed this specification of the *Orange Book II* standard; they generate an interim TOC which allows partially recorded discs to also be played on playback-only CD players.

Recorders adhering to (or exceeding) the *Orange Book II* standard are particularly useful because they obviate the need to create an edited master tape prior to CD recording. In particular, a recorder can record passages continuously or incrementally, track by track. If a passage is not wanted, it can be marked prior to writing the final TOC so that the CD-R recorder will not play it back. For example, dead air during a live performance or a bad take can be marked so it is deleted whenever the disc is played back. The data physically continues to exist on the disc, however.

The availability of low-cost, one-off CD-R recording systems has revolutionized the audio mastering industry. These CD recorders take editing to the next logical step: the ability to produce a finished compact disc at the end of the editing or pre-mastering session. Significantly, the CD can also be used as the source medium in CD replication. In other words, a recording/editing workstation fitting with a CD recorder forms a compact disc production chain which is entirely free of tape.

This type of system is ideal for short-run CD manufacturing jobs for which conventional CD mastering and replication is either too time-consuming or expensive. In particular, such systems are often used to create reference copies for producers and artists. Although DAT provides a digital reference, many people feel more comfortable with a CD because it has the exact look and feel of the finished product. And of course, CD players are more widely available in home, car, and portable environments where material will be auditioned. Other applications include archiving, film and television sound post-production, and CD-ROM production.

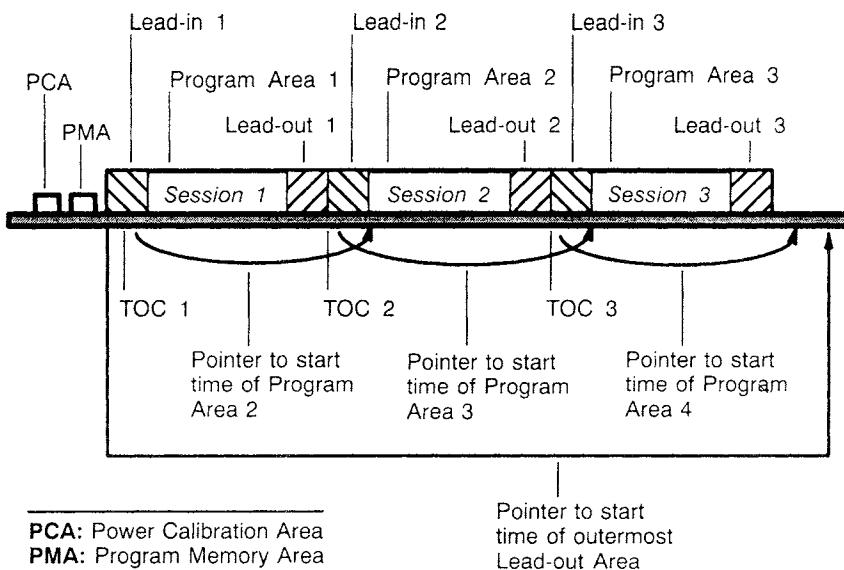


Figure 6.16 The Orange Book II *hybrid disc* uses a distributed table of contents structure in which each TOC points to the start of the next section. (Courtesy Eastman Kodak)

CD recorders are particularly targeted at radio broadcasters; jingles, sound effects, commercials, and music libraries can be recorded to disc for storage and quick access. The ability to incrementally record tracks to disc allows the recorder to function in lieu of a traditional tape cart machine. In addition, recording studios, post-production and mastering houses, record labels, as well as more than a few high-end consumers have all embraced these compact, easy-to-use CD-R recorders. Relatively low-cost hardware and blank media mean that any user who requires a CD can have one—in the time it takes to record it.

## ERASABLE DISCS

Several recordable/erasable optical media technologies have been introduced. These formats provide complete liberty in the reading and writing of data to optical disc. Among them are magneto-optical, phase-change, and dye-polymer recording. Magneto-optics is a system using a combination of vertical magnetic recording and laser optics; it is sometimes called

"optically-assisted magnetic recording." Magneto-optical recording is sometimes referred to as MOR and the discs as MODs.

Magnetic storage is an effective way to record and erase data, but it suffers from basic problems such as medium and head wear. In addition, magnetic storage has hitherto used longitudinal recording, in which the magnetic particles are laid flat across the medium; this limits the density of particles and the amount of information stored in an area. Optical storage offers longevity of medium and pickup and high data density as powerful assets. However, optical properties of materials are not as easily changeable as their magnetic properties. Magneto-optical storage merges the record/erase properties of magnetic materials with the high density and contactless pickup of optical materials.

### ***Magneto-Optical Recording***

Fundamentally, magneto-optical recording is the same as any magnetic recording. However, vertical (sometimes called perpendicular or VR) recording is used. This differs from regular, longitudinal recording. In a vertical medium, magnetic particles are placed perpendicularly to the surface. This allows for much greater particle density, shorter recorded wavelengths, and, consequently, greater recording density. Vertical recording actually becomes more robust as recording density increases; as the cylindrically shaped particles are packed more tightly, they must be made thinner, which increases their magnetic strength. However, this great recording density is not fully utilized by conventional magnetic heads whose recording flux fields cannot be focused sufficiently. In other words, the recorded area uses a far greater area than necessary.

With magneto-optics, a magnetic field is used to record data. By itself, it is too weak to affect the magnetic orientation of the particles. However, a unique property of magnetic materials is utilized: As they are heated, their coercivity (the magnetic field strength required to bring a saturated medium to erasure) suddenly drops close to zero. The temperature at which a material's coercivity drops suddenly is called its Curie point. A magnet, for instance, heated to its Curie point, would lose its ability to attract metals. In the case of magneto-optics, at the Curie temperature (typically 180 degrees centigrade), the magnetic orientation of the particles on the disc are altered by a weak field. A laser beam, precisely focused through an objective lens, is used to heat a minute spot of magnetic material to its Curie point. At that temperature only the few particles in that spot are affected by the magnetic field from the recording coil (perhaps 150 oersteds modulated at a maximum frequency of 720 kHz for EFM modulation), and a high-density recording results. After the laser pulse is withdrawn, the temperature decreases and the data is "frozen" into the magnetic layer. This occurs almost instantaneously because of the very low thermal capacity of the recording layer. This is accomplished in one disc rotation. As in any digital magnetic storage, saturation recording

is used. In this case, the magnetic field of the particles is reverse-oriented, as shown in figure 6.17.

### ***Magneto-Optical Reading***

Data readout utilizes the Kerr effect (or, alternatively, the Faraday effect), which describes the rotation of the plane of polarized light as it reflects from a magnetized material. Specifically, the reverse-oriented regions will reflect laser light differently than the unreversed regions. To read the disc, a laser is focused on it, and the angle of rotation, perhaps varying from 0.3 to 1.2 degrees, of its reflection is monitored, as shown in figure 6.18. An analyzer distinguishes between rotated and unrotated light and converts that information into a beam of varying light intensity. Data is then recovered from this modulated signal. The intensity of the reading laser is much lower (perhaps by a factor of ten) than the recording laser, so the magnetic information is not affected.

To write data again, a magnetic field is applied to the disc, along with the laser heating spot, as illustrated in figure 6.19, and new data is written. However, in some designs an entire track must be erased on one rotation of the disc prior to writing new data; new data can only be written on the next rotation. Updating thus requires two disc rotations, which slows data writing. Other recorders allow the write operation to be accomplished in

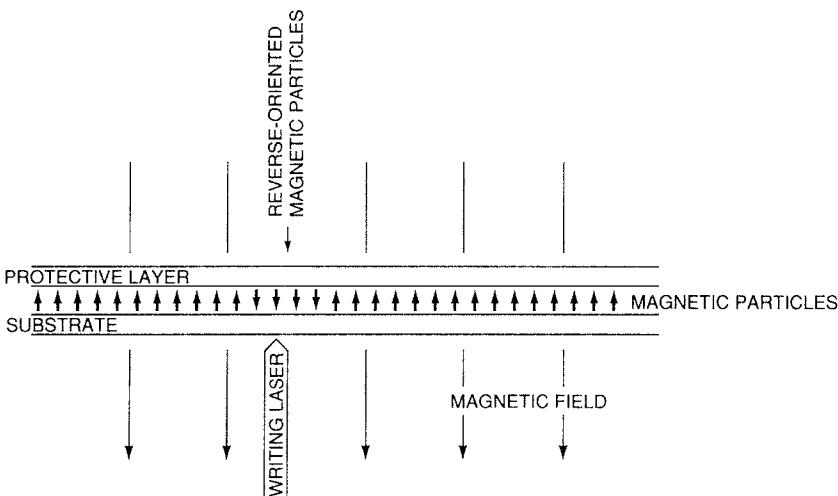


Figure 6.17    *Magneto-optical recording uses a laser beam to heat the recording layer and a magnetic field to magnetically orient the recording particles.*

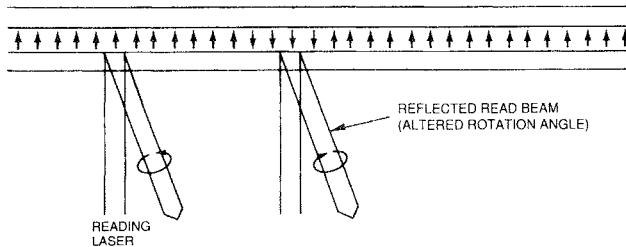


Figure 6.18    *Magneto-optical reading uses the angle of rotation of reflected light.*

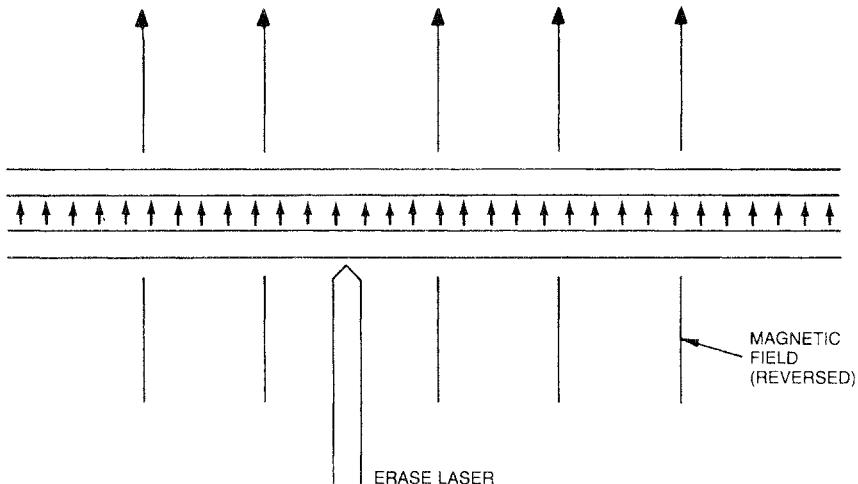


Figure 6.19    *Magneto-optical erasing is accomplished by writing new data.*

one pass, as with conventional magnetic media. Accidental erasure by exposure to high temperatures or strong magnetic fields is very unlikely.

#### *A Magneto-Optical System*

Any recording scheme is dependent on the modulation method used to write the data on the medium. There are several modulation methods suitable for magneto-optics. For example, the specific lengths of the input data code can be converted into a series of pulses which cause the recording laser to switch on and off, recording patterns of varying lengths corresponding to the input signal.

An important aspect of any recording medium is its compatibility with media from other recorders. To achieve this within the high tolerances of a magneto-optical medium, blank discs may be manufactured with pre-recorded and nonerasable addressing. This method, called hardware address sectoring, uses a grooved disc in which address information is physically formed in the groove and detected by light beam reflection. Using this system, any magneto-optical player will automatically track both address and data information contained on any magneto-optical disc. By superimposing the hardware addressing information on the recorded data signal, playing time is not sacrificed. As with standard CDs, MOD discs use a table of contents. However, the table on a MOD disc must be updated each time the disc leaves the record mode.

A magneto-optical disc retains the protective properties of other optical media. The recording layer is sandwiched between a transparent substrate and a protective layer; the thickness of the recording layer is typically about 50 nanometers. The laser light shines through the substrate, using its index of refraction to place surface dust and scratches out of focus with respect to the interior data. Although several magnetic materials can be used, the ultimate selection is based on orientation properties, signal-to-noise ratio, and long-term stability. Magneto-optical discs often use rare-earth transition metals (RETM) such as alloy coatings of gadolinium terbium iron, terbium iron cobalt, and terbium iron. Proprietary processing and chemical treatment have made these reactive elements stable. Tests indicate that a magneto-optical disc can be erased/recorded over ten million times and retain its data for ten years or more.

The optical head and magnetic coils of a magneto-optical system require sophisticated engineering, both in terms of their design and the hardware and software which control them. In addition, a complete signal encoding chain is contained in every recorder. Some of the tasks a magneto-optical recorder must deal with include audio signal processing for both recording and playback; constant linear velocity (CLV) control of the disc; drive control of the optical head block for the focusing servo, tracking servo, and traverse motor; magnetic field control; laser drive control; disc loading and clamping mechanism; and system control of recording, reproduction, erasing, and access. Because of the similarity of their tasks, magneto-optical recorders often borrow technology from the CD system and emulate the CD format in terms of sampling rate and word length.

The CD system is forward compatible with magneto-optical systems. Because much of the electronics are identical, CD format magneto-optical players can also play back CDs, in the same way that a cassette machine can record or play prerecorded tapes. They share a common optical head, and the difference between CD and magneto-optical discs is automatically detected because of the differences in reflectivity. Of course, to record magneto-optical discs, or to play back magneto-optical discs

that are already recorded, a magneto-optical recorder will be needed. Alternatively, CD players could be fitted with MOD pickups.

### *A Professional MOD Recorder*

A number of professional audio manufacturers have introduced disc recorders using 650 Mbyte magneto-optical drives, providing full recording, replaying, and erasing/re-recording of audio data. One model manufactured by Akai Corporation is a four-track recorder manipulating tracks as stereo pairs. One pair of tracks must be recorded simultaneously; however, either one or both pairs may be replayed together. Two track pairs can be recorded to a single track pair on another part of the disc. The dual-sided disc holds 26 minutes of stereo track time per side at a sampling rate of 48 kHz, 29 minutes at 44.1 kHz, and 40 minutes at 32 kHz.

An additional drive can be slaved to the recorder to increase recording time; by overlapping a recording between drives, an essentially unlimited recording time is possible. The disc itself is a standard 120-millimeter disc (looking very much like a CD but incompatible with it) enclosed in a protective plastic sleeve. It is completely removable at any point during a session.

The disc recorder is a rack-mounted unit with a front panel LCD and both hard-wired and soft buttons that provide complete control without need for external keyboards or CRTs. A personal computer interface facilitates use of daisy-chained recorders or applications involving many sound files. There are stereo balanced analog inputs and two pairs of balanced analog inputs; 16-bit converters are employed. In addition, there is an AES/EBU input, two AES/EBU outputs, and a fiber optic input. Other I/O includes SCSI, MIDI, SMPTE, RS422, word sync, and a printer port.

The disc recorder operates much like a tape recorder, with bargraph VU meters, level adjustment pots, and RAM scrubber. But it also provides the advantages of disc recording for random access, computer control for editing, and sampling for synthesis. Soundfiles are created and named on the Record page. Waveforms are displayed on the Edit page, and their horizontal scale may be varied using the Detail key. Edit points can be entered on the fly or input as timecode numbers via the keypad. General purpose marks may be dropped into a file to serve as sync points, for example, to lay in sound effects. A Repeat key is used to replay files. The Song file lets you sequentially combine sound files to create an entire song. The Track file is used to assemble a number of songs into an edited master with fade-up and fade-down, crossfades, inter-track gaps, etc. Alternatively, an edit list may be used to assemble a track from timecode. The Play Sheet lets you directly assign up to nine recordings to the numeric keypad (1 to 9) for instant access. This would permit easy manual

selection and playback as in a cart machine. In addition, these recordings can be called up by their timecode addresses. Hundreds of play sheets can be created and sequentially accessed by pressing the keypad's 0 key. This system can also be interfaced to DSP systems providing additional processing capabilities such as equalization, ambience simulation, and time expansion and compression.

Clearly, this system is most effective when a second disc drive is added to the basic recorder. In this way, large files can be moved from one location to another using the hardware digital bus and software drive assignment in the recorder. The alternative, a DAT recorder, would require separate control. However, although several recorders could be slaved to create a multitrack recorder, the device is not particularly effective in that application; it would make more sense to ally it with a digital multitrack tape recorder. Although not CD-compatible, recorders such as this provide instant and random access recording and editing, disc mastering, and digital optical recording.

## **MINI DISC**

The Mini Disc, developed by the Sony Corporation, is a 2.5 inch, recordable, erasable, optical disc format. It stores 74 minutes of stereo digital audio with a frequency response of 5 Hz to 20 kHz and dynamic range of 105 dB. The sampling frequency is 44.1 kHz. Data is encoded with EFM and error protected by CIRC, and the data track is recorded with constant linear velocity of 1.2 to 1.4 meters/second, depending on playing time. Because the MD format employs data reduction it is not CD-compatible. Remember the first time you saw a CD, and how bulky and clumsy the LP suddenly seemed? Now take a look at figure 6.20.

The MD system employs two kinds of media: a magneto-optical medium for recordable blank discs and a CD-type optical medium for pre-recorded software. Pre-recorded playback-only MD discs can be manufactured using existing CD pressing facilities. The magneto-optical (MOD) technology in MD is similar to others described above but introduces several new features. For example, it allows overwriting, in which previously recorded data can be erased and new data written simultaneously. As with other MOD systems, a magnetic head is positioned over the laser source and on the opposite side of the disc. To record, the laser heats the magnetic surface beyond its Curie point at 200 degrees centigrade so that the polarity of the heated magnetic spot is directed by the bathing magnetic field. As the disc rotates, and the heated spot moves away and cools, the magnetic information is stored. The size of the recorded spot is determined by the reversal cycle of the modulating magnetic field, as opposed to methods in which the laser is turned on and off. Because the laser source is always on, the controlling circuitry is simplified.

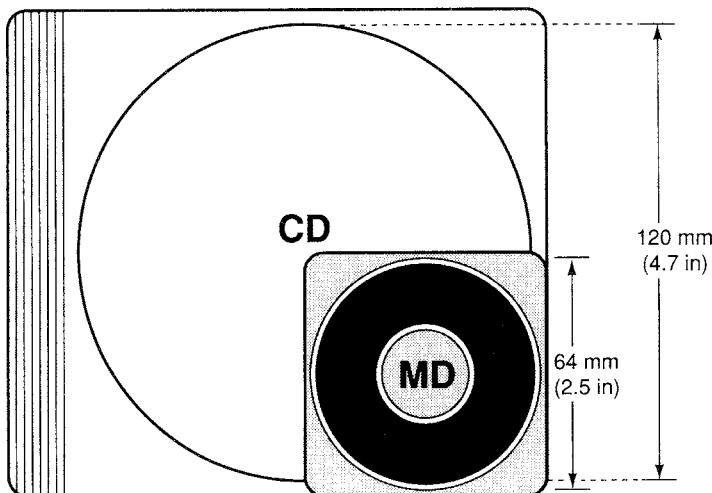


Figure 6.20     *A Mini Disc occupies about one-fourth the area of a compact disc yet provides an identical playing time through data reduction. (Courtesy Sony)*

The MOD disc is built on a polycarbonate substrate with a terbium ferrite cobalt magnetic recording layer covered by a reflective aluminum layer and top protective layer. The terbium ferrite cobalt recording layer changes polarity with 80 oersteds—about one-third the coercivity of other MOD media. This is important because the magnetic head does not touch the medium, and the need for stronger fields at the recording layer would entail higher heat generation and power consumption. The magnetic head itself is said to be particularly power-efficient and able to perform polarity reversals at a rate of 100 nanoseconds/cycle.

The dual-function 0.5 milliwatt laser can operate with both recordable and read-only MD media. Its design is essentially taken from a conventional CD pickup, with the addition of a MOD analyzer. When using a MOD disc, the pickup distinguishes the polarization angle of the reflected light, which is determined by the magnetization of the recording layer. The MOD analyzer converts the polarization angle into a light intensity, and light is directed to two photodiodes. These signals are subtracted to generate a positive or negative readout signal. When playing back a CD-type disc, the pickup reads the intensity of the reflected beam as modulated by the pit surface. The signal from the photodiodes is summed to generate a readout signal. In either case, the optical disc is captive in a protective caddy; the total package weighs about 0.6 ounce. The small disc size means quick access—less than 1 second to any data.

ATRAC (Adaptive Transform Acoustic Coding) data reduction is used to encode data on MD, reducing the data rate by nearly four-fifths, from

1.41 Mbits per second to 0.3 Mbits per second. During recording, analog signals are sampled at a rate of 44.1 kHz and quantized with a conventional A/D converter. The ATRAC encoder divides this PCM data into segments in intervals up to 20 milliseconds long. Fourier transform software analyzes the waveform data in each segment and generates corresponding frequency component data. Using psychoacoustic modeling, the system identifies the audio components that are audible and encodes them, assigning bits as needed according to the amplitude of audible frequency components. Other inaudible material is discarded. Following ATRAC encoding, data undergoes CIRC and EFM encoding and is recorded to disc along with subcode and address information.

Audio data reduction methods such as ATRAC are based on the workings of the human ear and assume that the human ear is capable of processing only a fraction of the sound from the 1.41 Mbits/second data stream from a compact disc. In short, the information capacity of the ear is assumed to be less than the CD standard provides. To ascertain which data may be discarded, psychoacoustic models of the ear have been devised which demonstrate that in theory a data rate of less than 100 kbps per channel may be adequate, if properly coded.

Much about the ear's performance is known. For example, sound below a certain level cannot be detected. Figure 6.21a shows the threshold of hearing curve as it varies across the audio band. Very low and very high frequencies are not audible unless their amplitude is relatively high. Any tones higher in amplitude are audible. Amplitude masking plays a major role in data reduction. Depending on relative amplitude, soft audio tones are masked by louder tones at a similar frequency (the critical bandwidth is about 100 Hz at low frequencies). Figure 6.21b shows that a loud tone shifts the audibility threshold upward, amplitude masking the nearby softer tone. Because the softer tone is now inaudible it can be ignored by the data reduction scheme and not coded. In the ear's critical frequency area and loudness level (1 to 4 kHz and 70 to 110 dB) differences between tones of 0.5 dB may be audible; however, in other areas the ear is much less sensitive.

Similarly, temporal pre-masking and post-masking take place. A louder tone appearing just before (within 3 milliseconds) or after (within 200 milliseconds) a softer tone overcomes the softer tone through temporal masking. Moreover, as overall sound level increases, the ear is relatively less sensitive, shifting the overall threshold curve upward. Using all of these dynamically changing psychoacoustical cues, inaudible components can be removed with minimal sound degradation.

The success of an audio data reduction system depends on the acuity of an individual's auditory perception. This is often difficult to assess, particularly because of the ear's ability to learn over time and sharpen its acuity. Success also depends on the type of audio program being processed. In general, audio signals contain less information at high frequen-

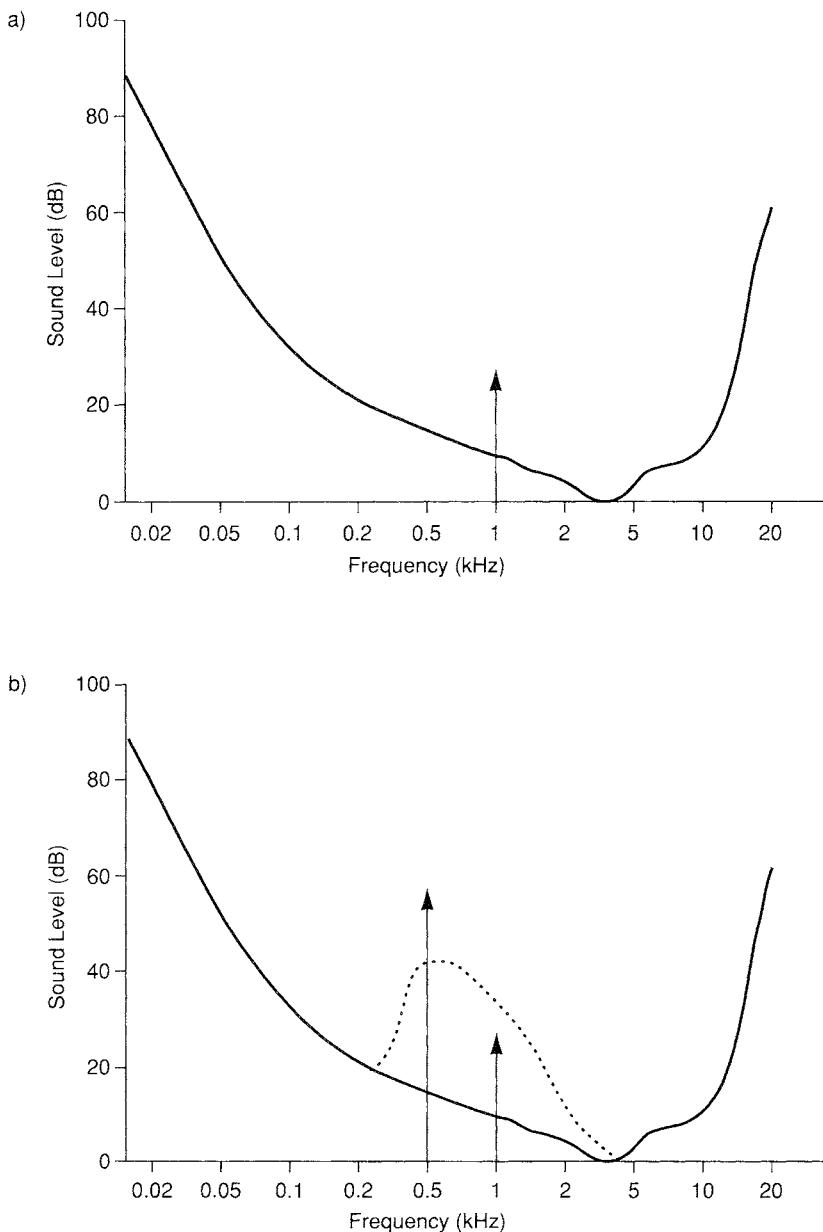


Figure 6.21 *The audibility limits of the human ear serve as a guide for psychoacoustic data reduction systems. Only sounds which exceed the threshold of hearing (a) are audible to the human ear. A loud sound causes the threshold of hearing curve to shift upward (b) thus masking an otherwise audible sound.*

cies. However, every type of music presents a different set of conditions to a data reduction scheme (and the ear). For example, it has been statistically estimated that 64 percent of the sound of a Stravinsky wind ensemble piece is masked and 73.4 percent of a Mozart wind ensemble piece is masked and hence inaudible to the typical listener.

During playback, following CIRC and EFM decoding, frequency information is deciphered by an ATRAC decoder and the 20-millisecond intervals are reconstructed into digital waveform data. This data is then output through a D/A converter. Because the data in the input and output sections of a MD recorder adheres to the standard 16-bit, 44.1 kHz data format, data can be digitally copied into and out of the MD recorder using CD and other digital media. The MD standard includes the Serial Copy Management System (SCMS) in which first-generation digital copies are enabled but not second-generation copies. In other words, you cannot make a copy of a copy.

Data reduction provides another important benefit. The pickup reads data from the disc at a rate of 1.41 Mbits/second. However, the ATRAC decoder requires only 0.3 Mbits/second. This low decoding rate permits efficient use of a look-ahead buffer; a 1-Mbit memory chip placed between the pickup and decoder stores 3 seconds of real-time audio. Data enters the buffer faster than it leaves; when the buffer is full, the pickup stops reading data until the buffer is ready to accept more data. If the player mistracks, the pickup has ample time (perhaps 3 seconds) to return to the correct tracking position. This is facilitated by a sector repositioning method in which address information is placed in the recorded bit stream every 13 milliseconds. When the pickup mistracks, the system detects the wrong address and returns the pickup to the correct address position. In other words, thanks to the low data rate, look-ahead buffer, and sector repositioning, the MD is very immune to shock and vibration during both recording and playback.

One question you might ask: Why develop an entirely new format incompatible with recordable CD? First, MD is a more portable product because of smaller disc diameter, address sectoring, and data buffering. More importantly, record labels are not anxious to promote any recordable format that matches the sound quality standards of the professional master recording. They are more supportive of new formats of slightly lower sound quality (specifically, noncloned, data-reduced music). The MD format provides random access, durability, portability, convenience, shock resistance, and recordability. Analog and digital cassette tapes come close but ultimately fail, especially in terms of durability and random access. The CD fares well in these criteria but is not as portable as one would like. MD merges analog cassette tape (emphasizing the importance of recordability and portability) and compact disc (durability and random access), resulting in a high-quality portable, recordable medium. Argu-

ably, the CD is thus the heir apparent to LP, and MD is the heir apparent to the analog cassette. Of course, proponents of the DCC (digital compact cassette) format would argue otherwise, particularly given the ability of DCC hardware to play back analog cassettes.

### ***Other Recording Technologies***

Magneto-optics is not the only optical recording system in use. Some systems use phase-change technology similar to that used in write-once systems. They use materials that exhibit a reversible crystalline/amorphous phase change when recorded at one temperature and erased at another. Typically, a high-reflectivity (crystalline) to low-reflectivity (amorphous) phase change is used to record data and the reverse to erase. Information is recorded by heating an area of the crystalline layer to a temperature slightly above its melting point. When the area solidifies, it has become amorphous, and the change in reflectivity can be detected. Because the crystalline form is more stable, the material will tend to change back to this form; thus, when the area is heated to a point just below its melting temperature, it will return to a crystalline state, erasing the data. A recording layer composed of gallium antimonide and indium antimonide has been developed. It has a long shelf life and is not affected by ambient temperatures and humidity. Over a thousand erasures can be achieved. Permanent recording may be achieved by simply increasing the power of the writing laser; this burns holes in the recording layer rather than changing its phase.

In another erasable technology, a dye-polymer recording layer is physically changed when recorded and changed again when erased. A dye-polymer disc is made by sandwiching two layers of plastic materials coated with a dye that absorbs laser light. The medium is written by laser heating. Laser light of a specific wavelength heats the two-layered medium, causing the lower layer to expand and push up the top layer. The top layer cools faster than the lower and retains the raised bump. This formation can be read by the read laser. Erasing and rerecording are easily accomplished: When the surfaces are again heated—this time with a laser of a different wavelength—the bump flattens out again, leaving the surface in its original form, ready for rewriting of new data.

The recording surface forms a bump with one laser wavelength and flattens out with another. Moreover, the wavelength that flattens a bump does not affect the surface if it is already recorded. Thus data can be overwritten on a single disc revolution as the medium passes under the smoothing laser just before it encounters the laser writing new bumps.

Significantly, because the recording medium is changed physically (that is, a pit is created), the disc can be read by a conventional CD player. Similarly, a conventional prerecorded CD can be played in the recorder. Thus the new technology is backward compatible.

However, a problem with erasable dye-polymer technology is that the layer can be erased only a few times, perhaps a hundred or less, before the recording layers become fatigued. While perhaps acceptable for some applications, this is certainly not acceptable for others. That problem, of course, has already been solved by magneto-optical recording.

Although the future is bright for recordable CDs, it is certain that playback-only CDs will remain popular as a medium for prerecorded optical reproduction. While many people will switch to CD-R and future CD-E recorders, one suspects that CD players will keep going strong, just as another great prerecorded software transport, the turntable, persisted for a number of years.

### **CD + G/M**

As noted in Chapter Four, every compact disc contains a subcode area. In addition to the P and Q channels used for conventional audio playback, there are other channels available for other applications. These six channels (R,S,T,U,V, and W) account for about 20 Mbytes of 8-bit storage per disc and are arranged in 98-bit words, as are the P and Q channels. The R-W subcode area was reserved as contingency storage by the original disc designers.

Compact Disc + Graphics/MIDI (CD + G/M) takes advantage of this nonaudio data area. Still color images, text, or other material may be stored on an audio compact disc and displayed on a television monitor while the music plays. Because this data area is already reserved in the CD-DA standard, no music playing time is sacrificed when graphics data is inserted. The graphic and text information on a CD + G/M disc is universally playable on NTSC, PAL/SECAM, or any television system. When a disc does not contain graphics information, the R-W bits are recorded with zeros and do not affect playback.

For video applications the data in the R-W subcode channels can be collected over thousands of subcode blocks to form a video image. A CD holding a full program of audio data could hold hundreds of still video images, drawing a new image from the R-W bits every few seconds. All synchronization, error correction, and video data are self-contained in the subcode field. However, appropriate editors and master tape recorder formats are required to encode the information onto the PCM tape master.

The possibilities for R-W graphics are varied; still pictures relating conceptually or directly to the audio material can be displayed, as can liner notes, lyrics, or other alphanumeric information. An opera recording may display the libretto in several languages. A popular disc might have pictures of rock 'n' roll stars. In some cases, CD-DA discs with CD + G data are known as CD-I-ready discs, because this auxiliary data is also readable by CD-I players. Other applications include display of guitar or

keyboard chords or MIDI (Musical Instrument Digital Interface) information, used to control external synthesizers.

In CD + G/M the R-W channels are grouped to form 6-bit R-W symbols. To provide useful information, these channels are accumulated over 98 frames. The subcode synchronization word occupies the first two frames; thus a subcode block contains 8 channels with 96 data bits per channel, as shown in figure 6.22. This data block is called a "packet," which is divided into 4 "packs." A pack is generated every 3.3 milliseconds.

The internal structure of each pack is shown in figure 6.23a. The Mode and Item switches are each 3 bits; five different Mode/Items are presently defined, as shown in figure 6.23b. Mode 1/Item 1 specifies TV graphics information, and Mode 3/Item 0 defines MIDI information. The "Instruction" frame further defines the nature of the "Data Field." The Q and P frames contain error detection and correction codes independent of the audio data's Reed-Solomon code. Specifically, the Q code protects the Mode/Item and Instruction fields, and the P code protects all data including the Q code. The Mode/Item and Instruction fields are thus doubly protected. The player decodes the G/M data separately from the audio data.

#### *Mode 1 Format*

In the CD + G/M Mode 1 format, the TV graphics mode uses a bit-mapped technique which displays text and graphics with image units called "fonts"; they comprise 6 by 12 pixel character units. A full screen consists of 50 vertical by 18 horizontal fonts. Outer rows and columns are intended for scroll action and are not visible. Thus the total visible picture contains

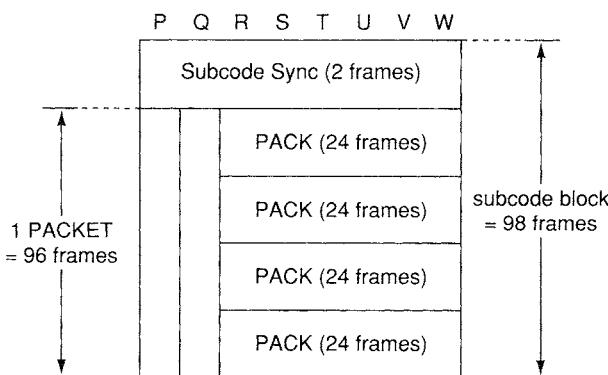


Figure 6.22    A CD + G/M subcode packet contains four data packs, generated at a rate of one every 3.3 milliseconds.

frame	R	S	T	U	V	W	MODE	ITEM	contents
0	MODE	ITEM					0	0	No data (Zero)
1			Instruction				1	0	Line Graphics
2-3			Q0, Q1 (2 frames)				1	1	TV Graphics
4			Data Field				3	0	MIDI Mode
↓							7	0	The User Mode
19									
20-23			P0-P3 (4 frames)						

Figure 6.23 *Each data pack is composed of six data types. The Mode and Item data are used to define the information type.*

$(50-2) \times 6 = 288$  vertical by  $(18-2) \times 12 = 192$  horizontal pixels. Font data is given a channel number from 0 to 15, and each channel can store individual images, for example:

- channel 0: graphics
- channel 1: English lyrics
- channel 2: French lyrics
- channel 3: German lyrics
- channel 4: guitar chords
- channel 5: rhythm guide

Color is obtained through the use of CLUT (color look-up table), and each screen can display 16 colors from among 4096 colors, with 4-bit gradations for each primary RGB color. In addition, the CD + G instructions can preset a memory with color, preset the border with color, write a font using two colors for foreground/background, exclusive-OR a font with two colors, load the CLUT, and soft scroll the screen.

CD + M adds MIDI information to the subcode area to control MIDI instruments or other music processors. Applications include “music minus one,” re-orchestration and restructuring of songs, printing or display of a score, revoicing of parts in an arrangement, change of tempi and key signature, and creation of a personal music database. MIDI messages are sent over any of 16 MIDI channels, using the signal format shown in figure 6.24 and adhering to MIDI protocol. The hardware MIDI interface

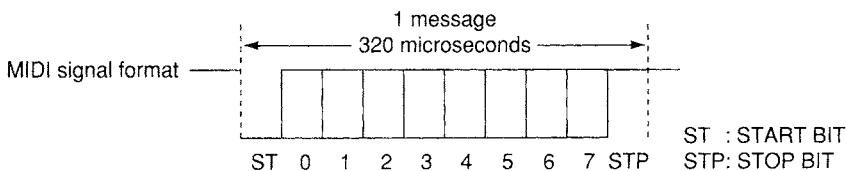


Figure 6.24     *MIDI messages are composed of a start and a stop bit and 8 data bits.*

operates at 31.25 ( $\pm$  1 percent) baud, asynchronous, using a start bit, 8 data bits (D0 to D7), and a stop bit. This yields a total of 10 bits for a period of 320 microseconds per serial byte. The start bit is a logical 0, and the stop bit is a logical 1. Bytes are sent LSB first. Communication uses multibyte messages consisting of 1 status byte followed by 1 or 2 data bytes (real-time and exclusive messages excepted). All common MIDI messages such as Note Off, Note On, Polyphonic Key Pressure, Control Change, Program Change, Channel Pressure, Pitch Bend, and system common messages and system real-time messages are supported in the CD + M format.

MIDI data is recorded into CD subcode using the same data packs as in CD + G, with MIDI data arranged in bytes 0 through 12 within packs, as shown in figure 6.25. To maintain the baud rate specified by the MIDI standard, 12 consecutive packs may contain no more than 125 MIDI bytes. If playback on equipment with up to + 12.5 percent pitch bend is used, it is recommended that the data be limited to a maximum of 110 bytes in any 12 consecutive packs.

### **Mode 2 Format**

Although the Mode 1 CD + G/M standard is quite suitable for many applications such as MIDI, its bit-mapped video presentation is limited. In many video applications, because of low text transmission speed, slow random access to information, and low picture quality both in terms of resolution and colors, Mode 1 is simply not adequate. Thus a Mode 2 CD + G/M standard was subsequently devised; its principle enhancement is data reduction. Because less data is required to record the same amount of information, more information can be recorded, identical data may be recorded at several locations on the disc and efficiently buffered for faster access, and readout is relatively faster, even though the data rate from the subcode channel is fixed. In addition, Mode 2 provides more extensive control sets for flexible menu capability.

FRAME	R	S	T	U	V	W	
0	0	1	1	0	0	0	: MODE/ITEM
1	0	0	n3	n2	n1	n0	: BYTES: n = Number of MIDI Bytes
2,3	PARITY Q0,1						
4	B	1					: MIDI Byte 1
5			B	2			: MIDI Byte 2
6					B	3	: MIDI Byte 3
7							
8	B	4					: MIDI Byte 4
9			B	5			: MIDI Byte 5
10					B	6	: MIDI Byte 6
11							
12	B	7					: MIDI Byte 7
13			B	8			: MIDI Byte 8
14					B	9	: MIDI Byte 9
15							
16	B	10					: MIDI Byte 10
17			B	11			: MIDI Byte 11
18					B	12	: MIDI Byte 12
19							
20	PARITY P 0 ~ 3						
23							

• Any 12 consecutive  
 packs may contain  
 max. 125 bytes  
 • All bits of unused  
 bytes are zero

Figure 6.25 In the CD + G/M format, MIDI data is arranged as 12 MIDI bytes per data pack.

The reduction methods employed in the Mode 2 format borrow from existing standards. Character codes for ASCII text use the ISO 8859-15 standard, mosaic and line graphics use the CCITT SG-100 video-text standard, and run-length codes and natural image data compression use the CD-I standard. Mode 2 provides three kinds of control sets to facilitate user interaction. The first control set contains GO TO commands; they can invoke a CD absolute time code or picture number in seek functions or select information from a cache memory. The second control set contains functions such as display on/off, select either text only or with underlaid video, scroll up/down, and select cursor modes and default positions. The third control set contains hardware commands such as play, pause, scan, and audio/video on/off.

In both Mode 1 and Mode 2, CD + G/M data is stored digitally and uses universal text/video formats independent of the television format in which the display is limited to 40 characters by 21 lines. As noted above, there are sixty-eight 6-bit symbols per packet, which is equivalent to fifty-one 8-bit bytes per packet. ASCII coding requires 1 byte per character. Thus 40 bytes are needed to encode one line of text. The remaining 11 bytes are used for two other fields: a 9-bit line-related control code and a 2-bit byte index number used to control a buffer memory. Packets can also control other information such as commands and other data.

The format allows information to be grouped into 63 main menus and 255 submenus per main menu. Menus can be used for a variety of applications, including separation of information by language. There is also a main\_submenu override for information that is accessible from all menus. In addition, branching is possible within the menu\_submenu structure. Menus are interleaved in the data stream and during playback withdrawn from the data stream and placed in a cache memory. This permits fast access to all menus as well as access even when the player is in a pause mode. Even if the cache memory's size is insufficient to simultaneously store all menus, any menu can be retrieved from the data stream. Mode 2 also allows use of blinking hot spot fonts, with a maximum of 420 ( $20 \times 21$ ) individual hot spots on a screen. Hot spots can be grouped to create any shape.

CD + G/M mastering requires use of a personal computer and appropriate hardware and software. For example, an IBM PC/AT computer with EGA board could be used to create all text and attribute and command data and output a CD A-timecode-accurate R-W data stream which would be merged with other PQ subcode data and placed on a master disc during lead-in and program areas.

CD + G/M players have a variety of outputs, including video, S-video, MIDI-OUT, and, of course, audio. European versions have a Euroconnector for digital RGB output. Alternatively, CD graphics data could be output from a player using a digital output port on the CD player. An

external decoder/adaptor box would accept the subcode signal from the player and convert it into a modulated video signal suitable for the television. In addition, all CD-I players play back CD + G information.

Clearly, CD + G/M is a “value-added” kind of feature; few people would buy a CD player for these options alone, but they do enhance a CD player’s desirability and utility. Conversely, this added information increases the salability of disc titles.

### CD-3

The data format of the CD-3 disc is identical to that of regular 12-centimeter diameter CDs; however, its diameter is only 8 centimeters, or about 3 inches. Its small size promotes greater portability. In addition, the CD-3 format can be used for CD-ROM applications. Because the CD data spiral begins at the innermost diameter, and all CD formats use this standard dimension, CD-3s are compatible with regular discs and players. Most players feature a disc drawer with concentric wells able to accommodate both disc diameters. In some cases, an outer adaptor ring must be used to enlarge the CD-3’s size so that players can properly load them.

Naturally, the CD-3 holds less music because of its smaller size; its maximum playing time is about 20 minutes. It is thus designed to replace the analog 7-inch, 45 rpm single. A variety of CD-3 portable players has appeared, achieving truly shirt pocket dimensions. In some cases, these tiny players are also able to play 12-centimeter discs by sliding the disc spindle motor to one side.

Portable CD players also come packaged along with amplifiers, loudspeakers, cassette recorders, radios, graphic equalizers, and other assorted audio gear, as well as a handle. These “boom boxes” are useful for applications in which you are eager to share your musical tastes with as many other people as possible, such as at the beach or on the subway.

### FOR FURTHER READING

Bouwhuis, G., J. Braat, A. Huijser, J. Pasman, G. van Rosmalen, and K. Schouhamer-Immink. *Principles of Optical Disc Systems*. Bristol and Boston: Hilger, 1985.

Bouwhuis, G., and P. Burgstede. “The Optical Scanning System of the Philips ‘VLP’ Record Player.” *Philips Technical Review* 33(7):186–89, 1973.

Brandenburg, K., and D. Seitzer. “Low Bit Rate Coding of High Quality Digital Audio: Algorithms and Evaluation of Quality.” In *Proceedings of*

*the AES Seventh International Conference—Audio in Digital Times.* K. C. Pohlmann, chair. New York: Audio Engineering Society, 1990.

Compaan, K., and P. Kramer. "The Philips 'VLP' System." *Philips Technical Review* 33(7):178–80, 1973.

"CD-I—The Future" (Conference Proceedings). OnLine Conference and Exhibitions, San Francisco, May 11–13, 1987.

Cornyn, S. "CD-I and the Media." In J. Strawn, ed. *Proceedings of the AES Fifth International Conference—Music and Digital Technology.* New York: Audio Engineering Society, 1987.

Denon Nippon Columbia. *CD-ROM Disc.* Technical information. March, 1986.

———. *CD-ROM Drive.* Technical information. March, 1986.

———. *Controller for CD-ROM Drive (SCSI/SASI Protocol).* Technical manual. March, 1986.

Dixon, D. F., S. J. Golin, and I. H. Hashfield. "DVI Video/Graphics." *Computer Graphics World*, July, 1987.

Eastman Kodak. "Photo CD: A Planning Guide for Developers," Rochester, N.Y., May, 1991.

Feiten, B. "Spectral Properties of Audio Signals and Masking with Aspect to Bit Rate Reduction." *AES Preprint no. 2795*, 86th Convention, 1989.

Fielder, L. D., "Human Auditory Capabilities and Their Consequences in Digital-Audio Converter Design." In *Proceedings of the AES Seventh International Conference—Audio in Digital Times.* K. C. Pohlmann, chair. New York: Audio Engineering Society, 1990.

Freese, R. P. "Optical Disks Become Recordable." *IEEE Spectrum* 36(2):41–45, February, 1988.

Hartmann, M., B. A. J. Jacobs, and J. J. M. Brant. "Erasable Magneto-Optical Recording." *Philips Technical Review* 42(2):37–47, August, 1985.

Isailovic, J. *Videodisc and Optical Memory Systems.* Englewood Cliffs, N.J.: Prentice-Hall, 1985.

ISO (International Standards Organization). "Volume and File Structure of CD-ROM for Information Interchange." *Standard ISO/DIS 9660.*

Janssen, P. J. M., and P. E. Day. "Control Mechanisms in the Philips 'VLP' Record Player." *Philips Technical Review* 33(7):190–93, 1973.

- Lambert, S., and S. Ropiequet. *CD-ROM, the New Papyrus*. Redmond, Wash.: Microsoft, 1986.
- Lambert, S., and J. Sallis, eds. *CD-I and Interactive Videodisc Technology*. Indianapolis: Sams, 1987.
- Nishiguchi, M., K. Akagiri, and T. Suzuki. "A New Audio Bit Rate Reduction System for the CD-I Format." *AES Preprint no. 2375*, 81st Convention, 1986.
- Philips International. *A General Introduction to CD-Interactive*. October, 1986.
- . *Optical Media in Perspective*. March, 1987.
- . "The CD Video System." Press release, Eindhoven, March, 1987.
- Philips International, ed. *Compact Disc-Interactive: A Designer's Overview*. New York: McGraw-Hill, 1988.
- Ropiequet, S. *CD-ROM Optical Publishing*, Volume 2. Redmond, Wash.: Microsoft, 1987.
- Rosen, D. "GE/RCA Announces DVI." *CD-I News* 1(6):1, 4–5, 8–9, 1987.
- Roth, J. P. *Essential Guide to CD-ROM*. Westport, Conn.: Meckler, 1986.
- Rothchild, E. S. "Optical Memory: Data Storage by Laser." *Byte* 9(11):215–24, October, 1984.
- Schouhamer-Immink, K. A., and J. J. M. Braat. "Experiments toward an Erasable Compact Disc Digital Audio System." *Journal of the Audio Engineering Society* 32(7/8):531–38, July/August, 1984.
- Schouhamer-Immink, K. A., A. H. Hoogendijk, and J. A. Kahlman. "Digital Audio Modulation in the PAL and NTSC Optical Video Disk Coding Formats." *Journal of the Audio Engineering Society* 32(11):883–88, November, 1984.
- Schwerin, J. B. *CD-ROM Standards: The Book*. Oxford (England): Learned Information, and Pittsfield, Vt.: Info Tech, 1986.
- Thomas, G. E. "Future Trends in Optical Recordings." *Philips Technical Review* 44(2):51–57, April, 1988.
- Van den Bussche, W., A. H. Hoogendijk, and J. H. Wessels. "Signal Processing in the Philips 'VLP' System." *Philips Technical Review* 33(7):181–85, 1973.



## *Chapter Seven*

# Disc Manufacturing

### **INTRODUCTION**

The compact disc has proved to be a technological wunderkind in the highly sophisticated and competitive field of music and data storage. Annual worldwide demand for CDs passed the 1 billion mark in 1991. Established injection molding factories are expanding, and new facilities are coming on-line. In this chapter we explore the CD manufacturing chain.

### **PRE-MASTERING**

Most CDs begin life as a master tape, and any CD is only as good as its master recording. Because mastering to CD cannot eliminate the faults on the master, the master must be selected carefully. With analog tapes, the original master should be used, never a safety copy. If the master is of poor quality, one should consider remixing. If such is the case, one should locate the original equalization settings; if they were used to compensate for LP deficiencies, new equalization should be considered. Never use a submaster equalized for LP. If the master is digital, one must avoid any transfer to the analog domain; for example, one should equalize digitally if possible. In accordance with *Red Book* standards, facilities can produce only CD-DA discs which are playable on any stereo or mono playback system without requiring modification. The audio program must fall into one of these categories: (1) regular stereo; (2) one mono signal, simultaneously encoded on both audio channels; and (3) one mono signal on one channel and no signal on the other channel. Thus the use of two audio channels for two different mono programs is not allowed. Nevertheless, some dual-mono CDs have been released.

In pre-mastering, the original master tape is copied to a specially prepared tape master, which is used to produce the master disc prior to replication. It is the first step in CD replication. The original tape (of any format) is referred to as the master tape. By means of a digital audio processor, the audio program is copied by a  $\frac{3}{4}$ -inch U-matic tape referenced with SMPTE timecode. This tape is called the CD master tape. When subcode data is added, it is called a CD tape master. This is the tape used to cut the CD master disc. The digital audio program is recorded on the helical scan video track, subcode is recorded on longitudinal channel 1, and timecode is recorded on longitudinal channel 2. Figure 7.1 illustrates the three main steps in tape mastering. Although mastering facilities will accept master tapes (of various formats) along with required documentation, pre-mastering may be accomplished in a properly equipped recording studio. To achieve a properly prepared tape, one must observe the cardinal rules of CD pre-mastering listed in figure 7.2.

While the pertinent facts are contained in the cardinal rules, some points in the pre-mastering process merit further discussion. No matter what the original medium, the de facto standard equipment for CD mastering is the Sony PCM-1630 digital audio processor format, using a 44.1 kHz sampling rate and a  $\frac{3}{4}$ -inch U-matic video recorder. Other (incompatible) digital audio processors are sometimes used. High-quality tape is crucial;

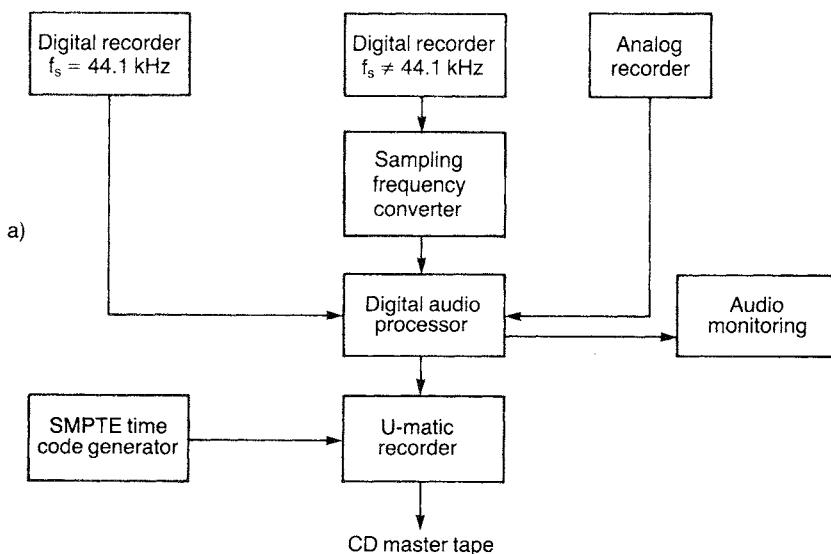


Figure 7.1    The tape mastering process consists of tape transcription (a), track assembly editing (b), and PQ code generation and editing (c).

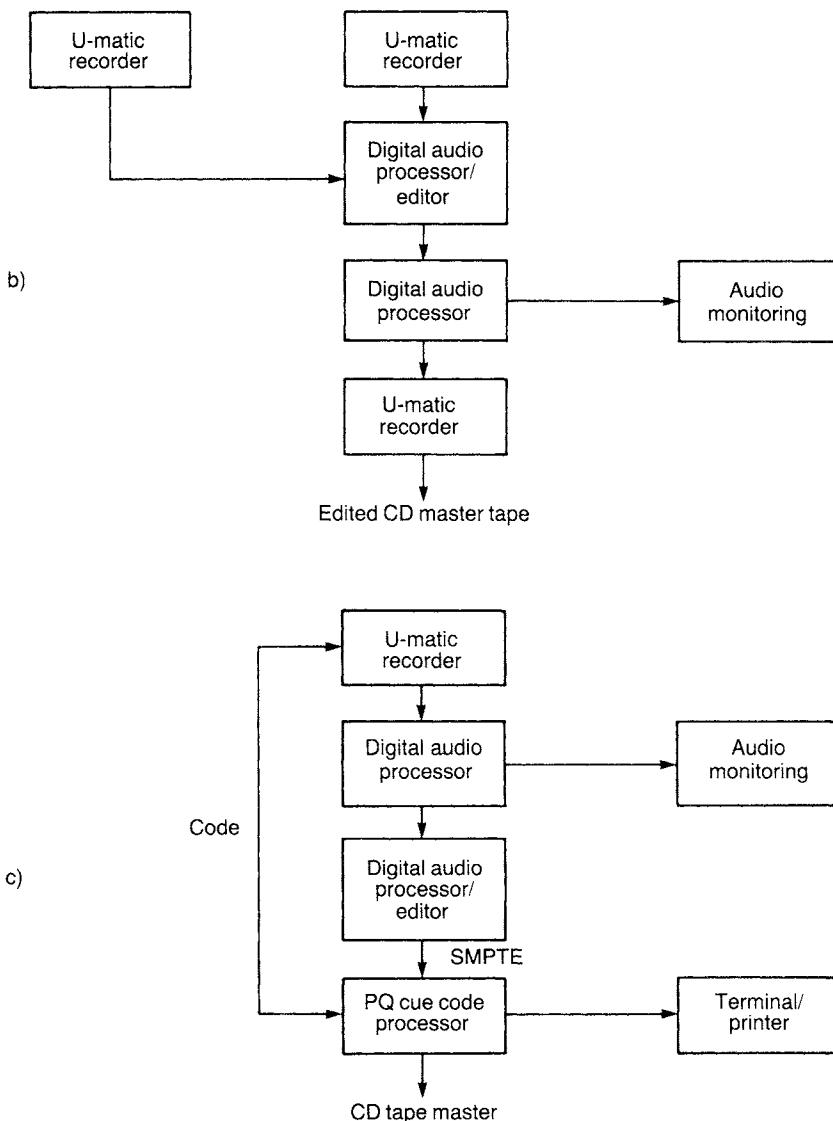


Figure 7.1 *Continued.*

although 80-minute tapes are available, programs longer than 60 minutes are sometimes supplied on two 60-minute tapes and reassembled during tape mastering. Any recording must be auditioned carefully, with special attention to phantom dropouts caused by loose particles moving around

1. Sampling rate must be 44.1 kHz.
2. Tape format must be  $\frac{3}{4}$ -inch U-matic NTSC standard.
3. Timecode must be SMPTE 60 Hz Non-Drop Frame timecode.
4. Timecode must be synchronized to the NTSC video signal.
5. Timecode should run continuously and uninterrupted. Code should not cross over 00 hours, 00 minutes, 00 seconds, 00 frames.
6. Timecode should be recorded on analog track #2 of the U-matic tape.
7. Record digital mute with timecode for a minimum of 1 minute before the first track and  $1\frac{1}{2}$  minutes after the last track.
8. A minimum of 2 seconds must be allowed for emphasis changes between tracks.
9. Prepare a list that is frame accurate to locate beginning and ending points.
10. Prepare a guideline of spurious noise information which indicates any unusual noises or general notes regarding the quality of the recording.

---

Figure 7.2     *Rules for CD master tape preparation.*

the tape. The U-matic recorder must be checked for dirty heads or tape path, worn heads, and mechanical alignment. The U-matic should be NTSC format with the dropout compensator (DOC) switched out of circuit. In the case of V0 type U-matics, a modification should shift the crossover position of the head-switching servo (this is indicated by a "DA" on the model number). The dropout mute sensor on the 1630 must be set correctly; this is variable by the use of DIP switches inside the processor—many engineers exercise the video tape by fast-forwarding it to the end and then rewinding it; this removes any packing irregularities.

### *Tape Preparation*

Whatever the master tape's format, the master CD will be cut from a U-matic tape through a PCM processor. To prepare a tape master, the timecode should first be recorded on channel 2 of the analog track at 0 VU,  $\pm 3$  dB on the U-matic meter, nondrop frame (i.e., full frame). It should run across the entire tape length, or at least from the beginning, to 1 minute after the end of the audio program. This provides a SMPTE code and video sync pulse; the digital audio program will be recorded as an insert-type edit, leaving the video control track and timecode track intact. Timecoding can be done using a 1630 set in dubbing mode, which produces a digital mute picture for the video transport. The processor automatically sets itself for 60 Hz, non-drop frame code. Major axis crossings such as 23 hours, 59 minutes, 59 seconds, 29 frames, should be avoided; they might disrupt the encoder. When using a digital audio editor, track assembly may be accomplished without first striping the tape with timecode.

The original master tape is copied to the timecoded CD tape master. The CD master tape is assembled by performing the correct edits and inserting pauses between tracks. When dubbing from noisy analog tapes, fast fade-ins and fade-outs are preferable to tight leadering in order to avoid abrupt transitions in noise level. Of course, the analog tape machine must be aligned to the tape itself.

At least 30 seconds of digital zero (PCM data with no modulation) must be recorded at the beginning and end of the tape. This is a lead-in and lead-out period. The lead-in avoids potentially defective tape at the beginning of the cassette (cautious clients use 2 minutes or more of digital zero here). This is easily achieved with the 1630 input set for dubbing and the playback processor set for pause. When the audio signal is transferred/edited through the 1630, watch the levels carefully, adjusting for a decibel or two of headroom. (A peak level of + 18 to + 20 on the 1630's bar-graphs will result in a fully modulated tape.) When digitally fading any program material, care must be taken to avoid granulation noise during fade-outs. If the editor does not use dither, fader positions of more than 60 dB below system headroom can produce granulation noise and should be faded through as quickly as possible. Beware of any lapses in audio signal or clicks between tracks. Leave up to two minutes of lead-out at the end of the program. No alignment tones are required on the CD master tape. And of course, LP side breaks are no longer needed.

The audio program can be recorded either with or without pre-emphasis, or it may change between tracks, provided the begin-access point of the track is placed at least 2 seconds after the end-access point of the previous track (see below). Pre-emphasis boosts high frequencies during encoding and attenuates them upon decoding (at the player); this results in a slightly better S/N ratio. However, the boost also necessitates slightly lower levels on the master tape and the resulting CDs.

### ***Subcode Editing***

After the master tape has been assembled, the music program is fully recorded on the U-matic tape. The next step in the pre-mastering process turns the CD master tape into a CD tape master, ready for mastering. Every CD contains a subcode (which was discussed in Chapter Three). Cue editors use user-input information and timecode numbers to generate the subcode. Using the CD subcode processor/editor, the engineer enters the preliminary required data that is requested by the system's menu: album title, artist, record label, catalog number, UPC number, analog or digital source, and mastering engineer's name. All track titles are entered, along with other individual track information, including the ISRC number, copy inhibit, pre-emphasis, and index points.

### *Access Times*

The timecode points at the beginnings and ends of individual tracks must be entered. The concept of using timecode to select the proper begin-access and end-access points is critical in preparing a CD tape master as well as its cue sheet. The start of the first track is located using the proper timecode numbers. Then the editor is used to rewind the tape 15 video frames, and this new location is entered into the subcode editor. The added time is required to allow a player sufficient time to unmute and begin playing after it has located a track's beginning. It also provides a comfortable margin for the listener from the time the play button is pressed to the beginning of playback, and it protects against offset between the audio and video heads on the U-matic. Of course, any ambient noise before or after a track must be included in the track's time markings.

Next, the end of the track is located. The end-access point is the point where the access signal to the player is encoded. It signals the player when to go to the next track when the player is in random mode playback. It is not equivalent to the actual end point. The end-access point must be placed after the end of the program (including its decay ambience) or else the player might prematurely skip ahead to the next track when it is in random access mode. Some facilities suggest a full second between actual end point and end-access point.

This procedure is graphically represented in figure 7.3. In the first example (figure 7.3a), the tracks (TNO) are separated by intervals of 2 seconds or more. If access-begin points are placed a full second before actual begin points, and access-end points placed a full second after actual end points, then a comfortable margin is assured. In the second example (figure 7.3b), the tracks are merged with crossfades. In this case, the first begin-access point and the last end-access point get their 1-second cushion. However, the begin-access points are placed 15 video frames (30 frames per second) before the crossfade begins, and access-end points may be omitted after tracks 1 and 2. In the third example (figure 7.3c), the intervals separating the tracks are less than 2 seconds. The inter-track begin-access points must be placed 15 frames prior to the actual begin points. Note that in this example the begin-access point for TNO 2 is placed while TNO 1 is still playing.

A few specific caveats are worth pointing out. A track time length must be at least 4 seconds, and individual duration should be at the time between the begin-access and the end-access points. When a track is recorded with emphasis (but the prior track is not), the begin-access of the track should be placed at least 2 seconds after the end-access point of the prior track. When the interval between two tracks is less than 2 seconds, omit the end-access point from the former track and place the begin-

access of the latter track 15 frames before the actual begin point. As in a crossfade, when the track blank is less than 2 seconds and the end-access point is omitted, the individual time should be between the begin-access points of the two tracks.

Disagreement appears to exist between one pressing facility and another regarding the number of frames between the begin-access point, the end-access point, the actual begin point, and the actual end point. The result is an offset in which some players access a track after the start of a program when they are in random access mode.

Some pressing facilities suggest that the begin-access point be placed 1 second before the actual starting point to provide the listener with a comfortable margin. This also aids slow players. However, this risks the possibility of an accurate player starting a track at a point where the previous track is still playing. In the case of a crossfade between two tracks, use 15 frames before the beginning of the next track (not 15 frames after the previous track's end). Other facilities call for only 5 frames between the begin-access point and actual begin point. In either case, the first track calls for a longer start delay than other tracks; 2 seconds or more give the listener a chance to get settled.

The CD standard calls for players to mute/unmute the audio signal within 14 milliseconds during track search operations and gives them 2 seconds to unmute at the first audio track. The standard also permits an inaccuracy of 150 milliseconds in locating a random start location. In practice, in any case, cheaper players can exceed tolerances. A difference of 5 frames generously covers these player tolerances. However, not all players are built to tolerance. When preparing your own CD tape master, the choice is yours. A consultative phone call to the intended pressing plant could prevent confusion.

When supplying a master tape to a CD pressing facility, and when in doubt, specify all timecodes without offset and clearly state this in your documentation. In any event, all timecodes must of course refer to the actual time of the program on the production tape master. Any error would result in errors in the disc's table of contents.

### ***Final Steps***

When all of the track access points have been entered, the subcode editor calculates the playing time for each track and the total disc playing time, including pauses in minutes, seconds, and frames. Finally, the master tape is rewound, and the subcode editor writes a 30-second header tone which is followed by the subcode information recorded on audio track 1 of the U-matic. The subcode section may extend past the start of the program; this is permissible. Subcode information is recorded at the beginning of

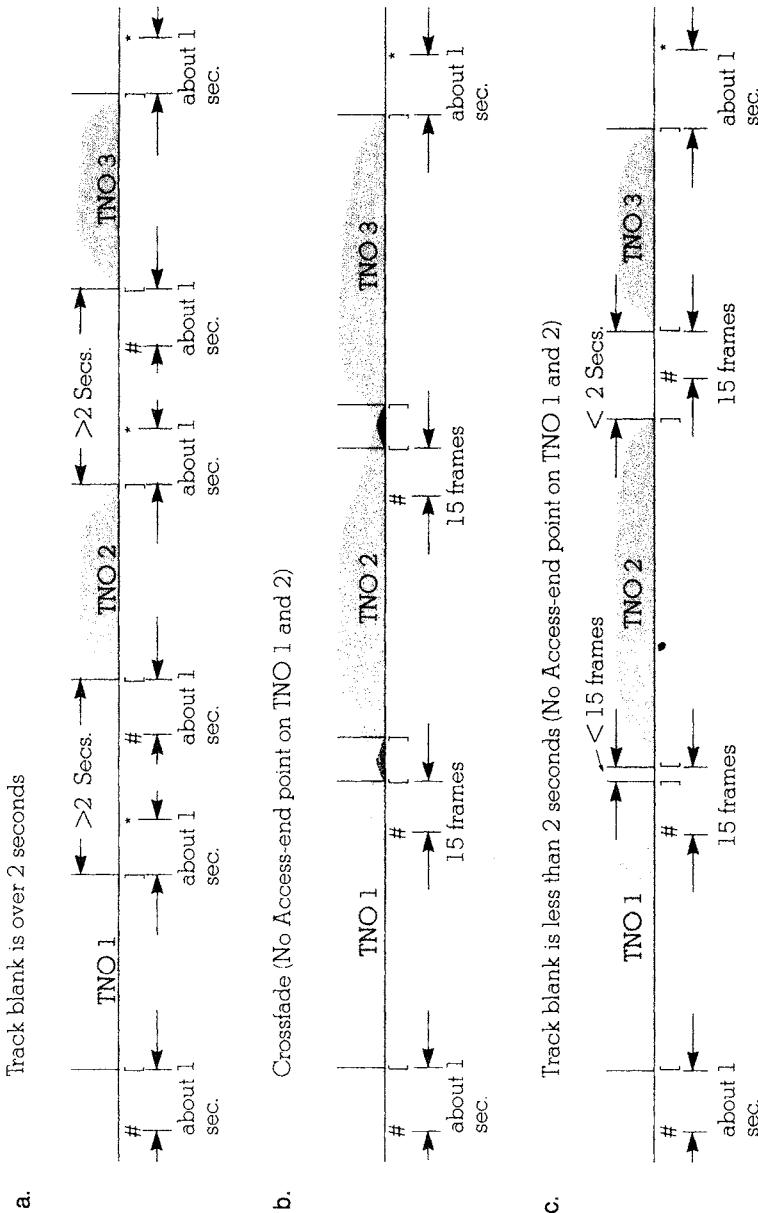


Figure 7.3 Three examples showing placement of access points relative to actual program times: (a) two tracks separated by 2 or more seconds; (b) tracks merged with a crossfade (no access-end points on tracks 1 and 2); (c) tracks separated by an interval less than 2 seconds (no access-end points on tracks 1 and 2). Key: # = access-begin point; \* = access-end point; [ = actual begin point; ] = actual end point. (Courtesy Sony)

the tape in a table of contents (TOC) format. This code is different from the disc subcode channels which are generated with the SMPTE timecode when the disc is cut.

A cue sheet contains a printout of the subcode from the editor. Copies should accompany the CD tape master to the pressing facility, a copy should accompany a safety copy, and the graphic designer could use a copy to properly list the titles and timings on the artwork. Naturally, the cue sheet should be checked carefully against the master tape for accuracy.

A spurious noise information sheet should accompany any tape sent to a pressing facility. The client should document the TNO and describe any noises on the tape (e.g., "TNO 3, click at 25 mins, 34 secs, 12 frames"). This lets the mastering engineer know that the tape has a problem that is already identified and that it has been decided to let it slide. Otherwise, the mastering engineer might contact the client to ask about the click, which would cause delay and further expense.

A summary of information to be supplied with the tape master is shown in figure 7.4.

Title, catalog number, name of customer, etc.

Lettering for center of disc (if different from catalog number)

Format of tape (PCM-1610/30, DASH, PD, DAT, analog, noise reduction, etc.)

Sampling rate

Whether timecode or cue format present (PCM-1610/30 tapes only)

Barcode (optional)

Total playing time, including pauses (optional)

Starting point for all timings (tapes without timecode only)

For each track:

Track number

Start time

Stop time

Times of index points (optional)

ISRC (optional)

Whether emphasized

Length of track (optional)

Title of track (optional)

Spurious noise information

---

Figure 7.4     *Summary of information that should accompany a CD tape master.*

Finally, after the tape master has been auditioned for dropouts, phantom or otherwise, and the cue sheet has been checked for accuracy, they may be forwarded to the CD pressing facility, along with the documentation. The only task left is the preparation of artwork; the pressing plant should be consulted regarding details of the disc label, the inlay card booklet, and restrictions on the use of the official CD logo.

## **DISC MASTERING**

The design of CD disc mastering systems presented a unique challenge to development engineers. Although much of the manufacturing chain could be assembled from existing equipment, such as injection molding, metallization, and label printing machines, the first link in the chain required a wholly new system. The precision required of the system is considerable; any defect introduced on the master disc will be replicated on all production discs.

CD mastering begins with a U-matic tape master and a glass disc. The PCM audio data contained on a tape master will be transferred to the glass master, where it will be represented as pits. Three mastering methods, using photoresist, nonphotoresist, and direct metal mastering, have been developed. Photoresist mastering is most commonly used. All CDs are ultimately derived from the resulting master disc. Figure 7.5 illustrates the manufacturing chain; in this case, photoresist mastering is shown.

### ***Glass Master Disc***

Disc mastering begins with a glass plate, about 240 millimeters in diameter and 5.9 millimeters thick. It is washed, lapped, and polished. An adhesive (chrome film or a silane coupling agent) is applied, followed by a coat of photoresist that is applied by a spinning developer machine. The plate is tested for optical dropouts with a laser; any burst dropouts in reflected intensity are cause for rejection of the plate. The plate is cured in an oven and stored with a shelf life of several weeks; it is ready for master cutting.

### ***Photoresist Mastering System***

In most cases, disc mastering is accomplished with a largely automated laser "cutting" machine which exposes the AZ photoresist on the master glass disc, followed by wet development. The encoded signal must meet the CD encoding standard, as discussed in Chapter Three. Figure 7.6 is a block diagram of a complete photoresist mastering system. The equipment is quite complex. The mastering machine is composed of two units:

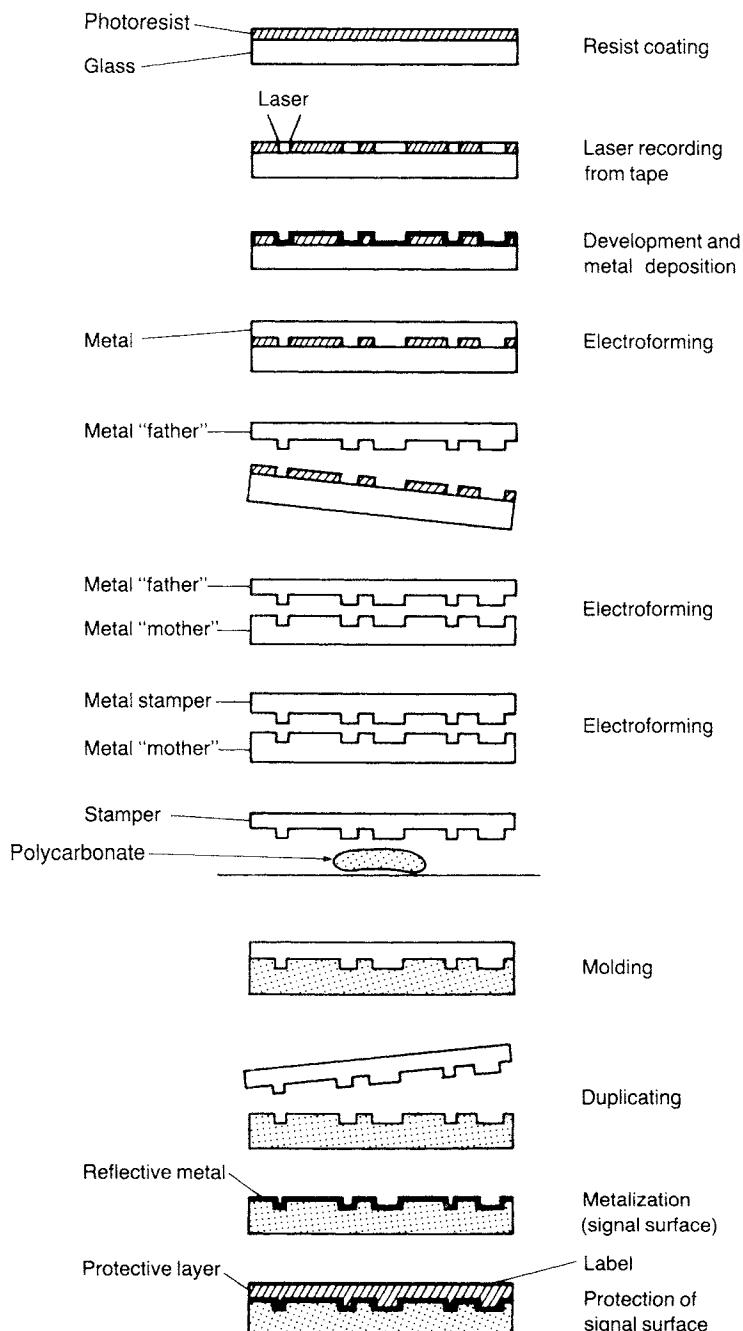


Figure 7.5 Summary of CD manufacturing chain.

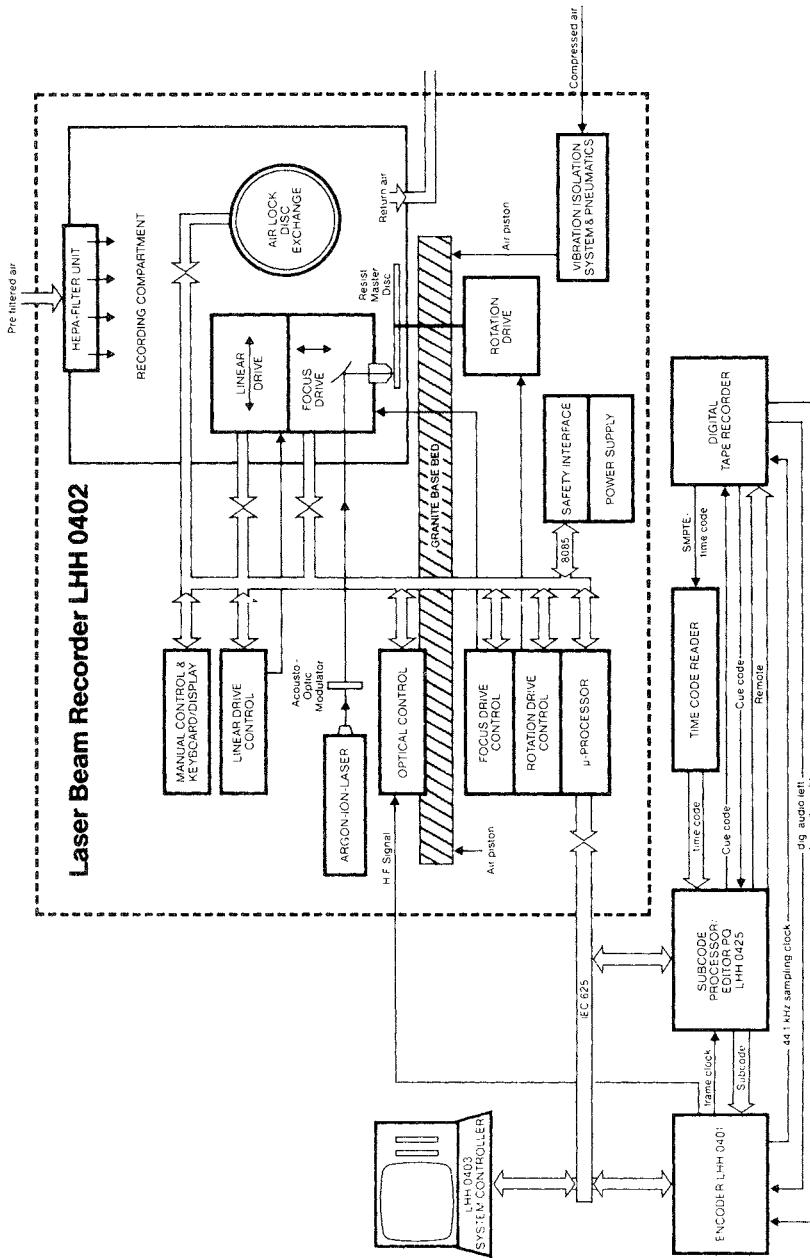


Figure 7.6 Block diagram of a laser beam recorder. (Courtesy Philips)

a control rack and a lathe. The control rack contains a minicomputer with video terminal and floppy disk drive, U-matic video transport, PCM audio processor, and diagnostic equipment. The master tape is loaded in the video transport, and the CD encoder uses subcode, timecode, and digital audio from the master tape to carry out multiplexing and CIRC encoding. The CD encoder generates the EFM channel bit stream signal. That signal modulates the cutting lathe's laser beam. In addition the CD encoder outputs signals used for automatic quality control.

The system controller in the control rack provides for automatic system operation. Recording operations are stored on floppy disk. Recording parameters, such as linear velocity, master identification number, and program length, are entered via controller keyboard. Track pitch, focus offset, and recording intensity can be altered from program defaults if necessary. A video display shows process status.

The master glass plate coated with photoresist is placed on the lathe and exposed with a "cutting" laser to form the spiral track, creating the disc contents in real time as the master tape is played through the PCM processor. The channel bit stream is input to an acousto-optical modulator that is used to intensity-modulate a laser which creates the cutting signal corresponding to the data on the original audio master tape. Audio data on the tape master is transferred, sample by sample, to the disc.

A second laser, which does not affect the photoresist, is used for focus and tracking. A focus monitor is used to confirm accuracy and recording spot quality. The spiral data track extending outward across the disc requires precise motion for both disc rotation and the linear speed of the sled carrying the cutting and focusing optics. To obtain frictionless motion, air bearings are used for both mechanisms.

Although the optics are similar to those found inside consumer CD players, the mechanisms are built on a grander scale, especially in terms of isolation from vibration. The entire cutting process is accomplished automatically under computer control. Not only are the contents of the CD exposed, but test signals are recorded on the inner and outer diameters of the master disc data area.

### ***Developing***

After exposure in the master cutter, the glass master is developed by an automatic developing machine. The exposed areas are etched away by the developing fluid, creating pits in the resist surface. During development a laser monitors pit depth and stops development when proper engraving depth has been reached. The optimum data signal would result when a pit caused an absence of reflected light from the disc, thus distinguishing it from the surrounding reflective land. This is theoretically achieved when the pit depth is one-quarter of the apparent wavelength of the laser pickup,

thus providing destructive interference, and the pit width is such that the intensity of the light reflected from the pit bottom equals the intensity of the light reflected from the surface. In practice, pit depth and width specifications must be modified slightly to account for pit geometry and to provide a more robust tracking signal. This was discussed in detail in Chapter Three.

### ***Final Preparation***

Following development, a metal coating, usually of silver, is evaporated onto the photoresist layer; the master disc is then ready for electroforming and replication. An important quality control check is performed at this point to ascertain the accuracy of the disc formation and pit geometry. The master disc is played on a master player, and test signals are derived to measure the high-frequency signal output. Track pitch and track stability are measured by monitoring the radial tracking signal during playback. In addition, errors are counted and subcode accuracy is verified. Finally, the master disc is auditioned for audio program quality.

### ***Facility Requirements***

Even though mastering sounds fairly straightforward, most CD mastering facilities will attest to the difficulties involved. A laser beam mastering facility requires numerous subsystems, including a resist master preparation system, master recording system, developer system, master disc player system, and disc master electroplating equipment. Other items from a long list include microscopes, ovens, chemical preparation equipment, diagnostic and test equipment, glassware, cleaning and protective materials, desks and hoods, audio monitoring equipment, and dust-free paper.

Site requirements are carefully specified. Clean air is critical. CD pits are among the smallest of all manufactured formations—about the size of a smoke particle. Thus the entire mastering process must be carried out in a clean room environment, with the size and number of particles in the air strictly regulated. Clean room classifications specify the number of particles present in the air and their size. Class 100,000 for example, specifies no more than 100,000 particles 0.5 to 5 micrometers in diameter per cubic foot of air per minute. The air inside a CD mastering lathe is specified to be Class 100.

Temperature and humidity, along with ambient air pollution levels, must also be specified. With some systems the glass disc moves from one process stage to the next in a sealed cartridge. With each process step, the disc is automatically removed from the cartridge and then returned. This minimizes manual contact and air exposure, thus reducing the chances for contamination of the disc.

Vibration would be disastrous to the cutting process. The laser beam recorder is mounted on a massive baseplate, made from materials such as cast iron or granite, with a pneumatic vibration isolation system. Other requirements include a clean electrical system, demineralized and hot water, compressed air, filtered air, nitrogen, and a system ensuring the exhaust of contaminated air.

The facility could be housed in a clean room of 200 square meters. However, many disc manufacturers specify larger rooms, with an eye toward additional mastering equipment to accommodate the production of audio discs as well as CD-ROM and future formats encompassing both constant linear velocity as well as constant angular velocity discs.

### ***DRAW Mastering System***

A DRAW compact master disc recorder system uses a process quite different from conventional AZ photoresist mastering equipment. A non-photoresist (NPR) recording medium is used. It employs a direct-read-after-write (DRAW) technology, which eliminates the developing processing required by photoresist methods after laser cutting. With a DRAW system, pits are formed on a plastic coated glass master disc; a 50 mW argon ion writing laser operating at a wavelength of 488 nanometers vaporizes points on the plastic layer to form the pits. A lower powered 5 mW helium neon reading laser follows the writing laser to control the focus of the writing laser and check the integrity of the cut immediately after it is made, reading back the cut signal. With modifications, cutting time can be decreased two or three times less than real time. The cutting lathe is enclosed in its own Class 100 air chamber with positive pressure, thus minimizing exterior room specifications. Because the lathe is designed to read the master disc while it is being written, exposure power is servo-controlled, based on playback of the master disc in real time. This eliminates the need for a separate master disc player for the quality control operation following master disc recording.

The DRAW mastering process, originally developed for video disc mastering, offers several advantages over conventional AZ photoresist technology; there are fewer steps of critical nature. Preparation of the glass master is simpler; no adhesive layers are required, and baking does not affect the sensitivity of NPR as it does AZ photoresist. Because NPR is self-developing, a separate developing step is not needed; moreover, disc exposure evaluation can take place in real time. The pit shape of an NPR master surface is uniform and not subject to variation due to exposure, baking, or development. All quality control can be carried out during data writing. One can expect improvement in yield, efficiency, and productivity.

### ***Direct Metal Mastering System***

Alternatively, a direct metal mastering (DMM) process has been developed for producing compact disc masters. Instead of using a laser to expose a photoresist disc, DMM embosses the CD pits directly on a metal master disc. Because the embossing process is largely impervious to contaminants on the disc surface, a clean room is not required.

The process employs a piezoelectric electromechanical transducer, such as the type used in videodisc recording. Figure 7.7 shows a piezoelectric longitudinal transducer with an embossing diamond stylus attached. Deformation of the element occurs through the piezoelectric effect, according to the polarity of the bipolar applied potential. The assembly rests on the disc, pressing on the recording surface because of gravity. Because the center of gravity of the assembly cannot change instantaneously on account of inertia, the longer dimension causes the cutting stylus to press into the recording surface while the shorter dimension causes it to move away, lifting it from the disc.

The stylus thus presses into the disc during positive pulses and skips above it during negative pulses. The difference in length in the piezoelectric element and the maximum possible time between positive pulses can be calculated. This, in turn, is correlated to the lowest data frequency of the CD signal (196 kHz). The stylus is thus designed to successfully jump over (and onto) the disc surface within the constraints of the CD's EFM signal. The result is the impression of pits on the disc surface.

The shape of the diamond stylus is crucial, for it must create a pit geometry compatible with the CD system. An embossed pit has a geometry similar to that of an analog groove; a cross section view would reveal a "V" shape with walls at a 45 degree angle. By selecting a suitable stylus angle and optimizing the speed with which the cutting stylus is moving, engineers can create pits which modulate the pickup's laser beam with the same intensity and a similar phase characteristic as the pits from conven-

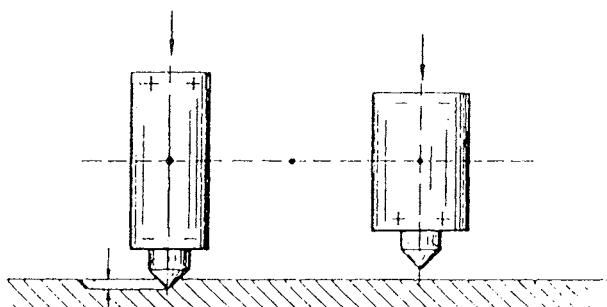


Figure 7.7    A piezoelectric embossing element is used in the DMM CD mastering process. (Courtesy Teldec)

tionally mastered CDs. An embossed pit also differs from a photoresist pit because embossing results in a displacement of material; because of material flow, a ridge is created along the lateral pit edges. These ridges produce a hollow area in the negative area of the stamper. This encourages polycarbonate flow and, consequently, a sharper pit edge.

In direct metal mastering, a metal surface is used for mastering. A glass substrate disc is prepared with a separation layer a few nanometers thick. A layer of copper approximately 300 nanometers thick is sputtered over the preparation layer, and the data is embossed on this copper surface.

When the embossing is complete, a galvanic process is used to prepare the master for replication. The copper surface is covered with a layer of precious metal (gold or rhodium) a few micrometers thick. This layer is then covered with a layer of nickel approximately 0.25 millimeters thick. Using the resulting metal parts, engineers use conventional galvanic processes to generate mother and stamper tools. If rhodium is used, the metal part may be used directly as a stamper, able to press 10,000 to 15,000 discs before deterioration.

## ELECTROFORMING

Electroforming, sometimes called matrixing, bears a close resemblance to the corresponding stages of vinyl record production. The various aspects of electroforming, such as composition of the plating solution, temperature control, solution flow, and plating current distribution, are optimized to obtain a fine surface for CD replication. These adjustments are completely different from those needed for an analog disc. The electroforming plating process ultimately results in metal stampers used to replicate CDs.

Following mastering and developing, a silver (or nickel) layer is deposited with vacuum evaporation over the photoresist layer. The master is played on a master player system to assess aural and measured quality; it also serves as a reference to evaluate the quality of the final production discs. The silvered master is then transferred to the electroforming room. The silvered master disc, which is now electrically conductive, is placed in a galvanic nickel electrolyte bath. The master disc is the cathode (-) of an electric circuit. From a nickel anode (+) a nickel layer is electroformed onto the master disc.

In a typical application, a reservoir holding the electrolyte solution (nickel sulphamate) is located outside the clean room. It is held at a constant temperature, filtered for 1 or 2 micron maximum particle size, electrolytically purified, pH checked and held at a constant value by adding sulphamic acid, and circulated constantly. The master is placed in a sink and etched (activated) with a solution of sulphamic acid. It is rinsed and put

in a holder to ensure uniform current flow over the master's face. The holder is placed on the shaft of the cathode drive in the plating tank, and the plating process is initiated.

The plating process starts with a low current which increases as the nickel plate grows thicker. The time required for plating depends on the thickness desired. Nominally, a maximum current of 100 amps produces a part in less than 2 hours.

After electroforming, the nickel part is separated from the glass master. It is rinsed in an electrolytic degreasing tank at high temperature, or with a solvent such as acetone, to remove any particles of photoresist. Since the master's photoresist layer is usually damaged when it is separated from the metal part, masters may be used only once.

Because the disc master has a positive impression of the CD pit track, the resulting nickel copy, called the "father," is a negative impression. In cases of limited production, the father can be used to replicate CDs. More generally, the father is used to galvanically generate a number of positive impression "mothers." Four or five mothers may be made from one father.

Each mother is inspected. If a mother is acceptable, it can then generate a number of negative impression nickel mold matrices, "sons" or "stampers," by the same process. Stampers are optically checked to ensure quality.

When it is separated from the mother, the stamper is rinsed and dried, and a protective layer (either a tape, form, or plastisol) is applied to its face. Next, a backsanding machine is used to polish its back, and it is put in a centering device. Using a reference mark on the stamper, the part is centered to within 1 micron, and the center hole is punched out. Next, the outside circumference is punched. The stamper is ready for mounting in an injection molding machine. Four or five stampers may be made from one mother. When mounted to a die, stampers are used in the replicating machines to produce CD discs.

The electroplating room requires a Class 1,000 environment. However, the electroplating process is carried out in enclosed electroforming consoles. They are placed in Class 100 laminar flow enclosures to maintain cleanliness. Often a HEPA (High Efficiency Particulate Air) filter and fan are used to circulate filtered air. Submicronic filtration systems are used to ensure that the chemical baths do not become contaminated. Moreover, the electroplating system must be able to produce disc molds that are flat to within  $\pm 3$  microns over the entire disc surface.

In new generations of electroforming equipment, the emphasis is on speed. In some cases the father, mother, and stamper processing can be completed in less than an hour. This is crucial when large disc runs require a large volume of stampers.

## INJECTION MOLDING

Following the mastering and electroforming processes, the disc is ready for replication. Injection molding techniques are commonly used. Molten plastic is injected into a mold cavity, with the stamper on one face producing a clear plastic disc with the pits impressed on one side. A polycarbonate plastic is used chiefly because of its high transparency, dimensional stability, accurate reproduction of the mold surface, minimum water absorption, good impact resistance, easy processing characteristics, and freedom from impurities. These characteristics can be achieved from a high-grade polycarbonate resin. However, polycarbonate material has certain inferior specifications, especially when handled by injection molding; we shall discuss these later in this section.

### *Molding Considerations*

The molding of compact discs presents great challenges; the disc must be flat and optically pure, and it must retain an accurate impression of the data pits. Furthermore, typical molding practices result in discs with deficient optical properties. To achieve satisfactory results, disc molding requires minimized plastic resin viscosity for good fluidity. To obtain low viscosity, good fluidity—and hence acceptable optical properties—the resin temperature must be raised considerably. However, the resin is easily decomposed, resulting in color change or bubbles. Because the disc volume is small and the amount of resin needed is small, heated resin retained in typical molding machines can easily be degraded or burned. Furthermore, high-speed passing of resin causes mechanical shearing heat, which is another factor to be controlled.

Although problematic, high-speed filling is desirable in injection molding. High-speed filling prevents a drop in temperature of the polycarbonate and enables uniform melting temperature in the cavity. This ensures uniform force and density of the molded disc and makes possible uniform cooling speed across the entire disc surface, which in turn allows uniform shrinkage. High-speed filling also prevents pressure drop and assures that adequate pressure is applied even at the far edges of the cavity. Distribution of the pressure in the cavity is uniform, and molding shrinkage is small. Molding can be performed with lower pressure, resulting in discs with minimum warpage or deformation. Finally, high-speed filling permits molding at lower temperature and faster cycle times.

As cavity filling speed is increased, however, two problems arise. First, it is difficult to accurately control the amount of polycarbonate to be injected. Second, it is difficult to discharge air from inside the cavity. Sophisticated control systems and, in some cases, a vacuum in the cavity can

be used to achieve very brief filling periods of 0.01 to 0.1 seconds at low pressure.

Because of these and other problems, use of typical injection molding machines results in discs with burned plastic and, consequently, contaminations or bubbles. If the resin temperature is lowered, strain or deformation of the disc after molding can result, along with a high birefringence. After experimentation with various polycarbonate resins, different kinds of injection molding machine designs, mold shapes, techniques for producing a single-piece polycarbonate disc were achieved. As a result, there is a considerable difference between standard polycarbonate and that used for making CDs in specifications such as melt flow rate.

### ***The Molding Machine***

The heart of an injection molding machine is its plasticizing unit, shown in figure 7.8. Because polycarbonate is a hygroscopic material, it must be dried, stored at a high temperature, and then used without being exposed to ambient air. Thus pellets of polycarbonate (cleaned and dried) are drawn directly through the hopper and into a heating barrel; a screw moves the pellets through a series of heating coils to heat the plastic quickly and uniformly to a high temperature (approximately 350 degrees centigrade) in order to achieve smooth flow properties into the mold cavity during injection. When the molten plastic is injected into the mold cavity at high pressure, it conforms to the stamper's contours, producing a substrate disc with pits. The mold is kept at a temperature of approximately 110 degrees centigrade.

Some systems use an injection-compression molding process. The molten plastic is first injected into the mold, which is not yet completely closed. The final shape of the disc is accomplished during the subsequent compression step. With injection-compression molding, molded-in stress is minimized, but cycle times are somewhat longer than conventional injection molding, perhaps 10 seconds as compared to 5 seconds.

In most cases, the center hole is formed before the disc is removed from the mold. In other systems, the center hole is punched out of the disc separately, after the top layer is applied but before the label is printed. Following molding, the warm disc is subject to static charge; any dust particles in the air are attracted to the disc. The molding room must be kept at Class 1,000, and special hoods are placed over the molding machines for an even cleaner environment at the molding head.

In summary, the injection molding process requires consideration of environmental cleanliness and temperature control, nozzle and hopper temperature regulation, mold temperature regulation, adjustment and stability of injection volume and time, removal of flashes from mold surface, and the quality of stampers.

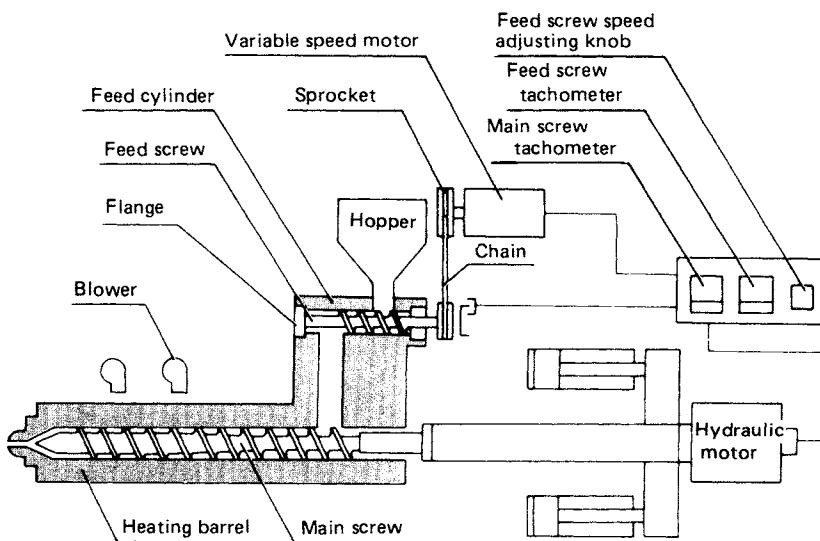


Figure 7.8     *Elements of an injection molding machine. (Courtesy Meiki)*

### Birefringence

A form of optical distortion measured as birefringence can be introduced during injection molding of the polycarbonate substrate. Birefringence is an optical measure of material stress, measuring by the differences in the index of refraction in materials through which light passes. These affect the phase relationship of laser light and can result in reading errors. The amount of birefringence in a plastic material is related to its inherent optical properties but more importantly to the internal stress in the plastic caused, in this case, by shrinkage, flow lines, and inclusions. Injection molding therefore greatly increases the amount of birefringence in polycarbonate.

Due to refraction, light rays bend as they pass from one medium to another with a different density. With birefringence (a kind of double refraction) light is refracted into two images. In refraction, light slows down as it enters a new medium. With birefringence, light is slowed down at two different rates, depending on the angle at which it passes through the material. More specifically, the light is split into two perpendicular polarized components with different refraction and different wavelengths, as shown in figure 7.9. Birefringence causes a reduction in the intensity of laser light that is read from the disc by the pickup. This

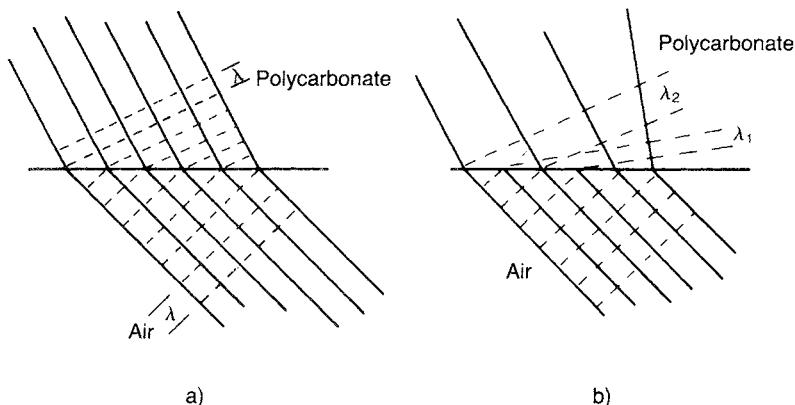


Figure 7.9 *Light rays bend as a result of refraction (a). Birefringence causes double refraction (b).*

reduction in intensity is caused by the turning of the polarization direction away from that of the polarizer. Because the polarization direction is not aligned with the polarizer, less light is allowed back into the pickup. The effect is similar to the way polarized sunglasses reduce glare; however, in the case of CD production, the effect is unwelcome.

Birefringence also changes the critical phasing of the light as it reflects off the disc. The area of the laser spot simultaneously striking a pit relative to the surrounding land (refer to figure 3.5) and the carefully calculated pit depth causes destructive interference. This yields intensity modulation of the reflected light. Since birefringence causes deteriorated interference, the quality of the modulated light is degraded. In either case, birefringence causes reduced signal response from the optical pickup. This can cause data reading errors, as well as tracking problems, in all CD players with laser pickups which use a polarized, beam-splitting head to read polarized light as it passes through the disc. Pickups using holographic optical elements (HEM) technology are not affected by birefringence.

The wavelength of a pickup's laser light is 780 nanometers in air and approximately 500 nanometers in polycarbonate (the index of refraction is 1.55). Birefringence is measured in terms of the phase shift between the  $x$  and  $y$  components of light and is expressed in nanometers. It is measured by finding the amount of reduction of polarized light intensity through the disc. The birefringence specification calls for a maximum deviation of 20 percent (100 nanometers) from the 500 nanometer wavelength. In any event, when birefringence of a disc exceeds a normal amount, there is increased risk that not all players will play the disc.

The first-generation methods of injection molding of polycarbonate were inefficient in several respects. Polycarbonate was difficult to mold without introducing optical distortions, and basic materials and equipment were expensive. Given the difficulties in this replication method, improvements and innovations in the process proceeded quickly. Experience has shown that birefringence is usually worst on outer disc circumferences and can be minimized with high melt temperature, rapid injection into the mold, low mold pressure, and short mold time. Low-viscosity polycarbonate is also helpful. In addition, a high chlorine content must be avoided, as this correlates with higher error rates in aging tests.

## METALLIZATION AND SPIN COATING

After leaving the molding machine, the disc is wholly formed but transparent; a player's laser beam could not read the impressed data because there would be no reflected beam to convey the information. Hence a reflective layer must be placed over the data pits. The reflective layer, typically aluminum, is very thin—on the order of 50 to 100 nanometers thick. To protect this thin layer from physical damage and oxidation, an acrylic plastic layer is applied over it.

### *Metallization Layer*

The reflection coefficient of the metal layer, including the polycarbonate substrate (note that the CD player laser must shine through the substrate to the metal layer), is specified to be between 70 and 90 percent. In addition, the metal must be chosen to be inert with the polycarbonate substrate. Three cost-effective metals qualify with the required reflectivity and stability: aluminum, copper, and silver. Gold could also be used as an expensive alternative. The fact that it is an inert metal may be of use in some applications. Reflectance values of 80 to 90 percent apply at the readout wavelength even with thin layers. Because of the physical appearance of the layer, aluminum and silver are preferred over copper. Metallization requires a clean room of Class 1,000.

### *Vacuum Evaporation*

With vacuum evaporation (vapor deposition) as the application method, aluminum has good adherence and is most commonly used. A spacer covers the center of the disc, and a ring masks the outer edge. Consequently, only the area containing the information is coated. In some systems, the evaporation is accomplished in a vacuum chamber with aluminum-coated

resistance elements. Large racks of discs are treated simultaneously, and a good mirror coating results. Metallization may take about 15 minutes, depending on the size of the vacuum chamber. However, a major production snag is encountered as the discs must be individually loaded onto the racks and batch-metallized. Then the pumping-down of the evaporation chamber must occur, followed by the readmission of air after deposition. These steps are not easily integrated into the otherwise continuous disc manufacturing process. Furthermore, any dust in the chamber, stirred up during pumping, may cause pinholes in the metal layer. This can be avoided only through stringent cleanliness.

Alternatively, some vapor deposition systems employ multiple pre-chambers in which a spindle loaded with discs may be undergoing metallization while another load in a prechamber is being evacuated. Discs are metallized individually, a process which takes about 3 seconds. With a capacity of 1,200 discs per hour and a target life of 40,000 hours, such a metallizer could run 40 hours without interruption. In still other systems, discs are moved one at a time through the metallizer on a conveyer. In this way, throughput is further increased.

### *Sputtering*

Magnetron sputtering is another metallization method used for CDs. A cold solid target is bombarded with ions, releasing metal molecules which coat the disc. Using high voltages, a discharge is formed between a cathode target and an anode. Powerful permanent magnets behind the cathode form a concentrated plasma discharge immediately above the target area. Argon ions are extracted from the plasma. They bombard the target surface, thus sputtering it. The CD is placed opposite the target and outside the plasma region.

Discs can be metallized one at a time by loading them into a fast pump-down vacuum chamber; metallization takes only a few seconds. Because each disc sees exactly the same conditions each time (versus differing disc-to-disc conditions in a batch process), consistency is enhanced. Alternatively, a continuous sputtering machine with a permanent vacuum can be used for mass production. Discs are loaded at one end of a conveyer and passed through a series of bulkheads which create pressure gradients from ambient pressure down to a vacuum and back again to ambient pressure conditions. Sputtering is faster than evaporation and allows a choice of different metals or alloys. A throughput of 2,400 discs per hour is possible.

### *Wet Silvering*

A metallic mirror can also be obtained by electroless silver deposition, often called wet silvering, a process long used to produce glass mirrors.

In this process wet chemicals are combined to cause the formation of a thin layer of silver over the disc surface. Since the process is accomplished by spraying liquids, a continuous production machine can be used; the discs are placed on a conveyer belt and ride through the sequential processing stages. Wet silvering takes less time per disc than vacuum evaporation, and continuous rinsing of the disc minimizes the formation of pinholes.

### *Spin Coating*

The metal layer is covered by a plastic layer with a spin coating machine. This layer protects the reflective layer from scratches and oxidation. The plastic layer can be nitrocellulose-based with air-drying or an acrylic plastic cured in ultraviolet light. The latter is advantageous because of very short curing time, good bonding ability to aluminum, and the fact that it is a harder, more scratch-resistant surface. However, it is more costly and difficult to print on.

At the spin coating machine, a small, metered amount of acrylic is sprayed on the disc just inside the inner diameter of the metal coat. The disc is then spun at high speed so that a very thin coat of lacquer, typically about 6 or 7 microns thick, remains over the disc surface. The layer is then cured.

### *Label Printing*

The label is printed over the top acrylic layer. Two types of printing are in common use: silk screening and pad printing. Silk screening is very good when large surfaces have to be covered with an opaque layer of ink. A silk screening machine is somewhat slower and more expensive. Pad printing may be superior for reproducing intricate details. Both systems can print up to four colors and can be easily changed to new labels.

In the silk screening process, a screen with light-sensitive emulsion is exposed to light in order to transfer the image to the screen. The screen is then used as a stencil to mask printing ink to the disc surface. The ink must dry quickly. In some systems, ultraviolet-curing ink is used.

With pad printing, a coat of printing ink is spread on a cliche with recessed characters. The ink on this cliche is then scraped away with a knife, leaving on the ink in the characters. A cushion or pad is pressed onto the cliche and then transferred to the disc where it is pressed down again. In this way the ink is forced to adhere to the surface of the disc and remains there when the pad is lifted off. The disc is conveyed from one print position to the next for multiple colors.

In general, silk screening is the preferred method of labeling. It proves to be faster, with better yield. For example, a silk screen printer may

process 2,500 discs per hour, as opposed to the pad printer's rate of 700 an hour. Printing may be carried out in a dirty room of Class 100,000. After printing, discs are ready for final quality control checks.

## QUALITY CONTROL

To ensure that all CDs successfully play on all CD players, a large range of optical, mechanical, and electrical criteria have been established for the CD system. The co-inventors and license holders of the compact disc, Philips and Sony, have published these specifications in the *Red Book*, the reference on standards available to all CD licensees. Some disc tolerances, such as disc eccentricity, have been established empirically; they represent a compromise between practicalities of manufacturing discs on one hand and those of manufacturing players on the other. Other specifications, such as error flags, are strictly theoretically determined. However, the implementation of this aspect, the number of errors permitted on any particular disc, again represents a compromise between media and players.

In any case, disc manufacturers and player manufacturers must agree on the facts to ensure that all disc tolerances are within the design limitations of all players. Moreover, quality control procedures are invaluable as an in-house means of diagnosing problems in the manufacturing process. Figure 7.10 shows the primary disc specifications and their relation to disc playback. Some of the primary specifications are examined below.

### *Mechanical and Optical Specifications*

Molded discs are checked for correct dimensions, lack of flash and burrs, birefringence, reflectivity, flatness (skew angle), and general appearance. The pit surface is checked for correct pit depth, correct pit volume, or pit form and dimensions. The metallized coating is checked for pinholes and uneven thickness. As many manufacturers have discovered, one of the easiest ways to detect dust in the clean room is the appearance of pinholes. Birefringence can be checked with a circularly polarized light used to convert the phase change to an intensity variation measured with a photodiode. The disc can be scanned, creating a map of birefringence versus radius.

Angle deviation measures the angle formed by the normal to the disc in the radial direction as shown in figure 7.11. This angle is critical because any deviation causes the reflected laser beam to deviate from its return path through the objective lens. This angle deviation could result from an improper manufacturing method; specifications call for a maximum angle of  $\pm 0.6$  degrees. Because warpage could be introduced later in the

DISC SPECIFICATION		PLAYER SYSTEM
Characteristics	Requirement	
MAIN PARAMETERS		
Outer diameter of disc	120 mm	
Center hole	15 mm	Disc clamping & mechanism
Eccentricity of track radius	max 70 $\mu\text{m}$	
Thickness of disc	1.2 mm	
Disc weight	14 ~ 33 g	
Angular deviation (skew)	max 0.6°	HF signal
Clamping area	$\varnothing$ 26 ~ $\varnothing$ 33 mm	
Recording area	$\varnothing$ 46 ~ $\varnothing$ 117 mm	
Track pitch	1.6 $\mu\text{m}$	
Reflection and substrate transmission	70 ~ 90 %	Focus servo
Birefringence	max 100 nm	Tracking servo
Scanning velocity	1.2 ~ 1.4 m/sec	Spindle servo
SIGNAL PERFORMANCE		
Asymmetry		
Crosstalk		
Push pull error		
Error rate	$3 \times 10^{-2}$ (block)	
Local defect		Access & control
SIGNAL FORMAT		
Frame format		
Modulation	EFM	
Error correction system	CIRC	Digital signal processing
Audio emphasis		
Control and display system		

Figure 7.10     Summary of primary disc tolerances and their relation to disc playback.

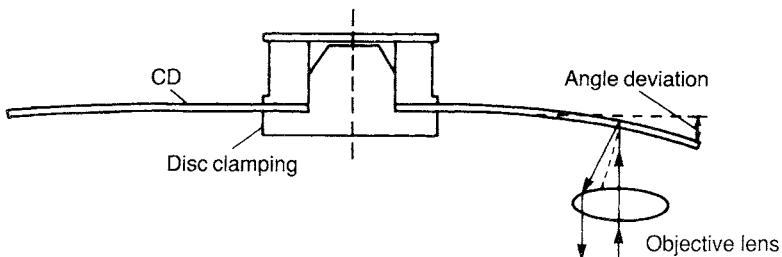


Figure 7.11     *Angle deviation causes reflected light to deviate from the objective lens.*

field (for example, the extreme temperatures found in a car interior might cause bending), the specification allows a safety factor of 0.3 degrees.

Disc eccentricity measures the deviation from circularity of the pit track and the positioning of the center hole. The electroforming and molding processes introduce some eccentricity in the shape of the pit track. In addition, the player's positioning of the disc in the drive might introduce eccentricity. If it is excessive, it could exceed the ability of the radial tracking servo of the player. Tolerances for deviation from circularity call for maximum eccentricity of 140 micrometers.

Disc eccentricity must also account for alignment of the center hole. Specifications call for a center hole tolerance of 0.1 millimeters. A hole that is off center would lead to disc imbalance and noise and resonance errors. In practice, an eccentricity of 1 millimeter would result in significant imbalance.

A distorted disc surface, such as one characterized by concentric waviness, results in reflection of the light beam at a deflected angle, as illustrated in figure 7.12. This, in turn, taxes the player's radial tracking servo circuit; a measurement of the radial error signal can be used to specify the tolerances for disc flatness. To measure the radial error signal for this specification (errors on the order of 0.1 to 1 millimeter) the error signal is filtered and processed to produce an RMS value. Standards call for a maximum radial tracking voltage of 250 millivolts. This represents the specification for radial errors, describing the maximum value the laser spot may deviate from the track. This maximum value is 0.03 micrometers. This specification also accounts for other factors which affect radial tracking, including eccentricity and impurities such as bubbles, occlusions, or scratches.

#### ***Error Correction Flags***

Error correction flags generated during CD playback from the CIRC circuitry can represent the error rate (from sources such as pit structure,

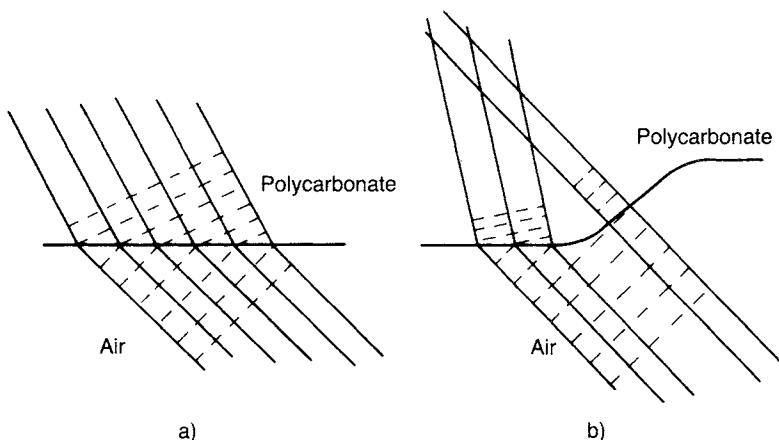


Figure 7.12 A flat disc surface (a) permits normal refraction. A distorted disc surface (b) reflects light at a deflected angle.

birefringence, pinholes, etc.) present on a disc. Thus the glass master, metal parts, and the finished disc may be checked for integrity via error correction flags. As discussed in Chapter Three, data is arranged in frames and corrected through two CIRC decoders, C1 and C2. The C1 decoder corrects minor errors and flags uncorrectable errors. The C2 decoder corrects larger errors, aided by the error flags. Uncorrected errors leaving C2 are flagged as well. Four flags are output from a measuring CD player: C1F1, C1F2, C2F1, and C2F2. They define the error conditions on the disc for every block of data as it is clocked through the decoders, as shown in table 7.1.

C1F1	C1F2	Status
L	L	No errors detected in C1 decoder
H	L	1 error corrected in C1
L	H	2 errors corrected in C1
H	H	More than 2 errors in C1, data passed to C2
C2F1	C2F2	Status
L	L	No errors detected in C2 decoder
H	L	1 error corrected in C2
L	H	2 errors corrected in C2
H	H	More than 2 errors in C2 (uncorrectable)

Table 7.1 Four error flags derived from the error correction strategy are used to diagnose error conditions.

These flags can be counted over a time interval to obtain a measure of disc quality, as shown in figure 7.13. Three signals are derived from the C1 decoder. The first signal, E11, signals one defect symbol in a frame. The second signal, E21, signals two defect symbols in a frame. The third signal, E31, signals more than three defect symbols in a block. To evaluate the magnitude and quality of correction at C1, these signals are added, thus providing the block error rate (BLER). This measures the number of blocks per second containing any C1 errors (one block equals 588 bits/33 symbols). It is thus a measure of both correctable and uncorrectable errors at that decoding stage. A maximum BLER value of 220 pulses per second is permitted at the C1 decoder. High-quality disc manufacturing facilities can achieve BLER values as low as 10 to 20.

The maximum value of 220 pulses per second shows that 0.03 percent of frames contain a defect. This defines the acceptable error limit, since greater frequency might lead to audible faults. The BLER does not provide information on individual defects between 100 to 300 micrometers, since the BLER responds to defects the size of one pit. BLER is often quoted as a 1-second actual value or as a sliding 10-second average across the disc; it is also quoted as the maximum BLER encountered during a single 10-second interval during the test.

There are two error signals at the C2 decoder. The first signal, E22, signals two defect symbols in a frame. Empirical studies of inaudible faults have shown that a count of 100 pulses per second is permitted for E22. A quality control specification may allow eighty pulses per second. The second signal, E32, signals more than two defect symbols in a frame. These errors, which leave C2 uncorrected, are sometimes classified as noise interpolation errors (NI). The player must then interpolate; if there are multiple interpolations, the player may mute, or audible clicking may be heard. Thus, because E32 signals may indicate an audible fault, they are not permitted.

The E21 and E22 signals may be combined to form a burst error count (BST). It measures the number of consecutive block errors exceeding a specified threshold number. For example, if there are fourteen consecutive block errors, and the threshold is set at seven blocks, there will be two BST errors indicated. This count is often tabulated as a total number over an entire disc. A good quality control specification would not allow any burst errors (seven frames) to occur.

### ***Production Checking***

In a production facility, finished discs are inspected for continuous and random defects using both automated and human checking. Birefringence, high-frequency signal, frame error rate, crosstalk, jitter, noise, presence of foreign particles, reflectivity, frame tracking, number of interpolations,

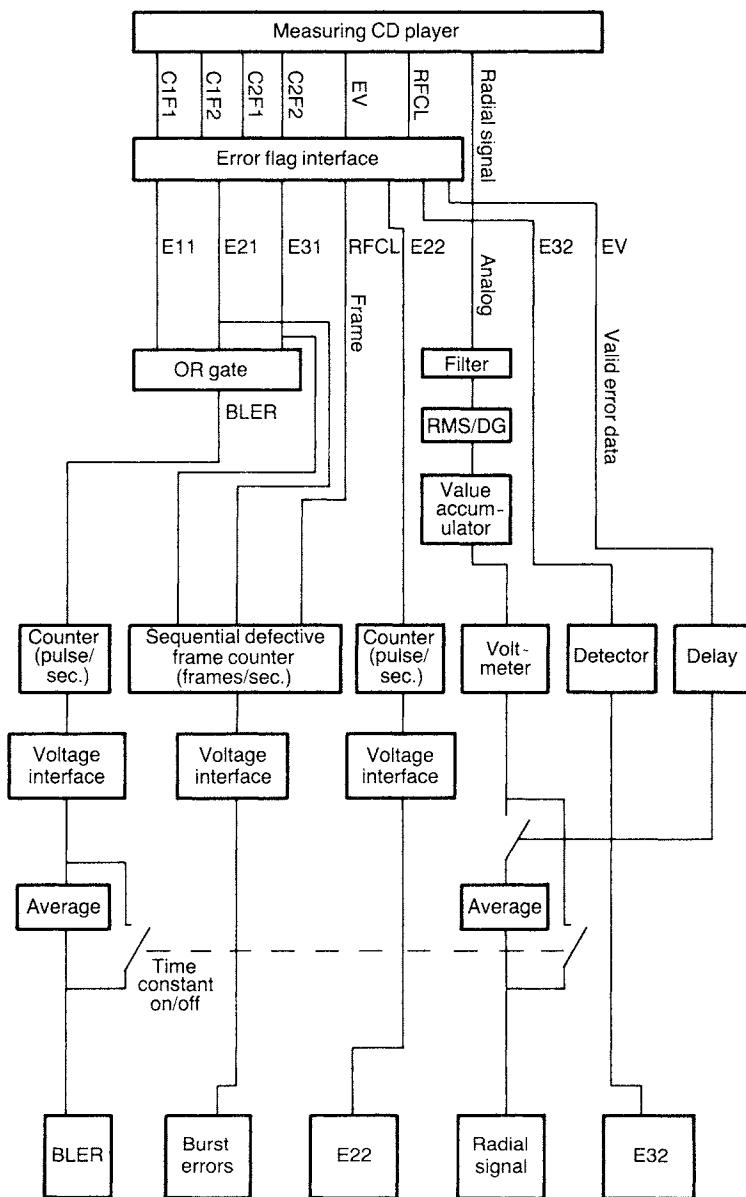


Figure 7.13 Block error rate, burst errors, and other diagnostic signals may be derived from the measuring player.

and skew are checked on selected discs. As part of the quality control of master discs and finished discs, CD analyzers are used. These special players display the data recorded on the disc and provide a simultaneous print-out. Often multiple players are used to check a percentage of several discs simultaneously, and then the results are combined for a composite report.

The analyzer indicates total playing time of the disc, the start time of each music selection as recorded in the table of contents, and the contents of the reference subcode. It checks for any subcode dropout or irregularities, noise generated by the error rate, and mistracking of the laser pickup. All pertinent information, such as the exact location of signal drop-outs, can be printed out. The human eye is also effective for spotting defects; under polarized light, an area of defective pits is visible as a blemish.

### ***100 Percent Automated Inspection***

In an automated quality control system, a rapid, automated device checks every disc for defects in less than 1 second. A laser beam scans each disc radially while the disc is rotated to produce 100 percent coverage of its surface. Defect detection is based on the normal reflection or transmission properties of the disc. Detection of all defects can be achieved only by using the optical properties of the disc itself, such as the diffraction grating formed by the pit track structure. To perform this analysis, some systems use two or more optical receivers in combination. Sensors are placed to detect light reflected in the zero- and first-order diffraction angles. This may be extended to higher-order diffraction angles as well.

Plane-polarized laser light is fitted to a polarization filter in the zero-order channel; it is sensitive to local birefringence and is thus able to detect small bubbles, inclusions, and other polycarbonate defects caused by localized stress. The first-order channel does not use a polarization filter; it is sensitive to metallization defects as well as faults in pit structure and track separation. Pinholes are detected with an optical transmission receiver shining on the disc.

Collecting the information from the light scattered by the different types of defects permits their classification and analysis. For example, damage on the upper surface of the disc produces a scattering in the incident light; this is recognized by a reduction in the amplitude of the first-order channel. Because the zero-order channel has a polarization filter, its amplitude will remain at zero unless birefringence causes a rotation in the plane of polarization. Thus an inclusion will cause an increase in the amplitude of the zero-order channel while decreasing the amplitude of the first-order channel.

Scanning optics are used to provide a normal angle of incidence in the radial and tangential planes across the disc surface. This is required to

reduce the refraction by the laser beam at the transitions between air, substrate, and metallized layer. Otherwise an error would occur in the measurement and location of defects due to distorting and de-focusing effects. A normal angle of incidence also prevents exaggeration of the local birefringence. The configuration of the laser scanning spot is crucial; experiments have shown that a rectangular spot is more sensitive to longitudinal defects, such as tangential scratches, than a circular spot of the same intensity. In one design, the rectangular scanning spot is thus made 10 micrometers radial by 150 micrometers tangential for audio CD checking. All optical scanning is controlled by extensive electronic evaluation equipment, which permits quick analysis and adjustment of scanning parameters as well as connection to a host computer for statistical compilation. In addition, options exist for reading a disc's alphanumeric code and automatically sorting discs from mixed production lines while inspecting each according to different levels of quality.

Many facilities utilize this automated inspection method. One hundred percent automated disc inspection is recommended in addition to human visual inspection for causes of rejection such as cosmetic defects. After all the quality control checks have been completed the disc is ready for packaging.

### ***Disc Packaging***

Most CDs are packaged in a plastic jewel box which is made of transparent polystyrene, with a tray typically of opaque polystyrene. Printed inlay cards are inserted into the box, the tray is put on, and the box closed. This can be done manually, incorporating a final visual inspection, or by an automatic machine which can package up to sixty discs per minute. The finished jewel box may be blister- or shrink-wrapped or inserted into a display container. Following insertion into a jewel box or other packaging and cellophane wrapping, the compact discs are ready for distribution to wholesalers.

### ***Manufacturing Summary***

CD plants built from 1982 to 1984, so-called first-generation plants, utilized large clean rooms, batch-type metallization systems, and little automation. Reject percentages were relatively high, and large amounts of downtime occurred. Plants constructed from 1984 to 1987 used second-generation technology with large clean rooms but with some isolation of heavy equipment outside of the clean areas to facilitate maintenance. Such plants generally employ in-line sputtering systems and some automation. However, similar to the first-generation plants, high energy consumption,

clean room costs, spoilage costs, amortization of larger capital investment, and downtime dictate higher operating costs.

The design of third-generation plants, first initiated in 1988, differs in a number of respects from earlier versions. First, they require about 25 percent less overall space, with very small clean rooms (about 10 percent of the space required by a second-generation facility). The replication processes are designed to be fully automated, including full integration of molding and metallizing as well as automatic quality control sampling. As a result, because standby capacity is provided, plant downtime is minimized. In addition, the high level of automation reduces the number of production workers and eliminates rejects caused by contamination of the discs by workers.

Third-generation plants employ a mono-line configuration, which combines all the phases of CD replication into a single module. The time from raw polycarbonate to finished disc, including injection molding, metallization, spin coating, and labeling, is less than 2 minutes. Annual disc capacity of two million discs may be expected from a single line. The module is housed within its own clean room, protected by Class 100 HEPA filters, thereby reducing clean room requirements. The station may occupy a 4 by 8 foot footprint. A module may be changed from one form of CD to another (e.g., 5-inch to 3-inch) within 1 hour.

## **ALTERNATIVE REPLICATION METHODS**

Conventional CD manufacturing methods described previously incur both technical difficulties (e.g., bubbles in the substrate and static charge on discs attracting dust particles) and production limitations (e.g., batch evaporation to metallize discs), both of which yield a relatively inefficient manufacturing process. As an alternative, new manufacturing methods with increased productivity have been investigated.

### ***2P Process***

The photopolymerization process (sometimes called the "2P" process) is not radically different from conventional injection molding of CDs. Instead of using polycarbonate thermoplastic resins, the 2P process uses a liquid photopolymer resin curable by ultraviolet light. Figure 7.14 shows the complete process. The liquid resin is injected into a mold with the pit stamper on one side and a prepared polycarbonate substrate on the other, and the mold is compressed. Ultraviolet light is radiated through the transparent substrate to cure the resin in the mold, and the hardened disc (pit formations on a 2P layer 30 micrometers thick, bonded to the substrate)

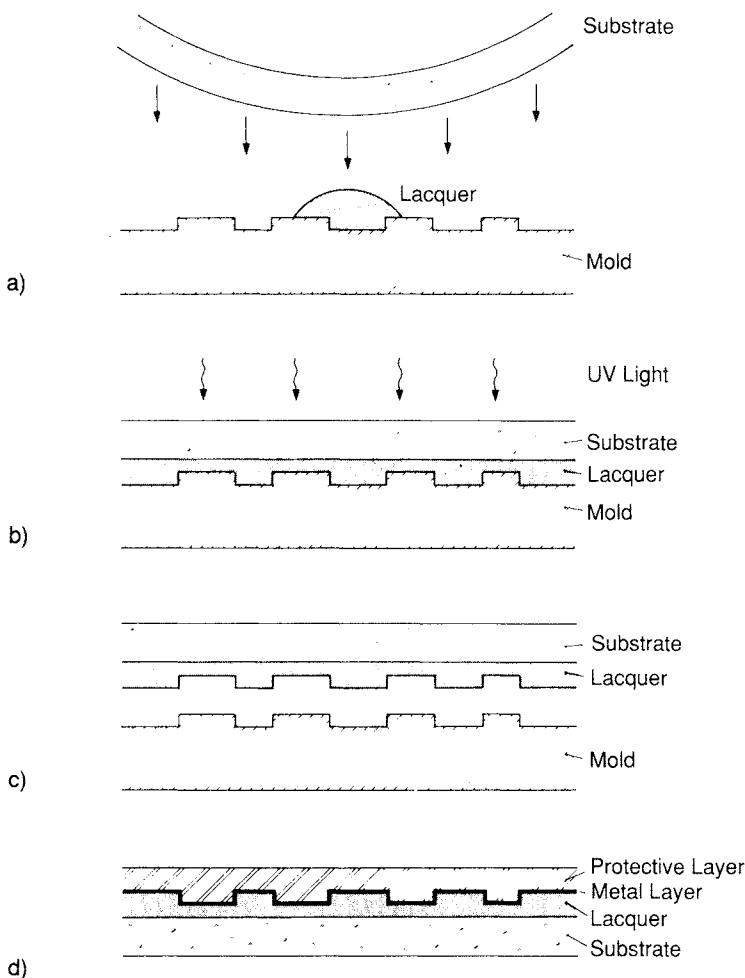


Figure 7.14 The four phases of the 2P process: the liquid lacquer is spread over the mold by a slightly deformed substrate (a); exposure to ultraviolet light to cure the lacquer (b); substrate with lacquer separated from the mold (c); lacquer coated with reflective and protective layers (d). (Courtesy Philips Technical Review)

is peeled from the mold. Subsequently, the disc is metallized, coated with a protective layer, and printed with a label, as with injection-molded CDs. Implementations of the 2P process have generally shown that it is cost-effective only for a relatively small production volume. Hence it is not an effective solution for the mass production of CDs.

### ***Continuous Methods***

Rather than relying on batch processing, in which CDs are manufactured through a sequence of individual steps, alternative methods strive to produce CDs in one continuous process. One such method utilizes a proprietary photo-thermographic transfer (PTT) process to produce stamped and laminated discs from extruded polycarbonate film and sheet. The process uses two layers of material, imprinted plastic film and polycarbonate, laminated together to form a disc, as shown in figure 7.15a. The pit data is imprinted on an optical-grade polycarbonate film with a nonmetallic stamper (made from a proprietary tool-making technique), and the film is then coated with a reflective material through vacuum deposition or sputtering, illustrated in figure 7.15b. In a separate process, the imprinted reflective film is coated with an upper protective layer and then laminated to a sheet of polycarbonate. Conventional ultraviolet curing or electron-beam laminating resins are used. Finally, the discs are printed with a label and cut from the sheet as shown in figure 7.15c. The finished discs reflect laser light from the pickup identically to discs made from injection molding.

In another method, a machine turns out CDs from a continuous roll of polycarbonate substrate, operating much like a rotary printing press. Aluminized mylar is laminated to embossed polycarbonate substrate, such as the one shown in figure 7.16a. The data is embossed on the substrate and then laminated to a reflective and protective mylar layer. The finished disc is punched from the continuous ribbon as shown in figure 7.16b. Although the reflective film is flat, the light from the laser pickup undergoes phase cancellation as with a regular CD. The reflection from the flat bottom of the pit (aided by the change in refractive index between the polycarbonate substrate and the pit) is out of phase with the reflection from the reflective film, creating phase cancellation when a pit is read and hence an intensity difference between pit and land.

This system uses the same master disc as the injection molding process but employs different stampers. Because the plastic substrate is first softened with a solvent and the stamping pressure is small, stampers have a long life expectancy. Because the pressing head is placed in a sealed environment, only a small, internal space requires clean room specifications. In theory, such a system could produce CDs at a rate of 1,000 per hour.

### ***Photolithographic Process***

Another alternative manufacturing method uses photolithography to replicate discs. As with conventional CD mastering, a glass plate is coated with photoresist and exposed by a laser beam to produce a positive-

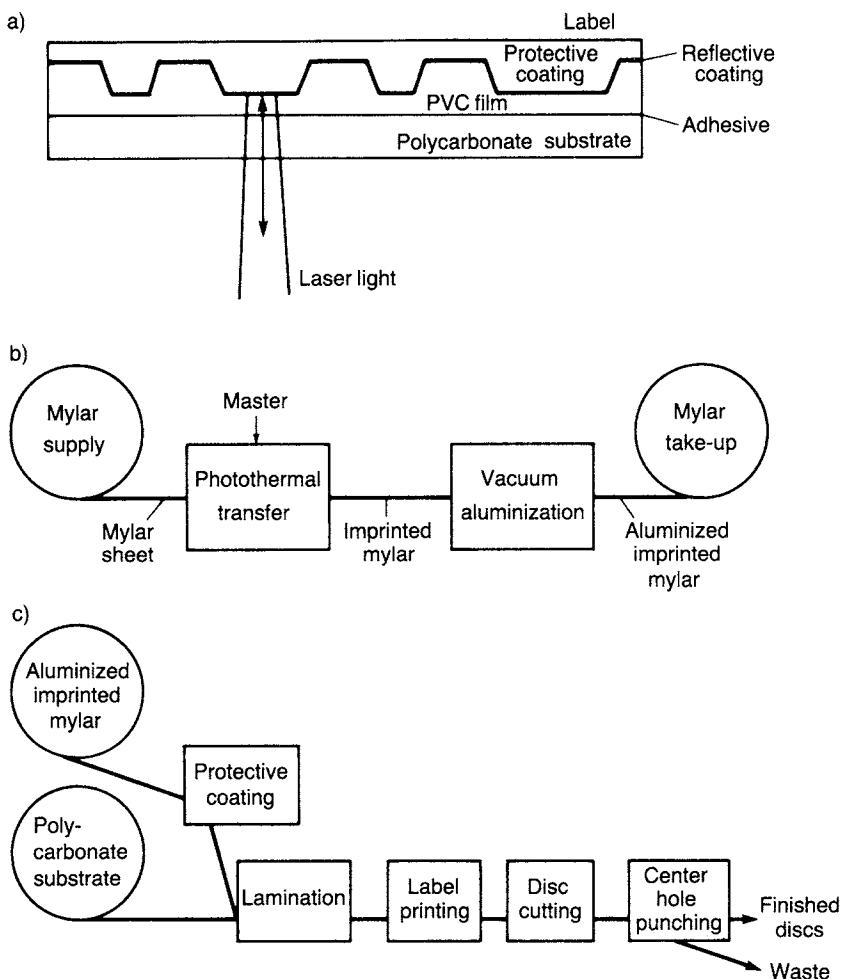


Figure 7.15 A continuous process in which an imprinted plastic film is laminated to a polycarbonate substrate (a). In the first manufacturing step, data is imprinted on the film and metallized (b). The film is coated with a protective layer and laminated to polycarbonate (c).

impression disc. However, instead of using electroplating and injection molding, the pattern of pits is contact printed, by means of an ultraviolet light photoresist and exposing process, onto a plastic disc substrate previously coated with reflective material. The process is, in fact, quite similar to the photographic process used to manufacture integrated circuits.

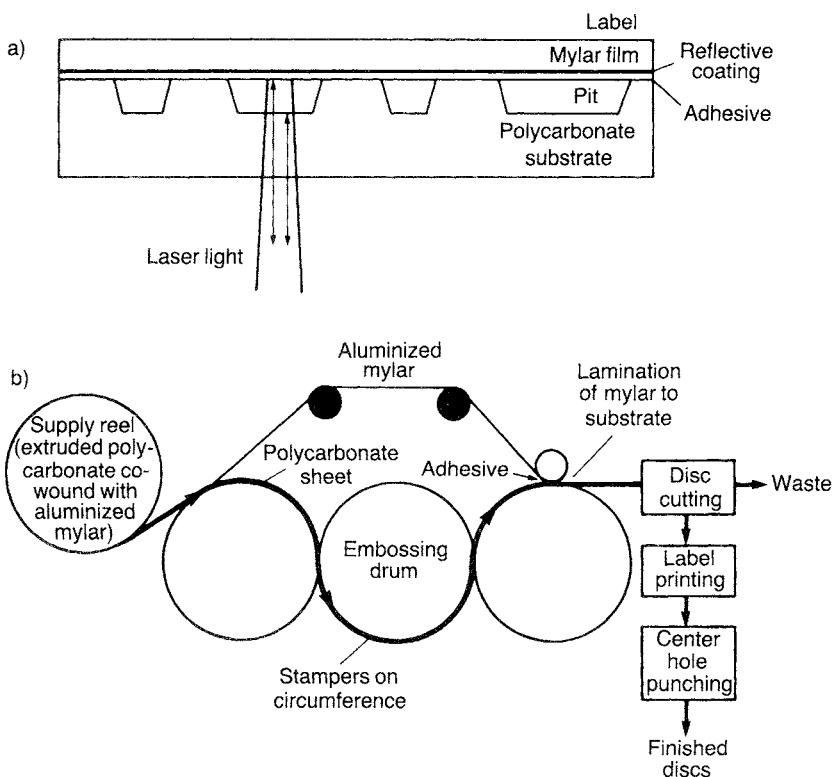


Figure 7.16 *A continuous process in which aluminized mylar is laminated to embossed polycarbonate (a). A rotary press is used to emboss the polycarbonate just prior to lamination (b).*

## **FOR FURTHER READING**

Foster, B. "CD Mastering: How It Should Happen." *Studio Sound* 26(10):76-81, October, 1984.

Fox, B. "CD Video—The Technology." *One to One Magazine*, 28-30, May, 1987.

Herring, P. "Commitment to CD." *One to One Magazine*, 42-44, February, 1987.

LeBlanc, A. "Quality Control and Testing Procedures for Compact Discs." *AES Preprint No. 2511*, 83rd Convention, 1987.

Leonarz, P. "Toolex Alpha: Manufacturing CDs." *One to One Magazine*, 50–53, May, 1987.

Miyaoka, S. "Manufacturing Technology of the Compact Disc." In Blessler, B., B. Locanthi, and T. G. Stockham, Jr., eds. *Collected Papers of the AES Premiere Conference—Digital Audio*. New York: Audio Engineering Society, 1983.

Redlich, H., and G. Joschko. "CD Direct Metal Mastering Technology: A Step toward a More Efficient Manufacturing Process for Compact Discs." *Journal of the Audio Engineering Society* 35(3):130–37, March, 1987.

Technoplas, Inc. "Cavity Pressure Control System for Injection Molding." November, 1986.

\_\_\_\_\_. "Fast-Cycle Precision Injection Molding Machine." November, 1986.

Turner, Ben. "CD Mastering: How It Happened." *Studio Sound* 26(10):82–88, October, 1984.

Verkaik, W., "Compact Disc (CD) Mastering—An Industrial Process." In Blessler, B., B. Locanthi, and T. G. Stockham, Jr., eds. *Collected Papers from the AES Premiere Conference—Digital Audio*. New York: Audio Engineering Society, 1983.



# Glossary of Technical Terms

## A

**Access Time** The time interval for the laser pickup to locate and begin reading data from any point on the disc.

**Acquisition Time** The time duration required for a sample-and-hold circuit to capture an input analog value; specifically, the time for the output to equal the input.

**A/D** See **Analog-to-Digital Converter**.

**Adaptive Delta Modulation (ADM)** A variation of delta modulation in which the step size may vary from sample to sample.

**Adaptive Delta Pulse Code Modulation (ADPCM)** A hybrid encoding technique that correlates successive data samples to adapt to changes in the signal.

**ADC** See **Analog-to-Digital Converter**.

**Address** A coded representation of the origin or destination of data.

**ADM** See **Adaptive Delta Modulation**.

**ADPCM** See **Adaptive Delta Pulse Code Modulation**.

**AES (Audio Engineering Society)** A professional engineering society that, among other activities, helps establish standards, including a professional digital audio interface standard.

**AES/EBU Digital I/O** A specification using bi-phase modulation for data and balanced line drivers to transmit digital audio data.

**Airy Pattern** A diffraction pattern with a central bright area and successively darker surrounding rings. A consequence of a diffraction-limited lens.

**Algorithm** A sequence of steps needed to perform a given task.

**Aliasing** Unwanted frequencies created when sampling a signal of a frequency higher than half the sampling rate (Nyquist frequency).

**Ambience** The acoustical characteristics of an environment.

**Amplifier** An electronic device for boosting signal strength.

**Analog** A physical quantity or data characterized by being continuously variable (rather than discrete).

**Analog-to-Digital Converter (A/D)** An electrical circuit which converts the instantaneous value of an analog signal to a digital word, represented as a binary number.

**Anti-Aliasing Filter** A low-pass filter used at the input of digital audio recorders to attenuate frequencies above the half-sampling frequency to prevent aliasing.

**Anti-Imaging Filter** A low-pass filter used at the output of digital audio reproducers to attenuate frequencies above the half-sampling frequency to eliminate image spectra present at multiples of the sampling frequency.

**Antinode** The point of maximum amplitude of vibration on a stationary wave.

**Aperture Circuit** The sample-

and-hold circuit used to buffer instability in the analog signal and correct for high-frequency roll-off following the D/A converter.

**Aperture Error** The high frequency roll-off occurring during de-sampling.

**ASCII (American Standard Code for Information Interchange)** A standard data transmission code consisting of 7 information bits and 1 parity bit. (ASCII is pronounced ask-ee).

**Astigmatism** The distorting property of certain lenses used in some optical pickups to generate a focus signal.

**ATRAC (Adaptive Transform Acoustic Coding)** An audio data reduction algorithm used in the Mini Disc format.

**Audio Frequencies** Frequencies that can be heard by the human ear, approximately 20 to 20,000 Hertz.

**Authoring System** Programs and equipment which enable nonprogrammers to write interactive and other programs, for example, for CD-I titles.

**Auto-Focus** The method used in CD player pickups to maintain the focus of the objective lens on the pit surface.

**Auto-Tracking** The method used in CD player pickups to maintain tracking of the laser spot on the pit track.

**Auxiliary Data Field** In CD-ROM Mode 1, a 288-byte field used for extended EDC/ECC. In Mode 2, a 280-byte field of user data.

## B

**Bandlimited** A signal occupying a defined and finite frequency band.

**Bandwidth** The range between lower and upper limiting frequencies; the width of a band of frequencies. Also, the maximum amount of information (e.g., bits) needed for transmission or storage.

**Base** See **Radix**.

**Base Case** The minimum equipment required to play a CD-I disc.

**Baud Rate** The transmitted data rate in bits per second (Baud).

**Beam** The laser light used in the optical pickup to recover data from the disc and to provide for focusing and tracking.

**Beam Splitter** See **Polarization Beam Splitter**.

**Binary** A condition in which there are two possible states; for example, the binary number system (base 2) using the digits 0 and 1.

**Birefringence** An optical dis-

tortion, such as the one in disc substrates.

**Bit** The smallest amount of digital information. A bit can store only two states, 0 or 1. Abbreviation for binary digit.

**Bit Error Rate (BER)** The number of bits processed before an erroneous bit is found (e.g.,  $10^{13}$ ), or the frequency of erroneous bits (e.g.,  $10^{-13}$ ).

**Bit-Mapped Display** A display in which each pixel's color and intensity data are stored in a separate memory location.

**Bit Rate** The rate or frequency at which bits appear in a bit stream. The bit rate of raw data from a CD is 4.3218 million bits/second.

**Bits per Second** A measure of the speed of transmission.

**Bit Stream** A binary signal without regard to grouping.

**Block** Data that is addressed or moved as a single unit. A data block is composed of sets of words to generate a data format. A typical block consists of data words, parity words, sub-code, synchronization, and error correction words.

**Block Error Rate (BLER)** The number of blocks at the C1 decoder with one or more errors.

**Blue Book** The Compact Disc-Video (CD-V) standards document.

**Brickwall** Refers to the slope of a filter with steep attenuation. Some anti-aliasing and anti-imaging filters have brickwall characteristics.

**Buffer** A memory used to provide consistent rate of data flow, such as from an inconsistent source. A crystal-clocked buffer virtually eliminates wow and flutter in digital audio systems.

**Burst Error** A large number of data bits lost on the medium because of excessive damage to or obstruction on the medium.

**Byte** A group of bits operating together, usually consisting of 8 bits.

## C

**Capacitor** An electrical component consisting of two metal plates separated by a nonconductor, used to store charge.

**Cassette** A self-contained package of magnetic tape.

**CAV** See **Constant Angular Velocity**.

**CD** See **Compact Disc**.

**CD-DA (Compact Disc Digital Audio)** System storing digital audio on a 12-centimeter diameter optical disc.

**CD + G/M (Compact Disc Graphics/MIDI)** System storing video and other nonaudio data in a disc's subcode area.

**CD + MIDI** System storing MIDI information in a disc's subcode area.

**CD-I (Compact Disc Interactive)** System storing digital audio, video, text, and graphics information interactively, with user control over presentation, on a 12-centimeter diameter optical disc.

**CD-PROM (Compact Disc Programmable Read Only Memory)** A write-once CD-ROM disc.

**CD-R (Compact Disc Recordable)** A system permitting audio or other data to be permanently written onto a 12-centimeter diameter optical disc and read by a CD-R or other CD player.

**CD-ROM (Compact Disc Read Only Memory)** A method of storing digitally coded information, such as computer information or database, on a 12-centimeter diameter optical disc.

**CD-ROM XA (Compact Disc Read Only Memory Extended Architecture)** An extended format of CD-ROM containing audio and video standards from the CD-I format.

**CD-RTOS (Compact Disc Real-Time Operating System)** The software required to operate and manipulate CD-I or CD-ROM systems, based on OS-9.

**CD-3** A system storing 20 minutes of digital audio on an 8-centimeter diameter optical disc.

**CDTV (Compact Disc Television)** An interactive, multimedia CD format storing digital audio, video, text, and graphics on a 12-centimeter diameter optical disc.

**CD-V (Compact Disc Video)** A system storing 5 minutes of analog video and digital audio plus 20 minutes of only digital audio on a 12-centimeter diameter optical disc, and longer times on 20- or 30-centimeter diameter optical discs.

**CD-WO (Compact Disc Write-Once)** See **CD-R**.

**Centi-** Prefix for one hundredth, abbreviated c.

**Charge** A quantity of electricity in a body.

**Chebyshev** A type of filter used in digital design which can provide a steep filter characteristic.

**Checksum** The use of redundant data for error detection.

**Chip** See **Integrated Circuit**.

**Chrominance** The color information in a color electronic image.

**CIRC** See **Cross Interleave Reed-Solomon Code**.

**Circuit** A system of conductors and electrical components through which electricity flows.

**Clean Room** An area with

strict cleanliness guidelines, used in CD manufacturing.

**Clock** A timing device that generates the basic periodic signal used as a source of synchronizing signals in digital equipment.

**Clock Rate** The time rate of pulses or events.

**CLUT** See **Color Lookup Table**.

**CLV** See **Constant Linear Velocity**.

**Code** A system of numerals or characters representing information.

**Collimator Lens** An optical component used in some pickups to make divergent light rays parallel.

**Color Lookup Table (CLUT)** A graphics encoding technique used in CD-I in which a table contains the colors for a particular image.

**Compact Disc (CD)** An optical disc storage system developed and licensed by Philips N.V. and Sony Corporation. The system stores 74 minutes, 33 seconds of digital audio and subcode information, or other nonaudio data, on a 12-centimeter diameter optical disc. The disc is made of plastic, with a top metallized layer, and is read by reflected laser light.

**Composite Video** A video signal in which luminance, chrominance, and synchronization are combined.

**Compression** See **Data Reduction**

**Concealment** Strategy used to supply approximate data in lieu of missing or incorrect data. See **Interpolation**.

**Constant Angular Velocity (CAV)** A disc rotating at a constant number of revolutions per minute. The LP is a CAV system at 33½ rpm.

**Constant Linear Velocity (CLV)** A disc rotating at varying numbers of revolutions per minute to maintain a constant relative velocity between pickup and track across the disc radius. The CD is a CLV system rotating from 200 to 500 rpm.

**CRCC** See **Cyclic Redundancy Check Code**.

**Cross Interleave Reed-Solomon Code (CIRC)** A method of error detection and correction using data delay, rearrangement, and the Reed-Solomon coding algorithm. Used in all CD formats.

**Crosstalk** The undesirable interference of one signal into another signal path.

**Crystal Oscillator** A circuit with an energized crystal producing fixed vibrations as in a clock.

**Cutoff Frequency** The frequency at which a filter effectively attenuates.

**Cycle** A complete sequence of a waveform that repeats at regular intervals.

**Cyclic Redundancy Check Code (CRCC)** An error detection code that is recorded with the original message information. CRCC is used in CD subcode data.

**Cylindrical Lens** An optical component used in some pickups to create an elliptical pattern through astigmatism to generate a focusing signal.

## D

**D/A** See **Digital-to-Analog Converter**.

**DAC** See **Digital-to-Analog Converter**.

**Data Reduction** The technique of reducing the amount of data required to represent a signal without overly affecting signal quality. Sometimes referred to as data compression.

**dB** See **Decibel**.

**Decibel (dB)** A measure of level equal to ten times the logarithm of a power ratio.

**Decimal** The base 10 number system, in which each numeral is multiplied by an appropriate power of ten.

**De-Emphasis** See **Emphasis**.

**De-Interleaving** See **Interleaving**.

**Delta Modulation (DM)** A single-bit coding technique in which a constant step size digitizes the input waveform. Past knowledge of the information permits encoding of the differences between consecutive values.

**Demodulation** See **Modulation**.

**Demultiplexing** See **Multiplexing**.

**Differential** The encoding technique in which the difference from bit to bit forms the basis of encoding.

**Diffraction** A modification in light or sound in which the incident rays are deflected.

**Diffraction Grating** An optical component used in three-beam pickups to split the laser beam into secondary beams to be used for tracking.

**Digital** The use of numbers (typically, binary numbers) to represent information.

**Digital Audio** The use of sampling and quantization techniques to store, process, or transmit audio information.

**Digital Filter** Any filter accomplished in the digital domain.

**Digital Signal** Any signal which is quantized at discrete points in time.

**Digital Signal Processing (DSP)**

The manipulation of signals in the digital domain by numerical methods.

**Digital-to-Analog Converter (D/A)**

An electrical circuit that converts a binary coded word to an equivalent analog voltage.

**Digital Video Interactive (DVI)**

A system using data compression techniques to store 70 minutes of digital video and audio information on a 12-centimeter diameter optical disc.

**Digitization** Any conversion of analog information into a digital form.

**DIN** Acronym for *Deutsche Industrie Norm*, the German standardization system.

**Direct Metal Master (DMM)** An alternative CD mastering process.

**Directory** Specification of the address of files on an electronic storage medium.

**Direct Read After Write (DRAW)**

An optical disc system in which data may be written (recorded) and immediately read (played back).

**Disc** A mechanical, optical, or magnetic storage medium.

**Disc Drive** Mechanism for supporting and rotating a disc.

**Disc Mastering** See **Mastering**.

**Discrete Time Sampling** The technique of representing a waveform at discrete instants of time.

**Disk Operating System (DOS)** A software program controlling data in memory and disk storage.

**Distortion** Undesirable change in an output signal relative to the input.

**Dither** Analog or digital noise added to the signal processing chain before quantization or re-quantization to minimize the effect of quantizing error at low-level signals.

**DMM** See **Direct Metal Master**.

**DRAW** See **Direct Read After Write**.

**Droop** An error condition in a sample-and-hold circuit in which the held analog value decreases.

**Dropout** An error condition in which bits are incorrect or lost from a medium. A disc defect or obstruction could cause a dropout.

**Dual Slope Integrating Digital-to-Analog Converter** A converter design which uses counters to time the output of integrators to yield an analog output.

**DVI** See **Digital Video Interactive**.

**Dye-Polymer** An optical stor-

age technology used in write-once and erasable systems.

**Dynamic Element Matching Digital-to-Analog Converter** A converter design which uses a series of current sources and switches to yield an analog output.

**Dynamic Range** The amplitude operating range of an electrical circuit, from noise floor to overload condition.

**DYUV** Delta YUV, a color encoding scheme for natural pictures with luminance encoded at full bandwidth and chrominance encoded at half bandwidth, storing only the differences between values.

## E

**EBU (European Broadcasting Union)** A professional society that, among other things, helps establish standards.

**ECC (Error Correction Code)**  
See **Error Correction**.

**EDC (Error Detection Code)**  
See **Error Detection**.

**EFM** See **Eight-to-Fourteen Modulation**.

**Eight-to-Fourteen Modulation**  
An encoding technique used in CD mastering to convert 8 bits of digital data to 14 bits. The player performs EFM demodulation.

**Electroforming** The production of metal impressions from master or submaster CD discs, to be used for disc molding.

**Emphasis** High frequencies in the audio signal can be boosted during recording, then de-emphasized or correspondingly attenuated during playback. This increases signal-to-noise ratio. A CD with emphasis is identified with a set emphasis bit, and the player automatically performs de-emphasis.

**Erasable Optical Storage** Optical media upon which data can be written, read, and then erased and written again.

**Erase** To destroy data stored on optical or magnetic media.

**Error** An incorrect data or signal value.

**Error Concealment** A method used to “repair” an error in the audio signal. Through interpolation, new data is derived from adjacent error-free data blocks.

**Error Correction** A method using a coding system to correct data errors by use of interleaved data or redundant data within a data block. Corrected data is identical to the original.

**Error Detection** A coding system which provides a method of determining errors that occur in a digital data stream. Accom-

plished prior to correction or concealment.

**Error Protection** The use of error detection, correction, and concealment to enhance the robustness of storage or transmission.

**Error Signal** A signal that is proportional to the difference between actual and desired conditions. Also, a signal that identifies the presence of an error.

**Extension** Programs added to an operating system to augment its capabilities (e.g., an MS-DOS extension to permit access of CD-ROM files).

**Eye Pattern** A display showing the high-frequency or radio-frequency data signal output from the optical pickup.

## F

**FET** Acronym for Field Effect Transistor, a unipolar semiconductor.

**Fiber Optics** The technology of using glass or plastic fibers to convey light and modulated information.

**File Descriptor Record** A sector found in CD-I files containing information on data segments, needed to access the file.

**File Structure Volume Descriptor** A record of the CD-I disc label describing files.

**Filter** A circuit designed to modify the amplitude or phase of a signal at certain frequencies.

**Flag** One or more bits used to store and indicate status of a circuit or system.

**Floating Point** An encoding technique in which a mantissa represents amplitude and an exponent provides a scaling factor.

**FMV** See **Full-Motion Video**.

**Foldover** See **Aliasing**.

**Format** The order in which data is organized or recorded on a medium.

**Four Quadrant Photodiode** A component found in some optical pickups used to convert reflected laser light into data and focusing signals.

**Fourier Theorem** A mathematical theorem stating that any periodic function may be resolved into sine and cosine terms with known amplitudes and phases.

**Frame** The basic unit of data representation in the CD format, containing synchronization, subcode, audio data, and parity.

**Frequency** The number of recurrences of a periodic waveform in time.

**FSFM (Full-Screen, Full-Motion) Video** See **Full-Motion Video**.

**Full-Motion Video** An exten-

sion to the CD-I format which permits moving pictures and the associated stereo audio signal to be digitally encoded on a CD using the MPEG standard.

## G

**Giga-** A prefix signifying one billion, abbreviated G.

**Glass Master** The glass disc used for CD mastering.

**Granulation Noise** An audible distortion resulting from quantization error.

**Green Book** The Compact Disc-Interactive (CD-I) standards document.

**Ground** The point in an electrical circuit at nominal zero voltage potential.

## H

**Hamming Codes** A class of error correction codes in which the error is self-locating.

**Hardware** The physical, mechanical, and electrical devices that form a system.

**Harmonics** A series of frequencies located at integer multiples of a fundamental frequency.

**Header Field** In CD-ROM, a 4-byte field of a data sector that

contains the absolute sector address and the mode byte.

**Hertz (Hz)** A measure of frequency (cycles per second).

**HF Signal** A high-frequency signal.

**High Sierra Group (HSG)** An ad hoc standards committee which established nominal data format and compatibility for CD-ROM.

**HSG** See **High Sierra Group**.

**Hygroscopic** Materials which readily take up and retain moisture.

**Hypermedia** Cross-linked databases possibly containing data in the form of text, pictures, and sound.

**Hz** See **Hertz**.

## I

**IC** See **Integrated Circuit**.

**IM** See **Intermodulation Distortion**.

**Image Pac** The data field used to store image information in the Photo CD format.

**Impulse Response** The output of a circuit or system based on an input impulse signal of infinitesimally narrow width. The response can be used to characterize the circuit or system.

**Index Numbers** Delineations found in the subcode of some CDs, marking points of reference in the audio program.

**Index of Refraction** The ratio of the velocity of light in one medium compared to a second medium.

**Information Exchange Protocol (IXP)** A protocol used in the mastering process for labeling data on a CD-I disc, describing media type and encoding techniques.

**Initialization** The process of reading the table of contents subcode when a CD is first loaded into the player, accomplished automatically.

**Injection Molding** Manufacturing process used to make CDs in which molten plastic is formed in a mold.

**Input** Any signal or data entering a circuit or system.

**Input Filter** See **Anti-Aliasing Filter**.

**Input/Output (I/O)** Equipment or data used to communicate from a circuit or system to other circuits or systems or to the outside world.

**Integrated Circuit** A solid state device with miniaturized active components on a single semiconductor material.

**Interactive Media** Programs or courses of instruction with content and output controlled by the viewer.

**Interleaving** The process of rearranging data in time. Upon de-interleaving, errors in consecutive bits or words are distributed to a wider area to guard against consecutive errors in the storage media.

**Intermodulation Distortion (IMD)** Undesired sum and difference frequencies in an output signal relative to the input.

**International Standard Recording Code (ISRC)** A code used by record manufacturers stating country of origin, owner, year of issue, and serial number of tracks, used optionally in CD subcode.

**Interpolation** The method used to conceal errors by using adjacent data to determine the approximate value of missing data. An error concealment technique.

**I/O** See **Input/Output**.

## J

**Jack** A socket into which connectors are inserted to make electrical contact.

**Jitter** Timing error.

## K

**kHz** Kilohertz, or 1,000 cycles per second.

**Kilo-** A prefix signifying one thousand, abbreviated k.

## L

**Land** The reflective substrate between pits on the CD data surface.

**Laser** See **Light Amplification by Stimulated Emission of Radiation**.

**Laservision** The trademark for an optical video disk format.

**LCD** See **Liquid Crystal Display**.

**Lead-In/Lead-Out** The nonaudio program areas on a disc preceding and succeeding the audio program.

**Least Significant Bit (LSB)** The bit within a digital word that is the least weighted.

**LED** See **Light Emitting Diode**.

**Light Amplification by Stimulated Emission of Radiation (Laser)** A device which generates coherent, monochromatic light waves. All CD players contain a semiconductor laser in their optical pickup.

**Light Emitting Diode (LED)** A self-lighting semiconductor di-

play of numerical or graphical information.

**Linear** A device having an output that varies in direct proportion to the input.

**Linearity** Often specifically used to describe the ability of a converter to accurately code low-level signals.

**Linear PCM** A pulse code modulation system in which the signal is converted directly to a PCM word without companding or other processing.

**Liquid Crystal Display (LCD)** A semiconductor display of numerical or graphical information, requiring ambient light or backlighting.

**Logic** Systematized rules or interconnections of digital switching circuits.

**Low-Bit Conversion** A technique used for both A/D and D/A conversion in which the sampling rate is increased and the word length is greatly reduced.

**Low-Bit D/A Converter** A converter architecture in which a few bits (e.g., 1 to 4) at a high oversampling rate are converted to form the analog output waveform. Also known as 1-bit conversion. See also **Noise-Shaping Circuit**.

**Low-Pass Filter** A filter with a characteristic that allows all frequencies below a specified cutoff

frequency to pass and attenuates all frequencies above. Anti-aliasing and anti-imaging filters are low-pass filters.

**LSB** See **Least Significant Bit**.

**Luminance** The intensity or brightness component of an electronic image.

## M

**Magnetic Tape or Disc** An analog or digital storage medium consisting of a magnetically coated substrate storing data as magnetic patterns.

**Magneto-Optical Recording (MOR)** An erasable optical disc system using magnetic media and laser reading/writing.

**MASH (Multi-Stage Noise Shaping)** A type of noise-shaping algorithm in A/D and D/A converters, employing third-order or higher-order shaping.

**Master** The original recording of a finished program.

**Mastering** The process of encoding audio and control data on a glass master disc prior to CD replication.

**Matrixing** See **Electroforming**.

**Medium, Media** A material or basis on which information is stored or transmitted.

**Mega-** A prefix signifying one million, abbreviated M.

**Memory** Hardware used to store information.

**Menu** A list of options displayed for a user.

**Menu-Driven** Software controlled through menus.

**Metallization** Step in the CD replication process in which the polycarbonate substrate is coated with metal to provide reflectivity.

**Micro-** Prefix for one thousandth, abbreviated  $\mu$ .

**Microprocessor** An integrated circuit that performs a variety of operations in accordance with a list of instructions.

**MIDI (Musical Instrument Digital Interface)** A standard code used to control operation of electronic music instruments.

**Milli-** Prefix for one thousandth, abbreviated m.

**Mini Disc** A recordable/erasable optical disc format using a 2.5-inch diameter magneto-optical media; borrowing from CD technology, but incompatible with it.

**Mode 1 and Mode 2** The two frame formats used in CD-ROM. CD-I is based on Mode 2.

**Modulation** The process of varying a carrier signal according to the instantaneous value of an input signal.

**Mono-Line** A CD manufacturing system in which replication processes are integrated, with automated control.

**MOR** See **Magneto-Optical Recording**.

**MOS** Acronym for Metal Oxide Semiconductor, a type of field effect transistor.

**Most Significant Bit (MSB)** The bit within a digital word that is the highest weighted.

**MPEG (Moving Pictures Expert Group)** A working committee and the audio and video compression standards it devised; implemented as an extension to the CD-I format.

**MSB** See **Most Significant Bit**.

**MS-DOS** A disk operating system used in many personal computers.

**Multibit D/A Converter** A converter architecture in which a number of bits (e.g., 16) are simultaneously converted to form the output analog waveform.

**Multiplexing** The process of combining independent signals and conveying them along a single conductor.

## N

**NA** See **Numerical Aperture**.

**Nano-** A prefix for one billionth, abbreviated n.

**Node** A point of zero amplitude of vibration on a stationary wave.

**Noise** Any spurious signal.

**Noise Floor** The lowest threshold of useful signal level.

**Noise-Shaping Circuit** A numerical feedback circuit used to increase the signal-to-noise ratio of low-bit D/A converters.

**Non-Photoresist (NPR)** An alternative CD mastering method.

**Non-Return-to-Zero (NRZ)** A digital code in which polarity changes positively with a 0/1 signal change and negatively with a 1/0 signal change.

**Non-Return-to-Zero-Inverted (NRZI)** A digital code in which polarity is reversed each time a 1 occurs and remains constant when a 0 occurs.

**NPR** See **Non-Photoresist**.

**NRZ** See **Non-Return-to-Zero**.

**NRZI** See **Non-Return-to-Zero-Inverted**.

**NTSC Video Format** The color television standard used in the U.S. and prepared by the National Television Systems Committee of the Electronics Industries Association (EIA).

**Numerical Aperture (NA)** A measure of the relative capacity

of an optical system to accept light from an outside medium; a measure of the acceptance angle of a lens.

**Nyquist Frequency** The highest frequency that may be accurately sampled. The Nyquist frequency is one-half the sampling frequency. The theoretical Nyquist frequency of the CD system is 22.05 kHz.

## O

**Objective Lens** The optical component used in pickups to focus and collect laser light on the disc data surface.

**OEM** See **Original Equipment Manufacturer**.

**Offset (Subcode)** The difference in the number of frames between access-begin time and actual start time. A safety margin for player muting/de-muting.

**Offset (Voltage)** A DC voltage that raises or lowers a signal from a zero or a reference level.

**Omni-Player** A CD player that will play discs of any CD format.

**Op Amp** See **Operational Amplifier**.

**Operational Amplifier** An analog device often used as a gain block.

**Optical Disc** Any read-only, write-once, or erasable medium

using laser light to convey data to or from the disc.

**Optical Pickup** The optical device using laser light to convey data from a disc.

**Opto-Isolator** Optical components which convert electrical signals to light and then back to electrical. Used in some CD players to isolate the noise of digital circuitry from analog circuitry.

**Orange Book** The Compact Disc-Recordable and Compact Disc-Write Once standards document.

**Original Equipment Manufacturer (OEM)** A term used to describe how one manufacturer obtains finished parts of its equipment from another manufacturer and then markets the assembled device under its own name.

**OS-9** The real-time operating system on which the CD-I operating system is based.

**Output** Any signal or data leaving a circuit or system.

**Output Filter** See **Anti-Imaging Filter**.

**Oversampling** The technique in which the number of samples is multiplied to permit digital filtering of the signal prior to conversion. A gentle analog low-pass filter is still required to remove extreme supersonic frequencies. Used in both A/D and D/A conversion.

## P

**PAL Video Format** The color television format used in Europe (except France), Africa, Australia, and South America. PAL is an acronym for phrase alternation line.

**PAM** See **Pulse Amplitude Modulation**.

**Parallel** The simultaneous presence of bits in a data word on multiple conductors.

**Parity** A redundant error detection method in which the total number of binary 1s (or 0s) is always even or odd.

**PBS** See **Polarization Beam Splitter**.

**P Channel** Subcode channel carrying information on lead-in, lead-out, and play areas of a CD. See **Subcode**.

**PCM** See **Pulse Code Modulation**.

**Peripheral** Equipment physically independent of a controller.

**Phase Change Medium** An optical storage method used in both write-once and erasable media in which a change in reflectivity is used to encode data.

**Phase Shift** A displacement in the time relationship between two signals or between two components in one signal.

**Photo CD** A method of digitally encoding up to 100 high-quality photographic images on a CD; compatible with the CD-I format.

**Photodiode** A semiconductor device which generates an electrical current in proportion to received illumination.

**Photopolymerization (2P)** An alternative CD manufacturing method.

**Photoresist** A substance which becomes soluble (or insoluble) when exposed to light. Used in CD glass disc mastering.

**Photo-YCC** The image encoding method used in the Photo CD format, employing an image data compression algorithm.

**Pickup** See **Optical Pickup**.

**Pico-** Prefix for one trillionth, abbreviated p.

**Pitch** The distance between adjacent tracks of a data track. The pitch of the pit spiral on CDs is nominally 1.6 microns.

**Pits** The physical impressions on a disc substrate. Viewed as bumps by the reading laser, they diffract the light, decreasing the intensity of light returned to the pickup.

**Pixel** A picture element, the smallest dot on a video screen.

**Placebo** In the context of CD treatments, a way to profit from the weak-minded.

**Player** The mechanical and electrical device used to recover data from a disc or other media.

**PNM** See **Pulse Number Modulation**.

**Polarization Beam Splitter (PBS)** An optical component used in some pickups to conduct light in one plane of polarization but reflect it in another.

**Polycarbonate** The transparent plastic material used to form the substrate for CDs.

**Potentiometer** A device with variable resistance.

**Power Supply** A circuit that supplies proper voltages and currents as required by the system.

**PPM** See **Pulse Position Modulation**.

**PQ Subcode** See **Subcode**.

**Pre-Emphasis** A high-frequency boost during recording, followed by de-emphasis during playback, designed to increase the S/N ratio.

**Pre-Mastering** The process of integrating digital data code, error correction codes, and sub-codes and converting these to the CD data format prior to disc mastering.

**Psychoacoustics** Study of the perception and response to sound.

**Pulse Amplitude Modulation (PAM)**

A modulation code in which the amplitudes of pulses in a pulse train represent the analog information.

**Pulse Code Modulation (PCM)**

A modulation code in which digital words in a bit stream represent samples of analog information. The basis of most digital audio systems.

**Pulse Density Modulation (PDM)**

A modulation code in which one bit may be averaged to form an analog waveform.

**Pulse Number Modulation (PNM)**

A modulation code in which the number of pulses in a pulse train represents the analog information.

**Pulse Parameter Modulation** The general type of modulation in which information is encoded by pulse representation.**Pulse Position Modulation (PPM)**  
A modulation code in which the positions of pulses in a pulse train represent the analog information.**Pulse Width Modulation (PWM)**

A modulation code in which the widths of pulses in a pulse train represent the analog information.

**PWM** See **Pulse Width Modulation**.

**Q**

**QC** See **Quality Control**.

**Q Channel** A subcode channel carrying information on track and index numbers, elapsed time, product codes, etc. See **Subcode**.

**Quality Control** The series of tests and measurements used to maintain manufacturing tolerances—for example, those of discs.

**Quantization** The process of converting an infinitely variable amplitude of an analog waveform to one of a finite series of discrete levels.

**Quantization Error** Error resulting from quantizing an analog waveform to a discrete level. The longer the word length, the less the error.

**Quarter Wave Plate (QWP)** An optical component used in some pickups to shift the plane of polarization of the laser light.

**QWP** See **Quarter Wave Plate**.

**R**

**Radial Tracking Error** A measurement of a player's radial tracking servo signal, used to determine tolerance for disc flatness.

**Radix** The number of symbols used in a number system.

**RAM** See **Random Access Memory.**

**Random Access Memory (RAM)** A memory in which data may be read and written regardless of address or location.

**Random Bit Error** An error in a digital data stream in which only a few bits are lost. Corrected by error correction.

**R-DAT** See **Rotary Head Digital Audio Tape Recorder.**

**Read** The nondestructive reading of data from a medium.

**Read-Only-Memory (ROM)** A memory from which data, after initial storage, may only be read. CD-DA is a read-only system.

**Real-Time** Operation which is perceived to be instantaneous to a user.

**Reconstruction** The final conversion step from de-sampled data to analog waveform, achieved through low-pass filtering.

**Red Book** The Compact Disc-Digital Audio (CD-DA) standards document.

**Reed-Solomon Code** A cyclic, multiple-error correction code used in CDs. See **Cross Interleave Reed-Solomon Code.**

**Refractive Index** See **Index of Refraction.**

**Reverberation** Repeating echoes of sound, increasing in density, decreasing in amplitude.

**RF Signal** Radio frequency signal.

**RGB (Red-Green-Blue)** A type of color output to a display consisting of separate signals for red, green, and blue—an alternative to composite video.

**RGB 5:5:5** An RGB encoding technique used in CD-I in which colors are each represented by 5 bits.

**ROM** See **Read-Only-Memory.**

**Root Directory** The highest-level directory on a CD-I disc.

**Rotary Head Digital Audio Tape Recorder (R-DAT or DAT)** A digital audio recorder utilizing a magnetic tape cassette system similar to that of a video recorder.

**RTOS (Real-Time Operating System)** An operating system with real-time interaction.

## S

**Sample-and-Hold Circuit (S/H)** A circuit which captures and holds an analog signal for a finite period of time. See **Aperature Circuit.**

**Sampling** The process of periodically examining a signal at particular points in time.

**Sampling Frequency** The frequency at which an analog signal is sampled. Expressed in Hertz. The sampling frequency used on compact discs is 44,100 samples per second (44.1 kHz).

**Sampling (Nyquist) Theorem** Theorem stating that a bandlimited continuous waveform may be represented by a series of discrete samples if the sampling frequency is at least twice the highest frequency contained in the waveform.

**SAR** See Successive Approximation Register.

**SCMS (Serial Copy Management System)** A copy inhibit method used in many consumer digital audio recorders to permit first-generation digital copying only.

**SCSI** See Small Computer Standard Interface.

**SECAM Video Format** The color television format used in France and the former Soviet Union in which the three primary colors are sent sequentially rather than nearly simultaneously. SECAM is an acronym for Séquentielle Couleur à la Mémoire.

**Sector** In CD-ROM, a block or frame complete with synchronization and header field; a 2352-byte block.

**S/E Ratio** See Signal-to-Error Ratio.

**Serial** Sequential presence of data along one conductor.

**Serial Transmission Format** See AES/EBU Digital I/O.

**Servo Mechanism** A control device or system which reads its own output to determine the degree of further output as compared to its input.

**Settling Time** The time required for a circuit, such as a D/A converter, to reach a stable output value.

**S/H Circuit** See Sample-and-Hold Circuit.

**Signal-to-Error Ratio (S/E)** The ratio of signal voltage to quantization error voltage, typically expressed in decibels.

**Signal-to-Noise Ratio (S/N)** The ratio of signal voltage to noise voltage, typically expressed in decibels.

**Sine Wave** A fundamental waveform corresponding to a single frequency, without harmonics.

**Single-Beam Tracking** An optical pickup design in which a single laser spot is used for data recovery, tracking, and focusing.

**Sled** A mechanical chassis upon which the optical pickup is supported and moved along the disc surface.

**Small Computer System Interface (SCSI)** A standard 8-bit parallel interface used to connect peripherals, such as connecting a CD-ROM player to a microcomputer. (SCSI is pronounced scuzzy.)

**Smoothing Filter** See **Anti-Imaging Filter**.

**SMPTE (Society of Motion Picture and Television Engineers)** A professional engineering society that, among other activities, helps establish standards, including a timecode standard.

**S/N Ratio** See **Signal-to-Noise Ratio**.

**Software** A set of instructions or programs used to operate a data processing system. Also, prerecorded data or music programs.

**Sound** A variation in air pressure within frequency limits from about 20 Hz to 20,000 Hz.

**Spectrum** A distribution of frequencies in a bandwidth.

**Spin Coating** A step in the CD replication process in which a protective layer is applied over the metallized disc surface.

**Spindle** Part of the drive that spins the disc.

**Spiral Track** The pit track spiral physically encoding data on a CD, running from the innermost region to the outermost.

**Sputtering** A method used in the CD replication process to metallize a disc data surface.

**Stamper** The metal mold used in polycarbonate plastic injection molding machines to replicate CDs.

**Subcode** Data encoded on a CD that contains definable information such as track number, times, copy inhibit, copyright, etc.

**Subheader** A field indicating the type of data in a sector.

**Substrate** The material upon which a device is constructed and supported. A CD uses a polycarbonate substrate.

**Successive Approximation Register (SAR)** A type of analog-to-digital converter using a digital-to-analog converter to determine the output word successively, bit by bit.

**Symbol** An 8-bit byte.

**Synchronous** A simultaneously timed operation in which a master clock is used to control events.

## T

**Table of Contents (TOC)** Nonaudio subcode data contained in the lead-in area of a CD, storing initialization data, including track and time information. Read during initialization.

**Test Disc** A specially prepared CD containing signals and program material designed to measure and evaluate CD players and other equipment.

**THD** See **Total Harmonic Distortion**.

**Three-Beam Tracking** An optical pickup design in which a laser beam is split into three spots, the center spot used for data recovery and focusing, and the secondary spots for tracking.

**Time Sampling** See **Discrete Time Sampling**.

#### **Total Harmonic Distortion**

**(THD)** Undesired harmonics in an output signal relative to the input. Often measured with corresponding noise (THD + N).

**Tracking Error** An error condition in which the optical pickup deviates from the pit track, perhaps caused by an obstruction or disc defect.

**Track Numbers** Numbering of songs or movements on a disc; this information is contained in the subcode.

**Track Pitch** See **Pitch**.

**Transversal Filter** A kind of digital filter used in CD players to perform low-pass filtering of the output signals; used in an oversampling-based architecture.

**Truncate** To eliminate without round-off some low-order bits

after performing an arithmetic computation.

**Two-Axis Actuator** Mechanical device holding the objective lens of an optical pickup, able to move the lens horizontally and vertically for tracking and focusing.

**Two's Complement** A binary representation that handles positive and negative values and facilitates arithmetic. The most significant bit is the sign bit.

## **U**

**Ultraviolet** Electromagnetic radiation at frequencies higher than those of visible light yet lower than those of x-rays.

**U-Matic Tape Recorder** A rotary head tape recorder used for CD mastering.

## **V**

**Velocity of Sound** The speed of sound, 344 meters per second at 21 degrees centigrade.

**VMS** An operating system used in some minicomputers.

## **W**

**Waveform** A periodic propagated disturbance in a medium.

**Weighted Resistor Converter** A type of digital-to-analog converter using resistors with values varying as powers of two, used to generate a current proportional to the input word.

**Wet Silvering** A method of metallizing a disc data surface.

**White Noise** A noise with equal energy at any frequency.

**Word** A convenient collection of bits, commonly 16 bits.

**Word Length** The number of bits in a word.

**Wow and Flutter** Variation in transport speed from mechanical error. Made negligible in the CD system through data buffers.

**Write** To record data on a medium.

**Write-Once Media** Systems in which data may be written once but not erased and rewritten.

## Y

**Yellow Book** The Compact Disc-Read Only Memory (CD-ROM) standards document.

**YUV** Video coding process used in CD-I in which the luminance signal (Y) is recorded at full bandwidth on each line and chroma values (U and V) are recorded at half bandwidth on alternate lines.



# Index

- Access  
random 47, 107, 119, 282  
time 47, 168, 192, 282
- Acousto-optical modulator 289
- A/D (see Analog-to-digital converter)
- Adaptive delta modulation 227
- Adaptive delta pulse code modulation  
225, 227–30, 245
- ADC (see Analog-to-digital converter)
- Address 92, 94–96, 215
- ADM (see Adaptive delta modulation)
- ADPCM (see Adaptive delta pulse code modulation)
- AES (Audio Engineering Society) 158
- AES/EBU digital I/O 157–60
- Airy pattern 57
- AlGaAs 108
- Algorithm 144, 150, 228, 244
- Aliasing 18, 22–27, 121, 157
- Ambience 199
- Amplifier 39, 44, 132, 178, 197
- Analog  
continuity 15  
filters 121–27  
recording 2–4  
systems 1  
to digital (see Analog-to-digital converter)
- Analog-to-digital converter 36, 41
- Angle of deviation 302
- ANSI 158
- Anti-aliasing filter 18, 26, 39
- Anti-imaging filter 19, 24, 43, 121–32,  
142, 145, 146, 151 (see also  
Filter and Oversampling)
- Aperture circuit 39, 43, 140–42, 152
- Aperture error 140–41
- Artwork 285, 286
- Astigmatism 107, 111
- Atmospheric pressure 2
- ATRAC (Adaptive transform acoustic coding) 263–65
- Audio  
aliasing (see Aliasing)  
band 123  
CD audio specifications 49  
chronology 8–12  
digitization systems 38–44  
frequencies 19, 180, 184  
measurements 178–95, 306  
modulation 35–38  
output 90, 146, 157, 201  
performance 167–78, 195–200  
resolution 33  
signal 17, 103, 178, 199
- Automatic  
focus 111, 117  
gain control 87  
inspection 308  
tracking 113, 117
- Auxiliary data field (CD-ROM) 216
- A-weighting 187
- Bandlimiting 18, 22, 123, 193
- Base (see Radix)
- Beam  
focusing 54, 107–8, 111–13, 117  
reflection 55, 107, 110  
tracking 54, 107, 113–15, 117–18
- Beam splitter (see Polarization beam splitter)
- Beethoven 1, 11
- Begin-access points 282

- Binary
  - digits 28
  - numbers 28
  - words 29, 74
- Birdies 32
- Birdsing 32
- Birefringence 295, 297–98, 302, 305, 308
- Bit error rate 73, 217
- Bit-mapped display 233, 268
- Bit rate 82, 91, 105
- Bit rate reduction 229 (see also Data reduction)
- Bits
  - merging 74, 80
  - parity 62, 67
  - subcode (see Subcode)
  - synchronization 82, 88, 91
- Bit stream 42, 89, 105
- Bitstream (Philips) 152
- Block 62
- Block error rate 73, 251, 306
- Block floating point 44
- Blue Book* 247
- Brickwall (see Anti-aliasing filter)
- Buffer 89, 145, 265
- Burst error 64, 72, 218, 306
- Byte 29
- CAV (see Constant angular velocity)
- CCIR 238
- CD (see Compact disc)
- CD-DA (see Compact disc)
- CD+G/M
  - graphics 99, 267–73
  - MIDI 269
  - Mode 1 268
  - Mode 2 270
  - reduction 272
- CD-I
  - applications 221
  - audio 224–30
  - base system 236
  - capacity 222
  - data format 222–24
  - FMV 231–33, 236
  - interactivity 234
  - MPEG 231–33, 236
  - player 234–37
  - standard 221
  - text 233–34
  - video 230–33
- CD+MIDI 269–72
- CD-R
  - ablation 249
  - disc layout 250
  - dye-polymer 249
  - PCA 250
  - phase change 249
  - PMA 250
  - standard 254
  - thin film 249
  - TOC 249
  - UTOC 250
- CD-ROM
  - applications 214
  - capacity 215
  - Data Discman 218
  - data format 215–17
  - encyclopedia 214
  - error correction 217–18
  - player 218
  - standards 220
- CD-ROM XA 220, 238
- CD-RTOS 234
- CD-3 273
- CDTV 242–43
- CD-V
  - CDV-EP 248
  - CDV-LP 248
  - CDV-Single 247
  - Combi-player 246
- CD-WO (see CD-R)
- Channel bits 82, 289
- Channel separation 180
- Character encoding (CD-I) 233
- CIRC (see Cross Interleave Reed-Solomon Code)
- Clean room 290, 294, 299, 309
- Clock 89, 105, 171
- CLUT (see Color lookup table)
- CLV (see Constant linear velocity)
- Code 35, 62
- Coercivity 256
- Collimator lens 109, 117
- Color lookup table 220, 231, 237, 269
- Compact disc
  - advantages 6–8, 47, 167
  - applications 47, 214, 221, 237, 242, 243, 246, 250, 260, 261, 267, 273
  - chip set 142–47
  - chronology 8–12
  - decoding 85–90

- digital audio 47  
disc 48–60  
encoding 74–85, 289  
erasable 255–67  
error correction 60–74, 304–6  
format family 223  
information capacity 83, 215, 218  
manufacturing 204, 277–314  
mastering 277–93  
player (see Player)  
specifications 48, 167, 178, 302  
subcode 76, 90–99, 253  
testing 178–95  
20-bit 99  
video 230, 231, 237, 242, 243, 267  
writable 249–55  
Companding 44, 229  
Compatibility 223, 234, 246  
Compression 196 (see also Reduction)  
Concealment (see Error concealment)  
Condensation 202  
Constant angular velocity 54, 248, 291  
Constant linear velocity 53, 248, 291  
Conversion  
    A/D (see Analog-to-digital converter)  
    D/A (see Digital-to-analog converter)  
    floating 138  
    resolution 30, 137–40, 187–89  
CRCC (see Cyclic redundancy check code)  
Cross Interleave Reed-Solomon Code  
    decoding 71–73, 305  
    encoding 69–71  
    function 62–64, 67, 217  
    performance 73–74, 305  
    polynomials 68  
Crossfade 282  
Crosstalk 180  
Cue sheet 285  
Curie point 256, 261  
Cutoff frequency 19, 123  
Cutting lathe 286–89, 291–93  
Cutting signal 289  
Cyclic redundancy check code 92, 159  
Cylindrical lens 111
- D/A (see Digital-to-analog converter)  
DAC (see Digital-to-analog converter)
- Data  
    blocks 62  
    decoding 85–90  
    disc 82  
    encoding 74–85  
    frame 74, 82  
    PCM (see Pulse code modulation)  
    reduction 22, 44, 224, 231–33, 238–41, 263–65, 272  
    storage 82  
Data Discman 218  
DCC (digital compact cassette) 266  
DC offset 87  
Decimal 29  
Decimation 173  
De-emphasis (see Emphasis)  
De-glitcher 43, 141  
De-interleaving (see Interleaving)  
Delta modulation 227  
Demodulation (see Modulation)  
Demultiplexer (see Multiplexer)  
Diffraction 52, 55–59, 109  
Diffraction grating 51, 57, 108, 308  
Digital  
    advantages 3, 47, 167  
    audio 3, 47  
    error 27, 73  
    error correction (see Error correction)  
    filter 127–32, 172 (see also Filter and Oversampling)  
    output 157–63  
    processor 278  
    recording 2, 249–67  
    signal 86–90, 157–63, 304–6  
    storage 48–56  
    systems 1, 38–44  
    to analog (see Digital-to-analog converter)  
    video 230–33, 237–48  
Digital compact cassette (DCC) 266  
Digital signal processing  
    applications 106, 119  
    error correction 60–74  
    evolution 106  
    filters 127–32  
    low-bit conversion 147–57  
    noise shaping 148–52  
    time delay 127  
Digital sum value 80  
Digital-to-analog converter  
    adaptive 138

- Digital-to-analog convert, (*continued*)  
 dynamic element matching 134, 146  
 function 36, 43, 132–40  
 integration 134  
 ladder 133  
 linearity 137–40, 187–89  
 low-bit 147–57
- Direct metal mastering 292–93
- Directory 220, 235, 241
- Direct read after write 291
- Disc  
 center hole 48, 205, 296, 302  
 clamping 49  
 cleaning 208  
 decoding 85–90  
 defects 204–7, 297, 302–9  
 developing 289  
 dimensions 48  
 drive 103–5 (see also Player)  
 encoding 74–85, 250, 289  
 handling 207–8  
 manufacturing 204–5, 277–314  
 mastering 74, 204, 277–93  
 operating system 234, 237, 241  
 packaging 208, 309  
 player (see Player)  
 pre-mastering 277–86  
 reflectivity 54, 58, 87, 299, 302  
 rings 208–11  
 specifications 49, 178  
 test signals 178  
 treatments 208–11  
 warpage 107, 111, 206, 304
- Discrete cosine transform 232
- Discrete time sampling (see Sampling)
- Dither  
 function 32–35, 152, 179, 187, 188  
 types 35
- DMM (see Direct metal mastering)
- DRAW (see Direct read after write)
- Droop 41
- Dropout 61, 65, 192–93, 280, 286
- Dropout compensator (U-matic) 280
- DSP (see Digital signal processing)
- DVI  
 data reduction 243  
 specification 243  
 video 244
- Dye-polymer 249, 266
- Dynamic range 30, 187
- DYUV (Delta YUV) 230, 231, 237
- EBU (European Broadcasting Union) 158
- ECC (see Error correction)
- Eccentricity 107, 304
- EDC (see Error detection)
- Editing 253, 281
- EFM (see Eight-to-fourteen modulation)
- EIA (Electronics Industries Association) 179
- EIAJ (Electronics Industries Association of Japan) 179, 183, 185, 192
- Eight-to-fourteen modulation 77, 86–89, 105, 144, 215, 223, 289
- Electroforming 293–94
- Embossing 292, 312
- Emphasis 83, 92, 159, 191, 281, 283
- End-access points 282
- Erasable optical storage 255–67
- Erasure 69
- Error(s)  
 burst 64, 72, 306  
 concealment 65–67, 145, 196–97  
 correction 60, 62–64, 144, 192–93, 196, 217  
 detection 61–62, 144, 217  
 flags 65, 71–72, 304  
 random 64, 72, 306  
 signals 304–6  
 testing 192–93, 302–6  
 timebase 88, 89
- Evaporation (metal) 290, 293
- Extended error correction 216–18, 223
- Eye pattern 86
- Faraday effect 257
- Father 294
- FCC (Federal Communications Commission) 201
- Fiber optics 163, 170, 201
- File 220, 235, 241
- File descriptor record 235
- File structure volume descriptor 235
- Filter  
 analog 39, 121–27, 172, 191  
 anti-aliasing (see Anti-aliasing filter)  
 anti-imaging (see Anti-imaging filter)  
 cascade 124

- digital 41, 127–32, 145, 172, 191  
FIR 130, 145  
ideal 123  
impulse response 121, 127, 129, 131  
low-pass (see Anti-aliasing filter and  
    Anti-imaging filter)  
transversal 128–32, 145, 191 (see  
    also Oversampling)  
Flag 71–73, 144, 304  
Floating point 44  
FMV (see Full motion video)  
Foldover (see Aliasing)  
Fonts 268  
Forms (CD-I) 223  
Foucault focusing 108, 117  
Fourier theorem 20  
Fourier transform 263  
Four quadrant photodiode 111  
Frame  
    CD 74, 82, 90, 94, 105  
    video 282  
Frequency  
    audio 19, 123, 184  
    bandlimiting (see Bandlimiting)  
    response 19, 180  
Full motion video 231–33, 236
- Gibbs phenomenon 191  
Glass master 286  
Glitching 141  
Granulation noise 32, 281  
Graphics  
    CD-I 230  
    CD+G/M 267–73  
*Green Book* 221  
Group delay 125
- Hamming codes 63  
Hardware address sectoring 259  
Harmonics 20, 25  
Header Field (CD-ROM) 215  
Hearing threshold 264  
HEM (holographic elements) 298  
HEPA (High Efficiency Particulate  
    Air) 294, 310  
HF signal (see RF signal)  
High Sierra Group 220, 224, 235, 237  
Host computer (CD-ROM) 217, 219  
Huffman code 233, 239  
Hygroscopic material (polycarbonate)  
    295
- IBM PC 245, 253, 272  
IEC (International Electrotechnical  
    Commission) 48, 158, 233, 241  
IEC serial format 159–63  
Impulse response 121, 127, 129, 131  
Index numbers 96  
Index of reflectivity 249  
Index of refraction 54, 110, 297, 312  
Information 1–4  
Information capacity 213–15, 222,  
    226, 231  
Initialization 51, 94, 113, 249, 252,  
    254
- Injection molding  
    description 205, 295–96  
    injection-compression 296  
    machine 296  
    performance 295
- Input filter (see Anti-aliasing filter)
- Input/Output 157–63
- Integrated circuits 4, 142–46, 236, 244
- Interactivity 234 (see also CD-I)
- Interference 170, 201
- Interleaving  
    error correction 69–73  
    interactivity 234
- Intermodulation distortion 122, 171,  
    184
- Interpolation 65, 73, 144, 304–6
- ISO (International Standards Organiza-  
    tion) 98, 220, 233, 234, 241,  
    272
- ISRC (International Standard Record-  
    ing Code) 92, 96, 98, 282
- Jewel box 309  
Jitter 41, 88, 168
- Kerr effect 257
- Label printing 301  
Land 54, 80, 111, 113–15, 117–18
- Laser  
    diode 108, 117, 174, 201  
    light 54, 86, 108–9, 203, 249, 256,  
        261  
    pickup 86, 108–18, 201
- Laservision 246, 248–49
- Lead-in/Lead-out 92, 94–96, 249, 281
- Least significant bit 29, 133

- Lens  
 collimator 109, 117  
 cylindrical 111  
 objective 107, 109
- Linear PCM (see Pulse code modulation)
- Lissajous pattern 189
- Longitudinal recording 256
- Low-bit conversion  
 conversion 147–57  
 noise shaping 148  
 pulse density modulation 152  
 pulse width modulation 157
- Low-level linearity 137–40, 187–89
- Low-pass filter (see Anti-aliasing filter and Anti-imaging filter)
- LP record 4–7, 54, 119, 277
- Magnetic storage 256
- Magneto-optical recording  
 disc (MOD) 256–58  
 mini disc 261–66  
 operation 255–57  
 professional model 260  
 system 258–59
- Maréchal criterion 58
- Masking 263
- Master  
 disc 286–93  
 tape 278–86
- Mastering 278–93
- Matrixing (see Electroforming)
- Menu 281
- Merging bits (see Bits)
- Metallization 50, 205, 299–301 (see also Electroforming)
- Microphonics 176
- Microprocessor 105, 119, 235, 237
- MIDI 243, 251, 260, 269
- Mini Disc  
 ATRAC 263–65  
 data reduction 263  
 masking 263  
 psychoacoustics 263  
 system 261–66
- MOD (see Magneto-optical recording, disc)
- Modes  
 CD+G/M 268–70  
 CD-ROM 216  
 interface 159  
 subcode 92–98
- Modulation  
 adaptive delta pulse code 228–30  
 amplitude 36  
 definition 36  
 delta 227  
 eight-to-fourteen (see Eight-to-fourteen modulation)  
 frequency 36  
 pulse amplitude modulation 36, 121  
 pulse code modulation 39–44, 74, 225, 280  
 pulse density modulation 153, 154  
 pulse number modulation 36  
 pulse width modulation 36, 155, 157  
 sampling 23
- Molding 295–99
- Mono-line 299, 310
- Most significant bit 29, 133, 137, 158, 189
- Mother 294
- MPEG (Motion Picture Expert Group) 226, 231–33, 236
- Multiplexing 38, 42, 90, 158, 174
- Mute 66, 90, 186, 282
- NA (see Numerical aperture)
- Noise  
 analog 3, 32  
 quantization 32, 184
- Noise shaping  
 second-order 152  
 theory 148  
 third-order 155
- Non-photoresist 291
- Non-return-to-zero 80, 87
- Non-return-to-zero-inverted 80, 87
- NTSC (National Television Standards Committee) 22, 230, 233, 237, 242, 246, 247, 248, 280
- Numerical aperture 54, 110
- Nyquist 9, 16
- Nyquist frequency 16, 20, 121, 123
- Nyquist theorem 16–22
- Object code 222
- Objective lens  
 cleaning 202  
 function 54, 107, 111, 113
- Offset  
 DC 87  
 subcode 282
- Omni-player 235

- One-beam pickup 115–18  
One-bit conversion (see Low-bit conversion)  
Optical pickup  
  beam 54–60, 107–18, 198  
  comparison 168–69  
  disc signal 86–89, 107  
  improvements 168  
  laser 86, 108–18, 201  
  signal output 85, 111, 115, 117, 169,  
    257, 304  
Optical storage  
  disc 48, 107, 204–11  
  erasable 255–67  
  read-only 48  
  write-once 249–55  
Opto-isolator 170  
*Orange Book* 238, 241, 254  
Orange peel 205  
OS-9 237  
Output  
  audio 43, 105, 157, 201  
  digital 157–63  
  filter (see Anti-imaging filter)  
  subcode 76, 90–99  
Oversampling 44, 90, 127–32, 145,  
  198  
Pad printing 301  
PAL 22, 230, 233, 237, 242, 246, 247,  
  248  
PAM (see Modulation, pulse amplitude)  
Parity 62–64, 67–76, 158  
P channel (see Subcode)  
PCM (see Pulse code modulation)  
Period 80, 86  
Phase  
  lock 87  
  shift 39, 125  
  testing 189  
Phase change medium 249–66  
Photo CD  
  Huffman code 239  
  Image pac 239  
  *Orange Book* 238, 241  
  Photo-YCC 238  
  subsampling 239  
  system 237–42  
Photodiode 108, 111, 114, 115–17, 174  
Photolithographic process 312  
Photopolymerization 310  
Photoresist 286  
Photo-thermographic transfer 312  
Pickup (see Optical pickup)  
Piezoelectric transducer 292  
Pinholes 205, 300, 308  
Pitch error (frequency) 191  
Pitch (track) 51, 52, 107, 193, 289,  
  304  
Pits 6, 48, 51–60, 80, 205, 289, 292,  
  297  
Placebos 208–11  
Player  
  buying 195  
  connections 201  
  design 103–63, 167, 197  
  features 195  
  installation 200–202  
  overview 103  
  performance 49, 196, 198  
  repairs 203  
  specifications 49  
  testing 178–95  
PNM (see Modulation, pulse number)  
Polarization beam splitter 109  
Polarized light 109, 257, 297, 308  
Polycarbonate 48, 109, 204, 207, 295–  
  98, 304, 312  
Power supply 168  
Preamble (AES/EBU) 158  
Pre-emphasis (see Emphasis)  
Pre-groove (CD-R) 249  
Pre-mastering 218, 277  
Processor, digital audio 278  
Psychoacoustics 232, 263  
Pulse code modulation 39–44, 74, 225,  
  280  
Pulse density modulation 153, 154  
Pulse width modulation 36, 155, 157  
Pyramid delta encoding (DVI) 244  
QC (see Quality control)  
Q channel (see Subcode)  
QHY (Quantized high-resolution Y)  
  231  
Quadraphonics 92  
Quality control 204, 302–9  
Quantization  
  distortion 31  
  effect 15, 27–31  
  error 30, 136, 182, 227  
  noise 32, 131, 136, 184, 187, 228  
Quarter wave plate 109

- Radial tracking error 107, 304
- Radix 28
- Random access memory 89, 145, 237, 265
- Random bit error 64, 72, 218, 306
- Red Book* 48, 250, 254, 277, 302
- Reduction
  - CD+G/M 272
  - CD-I 224
  - DVI 243
  - masking 263–65
  - MPEG 231–33
  - Photo CD 238–41
  - use 22, 44
- Reed-Solomon Code 63, 67–69 (see also Cross Interleave Reed-Solomon Code)
- Refraction 54, 297–98, 302
- Refractive index (see Index of Refraction)
- RF signal 86, 169, 193, 306
- RGB 230, 238, 242
- Ripple 124, 145, 171
- Sample-and-hold circuit 41, 43, 140–42
- Sampling
  - discrete time 3, 15–22
  - frequency 16, 21, 22, 49, 280
  - rate converter 229
  - theorem 16–22
- Scale factors 229
- SCMS (Serial Copy Management System) 161, 265
- SCSI (Small Computer System Interface) 243
- SECAM 22, 230, 233, 237, 242, 246, 247, 248
- Sector 215
- S/E ratio (see Signal-to-error ratio)
- Serial transmission format (see AES/EBU digital I/O)
- Servo
  - focus 107, 111, 117
  - mechanism 103, 169
  - tracking 107
- Settling time 43, 141
- Shannon 9
- S/H circuit (see Sample-and-hold circuit)
- Signal
  - CD format decoding 85–90
  - CD format encoding 74–85
  - path in player 103–6
- Signal-to-error ratio 30
- Signal-to-noise ratio 138, 185
- Silk screening 301
- Single-beam pickup 115–18
- Sin ( $x$ )/ $x$  121, 127, 129, 131, 140
- Skew 86, 302
- Sled 115, 118, 169
- Slope overload 227
- SMPTE (Society of Motion Picture and Television Engineers) 251, 260, 278, 280, 285
- S/N ratio (see Signal-to-noise ratio)
- Son 294
- Sound
  - channel separation 180
  - characteristics 178
  - dynamic range 187
  - frequency response 180
  - intermodulation distortion 184
  - listening 199
  - low-level linearity 187–89
  - phase 189
  - pressure 3
  - quantization noise 184
  - recording 4, 15
  - signal-to-noise ratio 185
  - total harmonic distortion 181–83
  - waveforms 15, 17
- SPARS (Society of Professional Audio Recording Services) 207
- SPDIF serial format 159–63
- Specifications 178, 302
- Spectrum analysis 183, 195
- Spin coating 286, 301
- Spindle 105
- Spiral track 51, 107, 289
- Spurious noise sheet 285
- Sputtering 300
- Stamper 204, 293–96
- Standards 48, 158, 220, 221, 231, 233, 238, 302
- Static charge 204, 296
- Subband coding 232
- Subcode
  - bits 76, 99
  - block 76, 90–91, 144, 161
  - editing 253, 281
  - encoding 76, 92, 253
  - graphics 99, 267–73
  - P channel 92

- Q channel 92–99, 215  
synchronization 79, 91
- Subheader (CD-I) 223
- Substrate 51, 54, 111, 204, 295–98,  
304, 310
- Superposition 121
- Symbol 76, 218
- Synchronization word  
AES/EBU format 158, 159  
CD frame 74, 82, 88, 91  
CD-ROM 215  
subcode 79, 91
- Synchronous clocks 171
- Syndrome 68
- Table of contents (disc) 94, 249, 254,  
285
- Tape master 277–80
- Tape preparation 280
- Television 230–34, 237–48
- Test discs 178, 179, 196
- Test equipment 180
- Test frequencies 184, 186
- THD (see Total harmonic distortion)
- Three-beam pickup 108–15
- Threshold of hearing 264
- Time  
delays 125, 127  
sampling, discrete 15–22
- Timebase correction 89
- Timecode 92, 159, 251, 260, 278, 280,  
285
- Timing 96, 105
- TOC (see Table of contents)
- Total harmonic distortion 181–83
- Track eccentricity 107, 304
- Tracking error 113, 117, 192, 196
- Track numbers 92, 282
- Transfer rate 218, 224, 230, 232, 245
- Transport screw(s) 200, 203
- Transversal filter (see Filter and Over-sampling)
- Two-axis actuator 115
- Two's complement 161
- Ultraviolet light 301, 310, 313
- U-Matic tape recorder 278, 280, 282,  
286, 289
- UPC (universal product code) 96, 281
- User information 158, 250
- User table of contents (CD-R) 250
- Vacuum evaporation 299
- Vertical recording 256
- Vibration  
damping 175–77, 197, 290  
sensitivity 175, 197
- Video frames 282
- Warpage, disc 107, 111, 206, 208, 304
- Wavelength 52, 55, 56, 108, 297, 298
- Wet silvering 300
- Word 29, 74
- Word length 30
- WORM 249 (see also CD-R)
- Wow and flutter 7, 43, 89, 191
- Write-once medium 249 (see also  
CD-R)
- Yellow Book* 220
- YUV 230
- Zero crossing 86





A standard linear barcode is positioned vertically on the right side of the page. It consists of vertical black bars of varying widths on a white background.

9 780198 163275