A Digital Audio Primer

Many people don’t care about the technology behind their stereo system. As long as it sounds good and they can press a button and listen to music, everything is fine. However, when you start working with audio on computers and the Internet, it’s important to understand a few key principles to achieve good results.

What is Sound?

Sound reaches our ears as waves of rapidly varying air pressure caused by a vibrating object, such as a guitar string. As the string moves in one direction, it pushes on nearby air molecules, causing them to move closer together. This creates a small region of high pressure on one side of the string and low pressure on the opposite side. As the string moves in the opposite direction, the areas of high and low pressure reverse.

Sound waves occur as these repeating cycles of higher and lower pressure move out and away from the vibrating object. The frequency (pitch) of a sound is the number of times per second that these cycles occur. The amplitude (intensity) of sound is the size of the variations.
Measuring Sound

Our ears respond to sound logarithmically. As a sound gets louder, increasingly larger changes in sound intensity must occur for us to perceive the same amount of change in loudness.

Decibels

The term decibel (dB) means one-tenth of a Bel—named after Alexander Graham Bell. (This is why the B in dB is capitalized). A Bel is the base 10 logarithm of the ratio between the power level of two sounds or signals.

Sound Pressure Level

The intensity of sound is called the sound pressure level (SPL) and is measured in decibels (dB SPL). Decibels are a logarithmic scale that represents how much a sound level or audio signal varies from another signal, or reference level. You might refer to a sound as being 10dB louder than another sound or 3dB softer. A 3dB change is about the minimum change in sound level that most of us can perceive. A 10dB change sounds about twice as loud.

Decibels are always relative. To use decibels to represent a specific quantity, you need to know the reference, or 0 dB level. In the case of sound intensity, 0 dB SPL represents the threshold of hearing of a young undamaged ear (a pressure of about 3 billionths of a pound per square
In this case, all sound pressure levels are positive numbers that show how much louder a sound is than the threshold of hearing.

**Figure 17 - Relative Loudness of Common Sounds**

<table>
<thead>
<tr>
<th>Relative Level</th>
<th>SPL</th>
<th>Sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>10,000,000x</td>
<td>140</td>
<td>Colt 45 Pistol (25 ft)</td>
</tr>
<tr>
<td></td>
<td>130</td>
<td>Fire Engine Siren (100 ft)</td>
</tr>
<tr>
<td>1,000,000x</td>
<td>120</td>
<td>Jet Takeoff (200 ft)</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>Rock Concert (10 ft)</td>
</tr>
<tr>
<td>100,000x</td>
<td>100</td>
<td>Loud Classical Music</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>Heavy Street Traffic (5 ft)</td>
</tr>
<tr>
<td>10,000x</td>
<td>80</td>
<td>Cabin of Cruising Jet Aircraft</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>Average Conversation (3 ft)</td>
</tr>
<tr>
<td>1,000x</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>Average Suburban Home (night)</td>
</tr>
<tr>
<td>100x</td>
<td>40</td>
<td>Quiet Auditorium</td>
</tr>
<tr>
<td>10x</td>
<td>20</td>
<td>Quiet Whisper (5 ft)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Rustling Leaves</td>
</tr>
</tbody>
</table>

**Loudness**

Loudness is subjectively how we perceive different sound intensities. The sound intensity of a jet taking off 200 feet away is about 120dB SPL, or a million times more intense than the threshold of hearing. The sound intensity of rustling leaves is about 20dB SPL, or 10 times higher than the threshold of hearing. The sound of the jet is 100,000 times more intense than the rustling leaves (100dB). We actually perceive the jet to be about 1000 times louder than rustling leaves rather than 100,000 times louder.
Frequency

Figure 18 - Octave Intervals and Frequencies for Musical Notes

The frequency of a sound is measured in Hertz (Hz), which means cycles per second. A kilohertz (kHz) is a thousand cycles per second. We perceive pitch exponentially. A unit of pitch all musicians are familiar with is the octave. An octave is the interval between any note and the next higher note with the same name. Notes that are one octave apart sound similar, but one is twice the frequency of the other. For example, the note A below middle C is at a frequency of 220Hz, the note A above middle C is at 440Hz, and the next higher A is at 880Hz.

Analog Audio

The term analog means something that is similar in function or position. The varying voltage produced by a microphone is analogous to the pressure variations of a sound wave. On a cassette tape, variations in magnetic flux in a metal coating on the tape represent pressure variations in the sound wave. On vinyl records, variations in the width of the groove correspond to the pressure variations. The position along the groove or tape corresponds to time.

In an analog audio system, voltages represent sound pressures. These signals are amplified from the millivolt level (1000th of a volt) produced by microphones, playback heads and phono cartridges by about 1000 times (60dB) to the levels found inside stereo preamps. A power amp boosts the voltage level from the preamp to a loudspeaker, which creates sound waves in the air by vibrating rapidly in response to the audio signal.
**Digital Audio**

In digital audio, the representation of the audio signal is no longer directly analogous to the sound wave. Instead, the value of the signal is sampled at regular intervals by an analog-to-digital (A/D) converter (or ADC), which produces numbers (digits) that represent the value of each sample. This stream of numbers represents a digital audio signal, which can be stored as a computer file and transmitted across a network.

In order to listen to a digital audio signal, it must be converted to analog by a digital-to-analog (D/A) converter (or DAC). In most home stereo systems, the D/A conversion takes place inside the CD player. Computer sound cards, MiniDisc recorders and DATs have both A/D converters (for recording) and D/A converters (for playback). Many home systems have a combination of digital and analog components, but all audio systems end with analog signals at the speakers or headphones.

**Sampling**

To convert an analog signal to a digital format, the voltage is sampled at regular intervals, thousands of times per second. The value of each sample is rounded to the nearest integer on a scale that varies according to the resolution of the signal. The integers are then converted to binary numbers.

*Figure 19 - Sampling and Converting a Waveform to PCM*

The sampling rate is how many times per second the voltage of the analog signal is measured. CD audio is sampled at a rate of 44,100 times per second (44.1 kHz). DAT (Digital Audio Tape) supports sampling rates of 32, 44.1 and 48 kHz. Other commonly used sampling rates are 22.05 kHz and 11.025 kHz.
The sampling rate must be at least twice as high as the highest frequency to be reproduced. The range of human hearing is roughly from 20 to 20,000 Hz, so a sampling rate of at least 40 kHz is needed to reproduce the full range.

Higher sampling rates allow the use of filters with a more gradual roll-off. This reduces phase shift, which can affect the stereo image at higher frequencies.

\textsuperscript{11} According to the Nyquist Theorem
The 44.1 kHz sampling rate for CDs was chosen to allow headroom for filters and other types of signal processing. MPEG AAC and DVD Audio support rates up to 96 kHz.

**Resolution**

The resolution of a digital signal is the range of numbers that can be assigned to each sample. CD audio uses 16 bits, which provides a range of binary values from 0 to 65,534 ($2^{16}$). The binary value of 0000000000000000 (zero) corresponds to -32,768 (the lowest possible level), and the value 1111111111111111 (65,535) corresponds to 32,767 (the highest possible level). Higher resolution increases the dynamic range and reduces quantization distortion and background noise.

**Quantization**

Quantization is the process of selecting whole numbers to represent the voltage level of each sample. The A/D converter must select a whole number that is closest to the signal level at the instant it’s sampled. This produces small rounding errors that cause distortion.

Quantization distortion increases at lower levels because the signal is using a smaller portion of the available dynamic range, so any errors are a greater percentage of the signal. A key advantage of audio encoding schemes, such as MP3, is that more bits can be allocated to low-level signals to reduce quantization errors.

**Dithering**

![Figure 21 - Quantization Errors](image)

A process called dithering introduces random noise into the signal to spread out the effects of quantization distortion and make it less
noticeable. Some audiophiles don’t like the notion of noise that is deliberately added to a signal, but the advantages of digital audio are so great that the end result is still better than most analog systems.

**Clipping**

Figure 22 - Clipping

Levels in a digital audio signal are usually expressed in dB, measured by their relationship to 0 dB, the highest possible level. One of the rules of digital audio is that a signal can never exceed 0 dB. If the level of a signal is raised too much, the peaks will be clipped at the 0 dB level. Clipping causes extreme distortion and should be avoided at all costs.

**Bit-rates**

The term “bit-rate” refers to how many bits (1s and 0s) are used each second to represent the signal. The bit-rate for digital audio is expressed in thousands of bits per second (kbps) and correlates directly to the file size and sound quality. Lower bit-rates result in smaller file sizes but poorer sound quality, and higher bit-rates result in better quality but larger files.

The bit-rate of uncompressed audio can be calculated by multiplying the sampling rate by the resolution (8-bit, 16-bit, etc.) and the number of channels. For example, CD Audio (or a WAV file extracted from a CD) has a sampling rate of 44,100 times per second, a resolution of 16 bits and two channels. The bit-rate would be approximately 1.4 million bits per second (1,411 kbps).

<table>
<thead>
<tr>
<th>Table Error! Bookmark not defined.</th>
<th>Calculating Bit-rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling Rate</td>
<td>x</td>
</tr>
<tr>
<td>44,100</td>
<td>x</td>
</tr>
</tbody>
</table>
Dynamic Range

Dynamic range is the range of the lowest to the highest level that can be reproduced by a system. Digital audio at 16-bit resolution has a theoretical dynamic range of 96 dB, but the actual dynamic range is usually lower because of overhead from filters. The dynamic range of vinyl records and cassette tapes is much lower than CDs and varies depending on the quality of the recording and playback equipment. The dynamic range of cassette tapes also varies depending on the type of tape.

Signal-to-noise Ratio

The signal-to-noise ratio is the ratio of the background noise (hiss, hum and static) level to the highest level that can be reproduced. Each additional bit of resolution corresponds to an increase of 6 dB in signal-to-noise ratio. Audio CDs achieve about a 90 dB signal-to-noise ratio.

Encoding

Encoding is the process of converting uncompressed digital audio to a compressed format such as MP3. The algorithm used in the encoding (and decoding) software is referred to as a codec—as in coding/decoding. There is often more than one codec for a particular format, and different codecs can vary widely in quality and speed, even for the same format.
Advantages of Digital Audio

For years, audiophiles and engineers have debated the merits of digital audio versus high-end analog systems, and to this day, there are audiophiles who swear by their analog systems. Digital audio has emerged as the winner by most accounts, but it’s still useful to understand the advantages of digital versus analog audio, because many audio systems contain a mix of digital and analog components.

The advantages of digital audio can be summed up as follows: wider dynamic range, increased resistance to noise, better copyability and the ability to use error correction to compensate for wear and tear. Many types of digital media, such as CDs and MiniDiscs, are also more durable than common analog media, such as vinyl records and cassette tapes.

Wider Dynamic Range

Digital audio at 16 bits theoretically can achieve a dynamic range of 96 dB, compared to less than 80 dB for the best analog systems. This is especially important for classical music where levels within the same composition can range from the relative quiet of a flute solo to the loudness of dozens of instruments playing simultaneously.

Increased Resistance to Noise

In analog systems, crackling noise and hum from electromagnetic frequency (EMF) interference is picked up along the way as the signal passes through analog circuits. Background hiss is also generated by thermal noise from analog components. Digital signals are virtually immune to picking up these types of noise, although any noise that enters the signal before it’s converted to digital will be reproduced along with the rest of the signal.

Better Copyability

Digital audio can be copied from one digital device to another without any loss of information, unlike analog recording, where information is lost and noise introduced with every copy. Even the best analog systems lose about 3dB of signal-to-noise ratio when a copy is recorded. After several generations of analog copies, the sound quality will deteriorate noticeably. With digital audio, unlimited generations of perfect copies can be made.
This ability to make perfect copies is one reason why the RIAA has gone to so much trouble to introduce the Serial Copy Management System (SCMS) for consumer audio equipment, and why they are so concerned about the proliferation of MP3 files. SCMS prevents multiple generations of copies (copies of copies) from an original and is required by the Audio Home Recording Act of 1992 to be used on all consumer digital audio recording devices sold in the United States. Currently, there is no way to prevent multiple generations of perfect copies from a single MP3 file.

Digital copies can also be made much faster than analog copies, which usually must be made in real time. For example, with an analog device like a cassette deck, it always takes at least 60 minutes to record 60 minutes of music from a CD. With digital audio, the same 60 minutes of music can be copied to a hard disk in as little as 5 minutes on a system with a fast CD-ROM drive.

Of course, if you are making an original recording with digital equipment, it will take the same amount of time as with analog equipment. (Uncle Jack playing the kazoo for half an hour still takes half an hour to record). But once a digital recording is on your PC, you can make a digital copy in a fraction of the time it would take to record a copy with analog equipment.

**Error Correction**

Most digital audio media, such as CDs and DATs, have built-in error correction. On an audio CD, approximately 25% of the disc is used for error correction data. If a bad scratch causes an error that can’t be corrected, the player will attempt to reconstruct the missing data by interpolation.

**Durability**

Digital media such as CDs and MiniDiscs are much more durable than any analog media. This improved durability is one of the main reasons people were so eager to migrate from vinyl records to CDs.

Each time you play a record or tape, microscopic bits of vinyl or oxide are scraped away, adding to the cumulative wear. Vinyl records are particularly prone to warping and scratching, and tapes gradually become demagnetized. A CD or MiniDisc can be played hundreds of
times, with no loss of quality, as long as there is not excessive physical damage.

Both analog and digital tapes can suffer degradation from magnetic fields, but some popular digital formats like DAT are much more durable than analog tapes (especially cassettes) because the tape is stronger and the oxide coating is thicker.

**File Size and Bandwidth**

Digital audio can create large files that quickly use up hard disk capacity and require a tremendous amount of bandwidth to transmit over a network. Network bandwidth is like a pipe that carries a stream of bits. The size of the pipe imposes a limit on how many bits can be moved in a given time period. Multiple users competing for the same bandwidth limit the amount of bandwidth available to any one user.

File sizes and bandwidth requirements for uncompressed audio can be calculated by multiplying the sampling rate by the resolution, the number of channels and the time in seconds. The bit-rate has a direct relation to the file size—if you do something that changes the bit-rate, the file size will change proportionally. The bandwidth requirement of a digital audio signal is the same as the bit-rate. This is true whether the signal is compressed or not. Table 9 shows the formula for calculating file sizes for uncompressed audio.

<table>
<thead>
<tr>
<th>Sampling Rate</th>
<th>x</th>
<th>Resolution</th>
<th>x</th>
<th>Number of Channels</th>
<th>x</th>
<th>Time in Seconds</th>
<th>/</th>
<th>Bits / Byte</th>
<th>=</th>
<th>File Size (in Bytes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>44,100</td>
<td>x</td>
<td>16</td>
<td>x</td>
<td>2</td>
<td>x</td>
<td>60</td>
<td>/</td>
<td>8</td>
<td></td>
<td>10,584,000</td>
</tr>
</tbody>
</table>

You can do several things to control the size of digital audio files, but there will always be a trade-off between file size and sound quality. Lowering the sampling rate will produce a smaller file, but will also lower the maximum frequency response. Lowering the resolution produces a smaller file but reduces the accuracy and allows more noise and distortion to be introduced due to increased quantization errors. A mono signal, used in place of stereo, will cut the size in half (uncompressed audio only).

Table 10 shows how different combinations of sampling rates, resolution and numbers of channels can be used to control file sizes.
### Table

<table>
<thead>
<tr>
<th>Sampling Rate</th>
<th>Resolution</th>
<th>Number of Channels</th>
<th>Bit-rate</th>
<th>File Size (in Bytes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>44,100</td>
<td>16</td>
<td>2</td>
<td>1,411,200</td>
<td>10,584,000</td>
</tr>
<tr>
<td>44,100</td>
<td>16</td>
<td>1</td>
<td>705,600</td>
<td>5,292,000</td>
</tr>
<tr>
<td>22,050</td>
<td>16</td>
<td>1</td>
<td>352,800</td>
<td>2,646,000</td>
</tr>
<tr>
<td>11.025</td>
<td>16</td>
<td>1</td>
<td>176,400</td>
<td>1,323,000</td>
</tr>
<tr>
<td>11.025</td>
<td>8</td>
<td>1</td>
<td>88,200</td>
<td>616,000</td>
</tr>
</tbody>
</table>

### Compression

Limited network bandwidth and hard disk capacity have been major driving factors behind the development of compressed audio formats. Until recently, only a small number of people used their computers to store CD-quality music. A few people would copy their favorite songs from a music CD and use a CD-Recordable drive to create a compilation CD, similar to the way many people make cassette tapes from prerecorded music.

Audio and electronics engineers have been working to solve the bandwidth bottleneck ever since networks were invented. They work on both sides of the problems by increasing bandwidth (larger pipe) and compressing data (higher pressure). High speed Internet connections such as cable modems and ASDL have been developed to increase the size of the pipe, and compression schemes such as JPEG and MPEG have been developed to squeeze more data through it.

MP3 provides relief by compressing files up to approximately 10=1 without significant loss of quality. Four minutes of CD audio (44.1 kHz 16-bit stereo) requires about 40MB of disk space and would take more than 3-½ hours to download with a 28.8 kbps modem. At this rate, a 2GB hard disk would hold about 50 four-minute songs.

With MP3 encoded at 128 kbps, each four-minute song would take up less than 4MB of space and could be downloaded in less than 20 minutes with a 28.8 kbps modem. A 2GB hard disk could now hold more than 500 songs. This much compression, coupled with the larger and cheaper hard disks that are now available, makes it possible to use a PC as a high-capacity, CD-quality jukebox in place of tape decks, turntables and CD players.
**Table** Error! Bookmark not defined. - **Typical Download Times** for Four-minute Songs

<table>
<thead>
<tr>
<th>Format</th>
<th>28.8 k Modem</th>
<th>56 k Modem</th>
<th>Dual ISDN 128 kbps</th>
<th>Cable 1.5 Mbps</th>
<th>T1 Line 1.5 Mbps</th>
<th>ADSL 500 kbps+</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD Audio</td>
<td>3.6 hrs</td>
<td>2 hrs</td>
<td>44 min</td>
<td>4 min</td>
<td>4 min</td>
<td>7 min</td>
</tr>
<tr>
<td>MP3 at 128 kbps</td>
<td>19.7 min</td>
<td>9 min</td>
<td>4 min</td>
<td>20 sec</td>
<td>20 sec</td>
<td>39 sec</td>
</tr>
</tbody>
</table>

* Actual speed will usually be less.

Newer generations of MPEG Audio, such as AAC (Advanced Audio Coding), offer even higher levels of compression and better sound quality but have not yet reached the consumer market because of high licensing costs.

**Lossy vs. Lossless Compression**

**Figure 24 - Typical MP3 Compression**

![Figure 24 - Typical MP3 Compression](image)

There are two basic categories of compression: lossless and lossy. Lossless compression works by encoding repetitive pieces of information with symbols and equations that take up less space but provide all the information needed to reconstruct an exact copy of the original. Lossy compression works by discarding unnecessary and redundant information (sounds that most people can’t hear) and then applying lossless compression techniques for further size reduction.
There is an ongoing debate among audiophiles about the merits of lossless versus lossy compression. With lossless compression, there is never a loss of fidelity (unless an error gets introduced during the process)—there is no debate about that. With lossy compression (such as MPEG Audio), there is always some loss of fidelity that becomes more noticeable as the compression ratio is increased. The goal then becomes producing sound where the losses are not noticeable, or noticeable but not annoying.

The highest compression ratio for lossless audio is about 2 to 1, but the quality will always be indistinguishable from the original. With lossy compression, the quality will vary according to factors such as the bit-rate, the complexity of the music and the quality of the encoding software. Some forms of lossy compression, such as MPEG AAC, can achieve compression ratios of up to 11 to 1, with quality indistinguishable from the original. Numerous controlled tests with trained listeners have verified this.

Even with the best lossy formats, a few people with very sensitive ears may be able to tell the difference between the original and encoded file when listening to critical material (complex music) on expensive hi-fi systems. Most people will not be able to detect any differences at the higher bit-rates, but a few people will always feel like they are being cheated when they know something has been taken away (even if they can’t tell the difference).

**Digital Audio Formats**

Digital audio comes in many different formats, and multiple formats will be a fact of life for the foreseeable future. Groups like MPEG have created open standards, but even formats based on the same MPEG
standard may not compatible with each other because of proprietary components.

Fortunately for consumers, many hardware and software players are able to support multiple formats—so if you purchase digital music in any of the major formats (MP3, WMA, etc.) you will be in a good shape. If a format does become obsolete, plenty of tools are available for converting digital audio to different formats.

**Digital Audio Files**

An audio file has two main parts: a header and the audio data. The header is used to store information about the file, including the resolution, sampling rate and type of compression. Often a “wrapper” is used to add features, such as license management information or streaming capability, to a digital audio file.

![Figure 25 - Digital Audio File Structure](image)

The format of a digital audio file refers to the type of audio data within the file. The file type refers to the structure of the data within the file. It is common for the same format to be used by more than one file type. For example, the PCM format is found in both WAV and AIFF files.
**Table** Error! Bookmark not defined. - **Common Digital Audio Formats**

<table>
<thead>
<tr>
<th>Type</th>
<th>Extensions</th>
<th>Codec</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIFF (Mac)</td>
<td>.aif, .aiff</td>
<td>*PCM</td>
</tr>
<tr>
<td>AU (Sun/Next)</td>
<td>.au</td>
<td>*u-law</td>
</tr>
<tr>
<td>CD audio (CDDA)</td>
<td>N/A</td>
<td>PCM</td>
</tr>
<tr>
<td>MP3</td>
<td>.mp3</td>
<td>MPEG Audio Layer-III</td>
</tr>
<tr>
<td>Windows Media Audio</td>
<td>.wma</td>
<td>Proprietary (Microsoft)</td>
</tr>
<tr>
<td>QuickTime</td>
<td>.qt</td>
<td>Proprietary (Apple Computer)</td>
</tr>
<tr>
<td>RealAudio</td>
<td>.ra, ram</td>
<td>Proprietary (Real Networks)</td>
</tr>
<tr>
<td>WAV</td>
<td>.wav</td>
<td>*PCM</td>
</tr>
</tbody>
</table>

* Can be used with other codecs.

**WAV**

WAV is the default format for digital audio on Windows PCs. WAV files are usually coded in PCM format, which means they are uncompressed and take up a lot of space. WAV files can also be coded in other formats, including MP3.

**AIFF and AU**

AIFF is the default audio format for the Macintosh, and AU is the default format for SUN systems. Both of these formats are supported on most other platforms and by most audio applications. Each of these formats can be compressed, but compression sometimes creates compatibility problems with other platforms.

**Streaming Audio**

Streaming audio avoids many of the problems of large audio files. Instead of having to wait for the entire file to download, you can listen to the sound as the data arrives at your computer.

Streaming audio players store several seconds worth of data in a buffer before beginning playback. The buffer absorbs the bursts of data as they are delivered by the Internet and releases it at a constant rate for smooth playback.

Many digital audio formats can be streamed by wrapping them in a streaming format, such as Microsoft’s ASF (Active Streaming Format), which can be used to stream MS Audio, MP3 and other formats.
Table  Error! Bookmark not defined. - Streaming Audio Systems

<table>
<thead>
<tr>
<th>Type</th>
<th>Primary Format</th>
<th>Developer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows Media Technologies</td>
<td>Windows Media Audio / Active Streaming Format (ASF)</td>
<td>Microsoft</td>
</tr>
<tr>
<td>Icecast (open source)</td>
<td>MP3</td>
<td>The Icecast Team</td>
</tr>
<tr>
<td>QuickTime</td>
<td>QuickTime</td>
<td>Apple Computer</td>
</tr>
<tr>
<td>RealSystem</td>
<td>RealAudio</td>
<td>RealNetworks</td>
</tr>
<tr>
<td>SHOUTcast</td>
<td>MP3</td>
<td>Nullsoft</td>
</tr>
</tbody>
</table>

Standard Formats

Standard formats make it easier for software developers and equipment manufacturers to produce products that are less costly and more compatible with each other. The compatibility provided by standard formats helps assure consumers that their music and equipment won’t become obsolete. Cassette tapes, compact discs and PCM are examples of standard audio formats that benefit both consumers and manufacturers.

PCM

PCM (Pulse Code Modulation) is a common method of storing and transmitting uncompressed digital audio. Since it is a generic format, it can be read by most audio applications—similar to the way a plain text file can be read by any word-processing program. PCM is used by Audio CDs and digital audio tapes (DATs). PCM is also a very common format for AIFF and WAV files.

![Figure 26 - Pulse Code Modulation](image)

PCM is a straight representation of the binary digits (1s and 0s) of sample values. When PCM audio is transmitted, each “1” is represented by a positive voltage pulse and each “0” is represented by the absence of a pulse. Figure 26 shows how binary data is converted to a PCM signal.
DPCM
DPCM (Differential Pulse Code Modulation) is a simple form of lossy compression that stores only the difference between consecutive samples. DCPM uses 4 bits to store the difference, regardless of the resolution of the original file. With DCPM, an 8-bit file would be compressed $2=1$, and a 16-bit file would be compressed $4=1$.

ADPCM
ADPCM (Adaptive Differential Pulse Code Modulation) is similar to DCPM except that the number of bits used to store the difference between samples is varied depending on the complexity of the signal. ADPCM works by analyzing a succession of samples and predicting the value of the next sample. It then stores the difference between the calculated value and the actual value.

u-law Compression
u-law (pronounced “mew-law”) is a common lossy compression scheme, similar to ADPCM, which can be used on AU, AIFF and WAV files.

MPEG Audio
MPEG Audio is a family of open standards for compressed audio that includes MP2, MP3 and AAC. (See Chapter 13 for more detailed information on MPEG Audio.)

MPEG-Based Proprietary Formats
Several proprietary formats are based on MPEG audio. Some of these are used in special applications, such as voice mail systems, high definition TV and satellite radio. Others compete directly with MP3 and are based on AAC or MP3, with proprietary wrappers. The sound quality of some of these is very good, but their proprietary nature makes them incompatible with many programs and portable players.

a2b
AT&T’s a2b music is a sophisticated music distribution system with many features, such as watermarking and encryption, to support copyright protection and royalty tracking. It is based on the MPEG-2 AAC Low Complexity Profile. The Policy Maker feature of a2b is a flexible electronic licensing system, which can control how music is
used and distributed. Music encoded with a2b can include artwork, credits, lyrics and links to the artist’s Web site.

**Proprietary Formats**

Even though MPEG Audio formats are based on open standards and widely used, many companies continue to develop proprietary digital audio formats. Proprietary formats can generate enormous profits for the developer if the format becomes a de facto standard. The marketplace, which tends to favor open standards like MPEG, will ultimately decide which formats will prevail.

**MP4**

Global Music Outlet uses the term MP4 to describe its proprietary music delivery system. It’s based on an enhanced version of MPEG AAC and includes an embedded player (each song is an .EXE file). Album graphics and links to the artist’s Web site can be embedded in the file.

**Liquid Audio**

Liquid Audio is a sophisticated music distribution system based on Dolby Digital and MPEG AAC. It supports both downloadable and streaming audio and uses watermarking and encryption for copyright protection. Music encoded with Liquid Audio can include artwork, lyrics, notes and pricing, along with links to a Web site where the song or album can be purchased. Liquid Audio has a playlist feature and allows you to burn songs to a CD if you have a supported CD-R drive.

**Apple QuickTime**

QuickTime is a widely used multimedia format from Apple Computer that supports both streaming audio and streaming video. Much of the MPEG-4 standard is based on QuickTime, and it is widely used for streaming video on the Web.

**Non-MPEG Proprietary Formats**

Several digital audio formats exist that are entirely proprietary. Many of these are quite good and are widely used.
Dolby Digital (Formerly AC-3)
Dolby Digital is a very high quality audio encoding and noise reduction system that is the audio component of High Definition Television (HDTV) and digital broadcast TV (DTV). It is also used in DVDs, laser discs, digital cable and direct broadcast satellite (DBS) systems.

EPAC
EPAC is a perceptual audio encoding scheme based on PAC—developed by Bell Labs, the research and development arm of Lucent Technologies. EPAC is reported to produce quality indistinguishable from the original CD at 128 kbps. However, I participated in one listening test where the audience was able to consistently tell the difference between original CD tracks and the same tracks encoded in EPAC at 160 kbps.

Windows Media Audio
Microsoft’s Windows Media Audio (WMA) format is a relatively late entry into the field of proprietary audio formats. WMA performs very good at lower bit-rates and is reported to produce quality indistinguishable from the original CD at 128 kbps. WMA is supported by most full-featured player programs and by many portable players. WMA is royalty-free when incorporated into software that runs on the Windows platform.

RealAudio
RealAudio was the first widely used system for streaming audio and video over the Internet. It is a proprietary format, but it is used by many online music stores for sample clips of songs. The RealPlayer also provides support for MP3.

TAC
TAC (Transparent Audio Compression) is a high-quality perceptual encoding scheme developed by K+K Research. TAC uses Adaptive Bit-rate Management (ABM), which is similar to VBR (variable bit-rate) encoding. TAC was developed as part of K+K Research’s MP02 (Music Publisher 02) software.

TwinVQ (VQF)
TwinVQ (Transform-domain Weighted Interleave Vector Quantization) is an encoding scheme developed by the NTT Human Interface Lab in Japan. TwinVQ is reported to provide higher quality than MP3, but encoding times are reported to be much longer, and CPU utilization is reported to be higher during playback.
Additional Resources

<table>
<thead>
<tr>
<th>Organization</th>
<th>Web Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio Engineering Society</td>
<td><a href="http://www.aes.org">www.aes.org</a></td>
</tr>
<tr>
<td>Internet Sound Institute</td>
<td><a href="http://www.soundinstitute.com">www.soundinstitute.com</a></td>
</tr>
<tr>
<td>MIT Media Lab – Machine Listening Group</td>
<td><a href="http://sound.media.mit.edu">http://sound.media.mit.edu</a></td>
</tr>
</tbody>
</table>

MPEG Audio

MPEG Standards

Organizations from all over the world are involved in developing MPEG standards. Fraunhofer-Gesellschaft of Germany and Thomson Multimedia of the United States provided key technology related to MPEG Audio Layer-III (MP3). Dolby Labs was heavily involved in the development of MPEG AAC. Each of these organizations holds multiple patents related to the technologies they contributed.

The MPEG committee works in phases and meets several times a year. To date, MPEG has released three families of standards: MPEG-1, MPEG-2 and MPEG-4. (If you are wondering about MPEG-3, it was merged into MPEG-2). All MPEG phases include standards for both audio and video. This chapter is concerned only with MPEG Audio.

It typically takes several years from when a standard is released to when consumer products that support it reach the market. MPEG-1, which includes MP3, was released in 1992. However, it took more than four years for software players, such as Winamp, to appear, and almost six years for the first portable MP3 players to become available.

MPEG standards for digital audio cover encoding of audio, either by itself or as the audio component of a multimedia file or stream. MPEG Audio is based on perceptual encoding techniques, which take advantage of the characteristics of human hearing and remove sounds that most people can’t hear.
MPEG-1

MPEG-1 (which includes MP3) was approved in November 1992. It works with bit-rates up to 1.5 mbps (million bits per second) and supports both mono and stereo audio, but not multi-channel surround sound. MPEG-1 supports sampling rates of 32, 44.1 and 48 kHz.

MPEG-2

MPEG-2 adds support for surround sound, lower sampling rates of 16, 22.05 and 24 kHz and bit-rates as low as 8 kbps. MPEG-2 can have up to five channels for surround sound and one low frequency enhancement channel for subwoofers. A multilingual extension adds support for up to seven more channels.

MPEG-4

MPEG-4 is intended to be an all-purpose encoding standard for multimedia systems of the future. It’s designed to handle applications ranging from simple voice systems that require very low bandwidth to high quality “audiophile” and professional sound systems. MPEG-4 can integrate synthetic and natural audio, including MIDI and text-to-speech systems. A large part of MPEG-4 is based on Apple’s QuickTime multimedia format.

MPEG-4 is made extensible by a language called MSDL (MPEG Syntax Description Language). The support for interactivity allows manipulation of the presentation of audio and visual data. MPEG-4 supports a wide range of storage and transmission media and will work over networks and wireless mobile connections.

MPEG-7

MPEG-7 is also referred to as Multimedia Content Description Interface. It defines a structure that supports searching, filtering and management of multimedia data. MPEG-7 is expected to be released in 2001.
Table  Error! Bookmark not defined. - MPEG Phases

<table>
<thead>
<tr>
<th>MPEG-1 (approved Nov. 1992)</th>
<th>Single (mono) and dual (stereo) channel encoding of audio at 32, 44.1 and 48 kHz sampling rates and bit-rates from 32 to 448 kbps.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPEG-2 (approved Nov. 1994)</td>
<td>A backwards compatible extension to MPEG1 with up to five channels, plus one low frequency enhancement channel. Adds support for 16, 22.05 and 24 kHz sampling rates for bit-rates between 32 to 256 kbps for Layer-I, up to 384 kbps for Layer-II, and from 8 to 320 kbps for Layer-III.</td>
</tr>
<tr>
<td>MPEG-2 AAC</td>
<td>Supports a wider range of sampling rates (from 8 kHz to 96 kHz) and up to 48 audio channels, plus up to 15 auxiliary low-frequency enhancement channels and up to 15 embedded data streams. AAC works at bit-rates from 8 kbps for mono speech and in excess of 320 kbps for very-high-quality audio.</td>
</tr>
<tr>
<td>MPEG4 Version 2 (scheduled to be approved Dec. 1999)</td>
<td>Builds on previous standards for digital television, interactive graphics applications and interactive multimedia.</td>
</tr>
<tr>
<td>MPEG-7 (scheduled to be approved July 2001)</td>
<td>Also called Multimedia Content Description Interface. Provides information search, filtering and management for multimedia data.</td>
</tr>
</tbody>
</table>

Licensing

Once a standard is released, it is up to private industry to develop products and technologies to take advantage of it. Often, these companies are required to pay licensing fees to companies that hold patents on technologies related to the standard. The only requirement from MPEG is that any licensing fees be fair and equitable.

Many people are surprised to learn that licensing fees are required to develop products based on an open standard. These licensing fees help compensate companies that contribute technology and other resources towards developing MPEG standards. If these companies had no way to recoup their investment, there would be little incentive for them to spend money developing technologies that their competitors could then use free of charge.
MPEG 2.5

A non-ISO extension called “MPEG 2.5” was created by the Fraunhofer Institute to improve performance at lower bit-rates. At lower bit-rates, this extension allows sampling rates of 8, 11.025 and 24 kHz. A high sampling rate at a very low bit-rate requires a trade-off in reduced resolution. Lowering the sampling rate reduces the frequency response but allows the frequency resolution to be increased, so the result is a file with significantly better quality.

MPEG Layers

Several related audio encoding schemes fall under the MPEG umbrella. These are referred to as Layers I, II and III, which exist under both MPEG-1 and MPEG-2. (Another audio encoding scheme that’s part of MPEG-2 is MPEG AAC, which is not compatible with Layers I - III.)

Each layer uses the same basic structure and includes the features of the layers below it. Higher layers offer progressively better sound quality at comparable bit-rates and require increasingly complex encoding software. This, in turn, requires more processing power for encoding and decoding the audio.

Layer-I

Layer-I was originally designed for the Digital Compact Cassette (DCC) and is not widely used.

Layer-II

Layer-II (also referred to as MP2) is widely used within the broadcasting industry. It was designed as a trade-off between complexity and performance and offers very high quality sound at higher bit-rates. It also has lower encoding delays than MP3, which is important for live broadcasting.

Layer-III

Layer-III (MP3) was designed for better quality at lower bit-rates. The high level of compression achieved by MP3 is very important because of the limited bandwidth of the Internet and the limited space of hard disks. This compression also makes MP3 well suited for portable players that use expensive solid-state memory cards.
AAC

AAC (Advanced Audio Coding) is not a MPEG layer, although it is based on a psycho-acoustic model. Sometimes referred to as MP4, AAC provides significantly better quality at lower bit-rates than MP3. AAC was developed under MPEG-2 and also exists under MPEG-4.

AAC supports a wider range of sampling rates (from 8 kHz to 96 kHz) and up to 48 audio channels, plus up to 15 auxiliary low frequency enhancement channels and up to 15 embedded data streams. AAC works at bit rates from 8 kbps for mono speech and up to in excess of 320 kbps for high-quality audio. Three profiles of AAC provide varying levels of complexity and scalability.

AAC software is much more expensive to license than MP3 because the companies that hold related patents decided to keep a tighter reign on it. Most AAC software is geared towards professional applications and secure music distribution systems, so it may be a while before you see AAC in consumer-oriented products.

AT&T’s a2b music, Global Music’s MP4 and Liquid Audio are systems for music delivery that are based on AAC. They both include schemes for copyright identification, encryption and royalty tracking. It’s important to remember that these systems are proprietary, even though they are based on an open standard.

Even though AAC is a better format for digital audio, it’s not clear whether or not it will eclipse MP3 in consumer products. MP3 can sound just as good as AAC at the expense of using more disk space, and disk space is getting cheaper all the time.

Compatibility

The various flavors of MPEG Audio are compatible with each other to some degree. Layers I, II and III are backward compatible. For example, a Layer-III decoder should also be able to decode a Layer-I or II stream, and a Layer-II decoder should be able to decode a Layer-I stream. AAC is not backward compatible with any of the MPEG layers and is sometimes referred to as “NBC,” or “not backward compatible.”

MPEG-1 layers, and the same layers under MPEG-2, are compatible with each other to a limited degree. MPEG-2 decoders must be able to
decode MPEG-1 files, and MPEG-1 decoders should be able to play the left and right channels of an MPEG-2 signal.

Most MP3 players are compatible with both MPEG-1 and MPEG-2 files, and most mainstream MP3 encoders and players are compatible with each other (though there have been compatibility issues reported with a few of the freeware encoders and some players).

Compatibility between proprietary formats based on MPEG is another story. Most of the proprietary formats based on MPEG Audio, such as AT&T’s a2b music and Liquid Audio, are not compatible with each other or with software that supports only pure MPEG formats.

Some features added to MPEG Audio (such as watermarking) should not affect compatibility, but many proprietary formats use encryption. And any form of encryption is likely to make these formats incompatible with each other and with products that support only pure MPEG Audio.

**MPEG Encoding**

MPEG Audio uses what’s referred to as perceptual encoding (a type of “lossy” compression. To compress audio, MPEG encoders first apply a psycho-acoustic model to identify parts of the signal that most people can’t hear. The encoder removes these sounds from the signal and then applies standard lossless data compression techniques.

This technique does not work perfectly because the sensitivity of each person’s hearing is different. But the sensitivity of human hearing does fall within a finite range, and researchers can determine a range that applies to the vast majority of people.

**Sub-bands**

The encoder first divides the signal into multiple sub-bands, so the encoded signal can be better optimized to the response of the human ear. For example, most of stereo information below 100 Hz can be discarded because the ear cannot determine the direction of very low frequency sounds; but at higher frequencies the ear is more sensitive to direction of sounds, so more stereo information needs to be retained.
Minimum Audible Threshold

The level below which all sounds are inaudible to the human ear is called the threshold of hearing, or minimum audible threshold. This threshold varies according to frequency because the human ear does not have a linear response.

Sounds below this threshold can be removed by the encoder, and most listeners will not detect any difference between the encoded signal and the original. The ear is most sensitive to frequencies between 2 kHz and 4 kHz, so less information can be removed from this range without affecting the quality of the sound.

Figure 27 - MPEG Audio Encoding

Figure 28 shows the Fletcher-Munsen curve, which illustrates how the threshold of human hearing varies according to frequency.

Masking Effect

Quiet sounds are “masked” by louder sounds that are close to them in frequency and time. Since you can’t hear these sounds, they can be removed from the signal without affecting the perceived quality. An example is the hiss and other background noise you hear when a song is paused or blank tape is playing. When the music plays above a certain level you can no longer hear this background noise, but it is still there in the signal.
Reservoir of Bits

Certain musical passages need to be encoded at higher bit-rates to maintain fidelity, so MP3 creates a reservoir by setting aside bits from less complex passages. These extra bits can then be applied to more complex passages, where they are needed more. This is different from variable bit-rate encoding, because a fixed number of bits are allocated—they just are shifted to where they are needed most.

Stereo Modes

Stereo audio normally requires twice the bandwidth of mono because it uses two separate channels. Much of the information is identical on both channels. For example, any sounds positioned at the center of the stereo image will be carried by both channels. This wastes a lot of space because the information is identical. MPEG Audio has several ways of handling stereo information. Each method varies in the amount of compression and the fidelity to the stereo image.

Simple Stereo (mode 0) is the closest to a normal stereo signal. It uses independent channels; therefore, any duplicate information will be
retained, and some bandwidth will be wasted. The MPEG encoder can vary the allocation of bits between channels according to the complexity of the signal. The overall bit-rate remains constant, but the split between the channels varies according to the dynamic range of each channel.

**Joint Stereo** (mode 1) uses MS (middle/side) Stereo, where one channel carries the information that is identical on both channels and the other carries the difference. Joint Stereo retains all the original stereo information and uses bandwidth very efficiently.

**Intensity Stereo** encodes only the stereo information that is perceived as important to the stereo image. Intensity Stereo provides the highest level of compression, but the stereo image will suffer at lower bit-rates.

Although Simple Stereo is the closest to a normal stereo signal, it is not the best option to use with MPEG Audio. In most cases, Joint Stereo will produce higher quality sound because the bits can be allocated more efficiently.

**Huffman Encoding**

In any musical composition, certain sound patterns are repeated—some more often than others. These patterns can be coded with symbols to save space, then decoded into the original pattern when played. Huffman encoding increases compression by using shorter codes for more common sound patterns. It’s similar to replacing every word in a document with a number and using the smaller numbers for the most common words.

**Bit-rates**

MPEG Audio supports constant and variable bit-rates ranging from 8 kbps to 1.5 mbps. Just as with uncompressed audio, the bit-rate of MPEG Audio has a direct relationship to sound quality and file size.

Constant bit-rate (CBR) encoding is not very efficient because it uses the same number of bits, regardless of how complex or simple the passage is. Variable bit-rate (VBR) encoding varies the number of bits depending on the complexity of the music and is more efficient than CBR. For example, a simple passage with just a vocalist and acoustic guitar needs fewer bits than a passage with a full symphony.
**Resolution and MPEG Audio**

MPEG encoders rely on the resolution used in the uncompressed audio file to set the range of resolution that will be used for the encoded file. The resolution of the encoded file is varied according to the complexity of the signal to achieve compression. Many encoders are optimized to work with 16-bit resolution input, and some will only accept 44. kHz, 16-bit WAV files as input.

Table 15 shows the file sizes and relative amounts of compression for different bit-rates. As the bit-rate increases, so does the sound quality, along with the file size. This table also shows how many hours of audio or four-minute songs, a 1GB hard disk will hold at each rate.

**Table**  Error! Bookmark not defined. - **File Size vs. Bit-rate**

<table>
<thead>
<tr>
<th>Bit-rate</th>
<th>File Size (4-min. song)</th>
<th>MB per Minute</th>
<th>Compression Ratio</th>
<th>Hours per GB</th>
<th>4-min. Songs per GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,411 kbps (CD Audio)</td>
<td>41.3MB</td>
<td>10.3</td>
<td>None</td>
<td>1.7</td>
<td>25</td>
</tr>
<tr>
<td>80 kbps</td>
<td>2.3MB</td>
<td>0.6</td>
<td>7.6 = 1</td>
<td>29.1</td>
<td>437</td>
</tr>
<tr>
<td>128 kbps</td>
<td>3.8MB</td>
<td>0.9</td>
<td>11.0 = 1</td>
<td>18.2</td>
<td>273</td>
</tr>
<tr>
<td>160 kbps</td>
<td>4.7MB</td>
<td>1.2</td>
<td>8.8 = 1</td>
<td>14.6</td>
<td>218</td>
</tr>
<tr>
<td>192 kbps</td>
<td>5.6MB</td>
<td>1.4</td>
<td>7.3 = 1</td>
<td>12.1</td>
<td>182</td>
</tr>
<tr>
<td>256 kbps</td>
<td>7.5MB</td>
<td>1.9</td>
<td>5.5 = 1</td>
<td>9.1</td>
<td>137</td>
</tr>
<tr>
<td>320 kbps</td>
<td>9.4MB</td>
<td>2.3</td>
<td>4.4 = 1</td>
<td>7.3</td>
<td>109</td>
</tr>
</tbody>
</table>

**Signal Delays**

The process of encoding and decoding audio introduces a slight delay into the signal. This is not a problem for home use, but it is a factor for applications where a short delay is critical, such as two-way voice
conversations, where a delay of more than 10 ms (milliseconds) can be disturbing. Delays for MPEG Audio typically range from 19 ms for Layer-I to more than 60 ms for Layer-III and AAC. The actual delay depends on the hardware and software used.

**Embedded Data (ID3 Tags)**

MPEG Audio is frame-based, which allows it to support the insertion of additional program information in the form of text, graphics and other data. The standard is flexible enough that software developers can include almost any type of data, such as copyright information, lyrics, album artwork and even links to artist’s Web sites.

**ID3 Tags**

An informal standard called ID3 tagging has emerged that specifies a format for storing non-audio data inside MP3 files. The ID3 information can be displayed and edited by MP3 players such as Winamp. The ID3 tag is placed at the very end of the MP3 file, which makes it unsuitable for streaming audio.

ID3 Version 1 is limited to 128 bytes of data and 30 characters per field and contains fixed length fields for title, artist, album, year, comments, track number and genre. Audio CDs do not contain this information, so it needs to be entered manually or obtained from a database, such as the CDDB (see Chapter 9, *Organizing and Playing Music*). The identification field must contain the characters “TAG” to indicate ID3 version 1 compliance.

ID3 Version 1.1 takes the last two characters of the comments field and uses them for the number of the CD track that the song originated from.

**Table** Error! Bookmark not defined. - **ID3 Tag Version 1.1 Fields**

<table>
<thead>
<tr>
<th>Position</th>
<th>Length (Bytes)</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-2</td>
<td>3</td>
<td>Identification</td>
</tr>
<tr>
<td>3-32</td>
<td>30</td>
<td>Title</td>
</tr>
<tr>
<td>33-62</td>
<td>30</td>
<td>Artist</td>
</tr>
</tbody>
</table>
ID3v2

ID3 Version 2 is designed to be more flexible and expandable than version 1.1. Each tag contains smaller chunks of data, called frames. Each frame can contain any type of data, such as lyrics, album cover graphics and links to a band’s Web site. The ID3v2 tag is placed at the beginning of the file, which makes it useful for streaming applications. A unique feature called the Popularimeter can be used to keep track of how often you listen to each song, and this information could be used to automatically construct playlists based on your personal tastes.
Measuring Sound Quality

Sound quality is subjective, so traditional measures like total harmonic distortion (THD) and signal-to-noise ratio are not useful for rating perceptual encoding schemes. The perceived quality of the sound is more important than any characteristic that can be measured with test equipment. Controlled tests with trained listeners are the best way of measuring the performance of perceptual encoding schemes.

During the MPEG-1 development process, three international listening tests were performed using the CCIR (Centre for Communication Interface Research) impairment scale shown in Table 17. At 128 kbps, MP3 scored between 3.6 and 3.8. This indicates that listeners detected a difference between the MP3 and the original but the difference was not annoying. At 240 kbps and above, MP3 scored at the high end of the scale, and most listeners found it difficult to distinguish between the MP3 and the original version.

- Has an “unsynchronization” feature to prevent ID3v2 incompatible players from attempting to play the tag.
- Maximum tag size is 256MB; maximum frame size is 16MB.
- Supports Unicode and the capability to compress data.
- Has several new text fields such as composer, conductor, media type, BPM, copyright message, etc.
- Able to contain both plain and synchronized lyrics (for karaoke).
- Can contain volume, balance and equalizer settings.
- Supports encrypted information, images and hyperlinks.
### CCIR Impairment Scale

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0</td>
<td>Imperceptible (indistinguishable from the original)</td>
</tr>
<tr>
<td>4.0</td>
<td>Perceptible (perceptible difference, but not annoying)</td>
</tr>
<tr>
<td>3.0</td>
<td>Slightly annoying</td>
</tr>
<tr>
<td>2.0</td>
<td>Annoying</td>
</tr>
<tr>
<td>1.0</td>
<td>Very annoying</td>
</tr>
</tbody>
</table>

### Variables That Affect Sound Quality

The major variables that affect the sound quality of encoded audio are the type of encoder, the bit-rate, the type of music and the sensitivity of the listener’s hearing. The quality of commercially available encoders is generally very good, and most people would find it difficult to tell the difference between two MP3 files encoded from the same song by different encoders. Assuming you’ve already decided on using MP3, the bit-rate is the biggest factor that you can control.

In general, music that is more complex will require higher bit-rates. A good example is classical (or symphonic) music. Classical music is generally more complex, because there are more instruments and a wider dynamic range compared to most other types of music, such as blues and rock. Variable bit-rate encoding is a good choice for all types of music because it provides significantly better quality than constant bit-rate encoding at a similar rate. This is because the bits are allocated where they are needed most, which also helps maintain a more constant signal-to-noise ratio.

Table 18 shows the bit-rates for various digital audio formats that will produce high quality sound for most types of music.

### Bit-rates for High Quality Sound

<table>
<thead>
<tr>
<th>Format</th>
<th>Bit-rate</th>
<th>Compression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Book (CD)</td>
<td>1.4Mbps</td>
<td>None</td>
</tr>
<tr>
<td>MPEG Layer-I</td>
<td>384 kbps</td>
<td>3.6=1</td>
</tr>
<tr>
<td>MPEG Layer-II</td>
<td>256 kbps</td>
<td>5.5=1</td>
</tr>
<tr>
<td>MPEG Layer-III (MP3)</td>
<td>192 kbps</td>
<td>7.3=1</td>
</tr>
<tr>
<td>MPEG Layer-III (MP3)</td>
<td>VBR Normal/High</td>
<td>7=1 to 10=1</td>
</tr>
<tr>
<td>MPEG AAC</td>
<td>128 kbps</td>
<td>11=1</td>
</tr>
<tr>
<td>ID</td>
<td>Genre</td>
<td>ID</td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>-----</td>
</tr>
<tr>
<td>0</td>
<td>Blues</td>
<td>20</td>
</tr>
<tr>
<td>1</td>
<td>Classic Rock</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>Country</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>Dance</td>
<td>23</td>
</tr>
<tr>
<td>4</td>
<td>Disco</td>
<td>24</td>
</tr>
<tr>
<td>5</td>
<td>Funk</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>Grunge</td>
<td>26</td>
</tr>
<tr>
<td>7</td>
<td>Hip-Hop</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>Jazz</td>
<td>28</td>
</tr>
<tr>
<td>9</td>
<td>Metal</td>
<td>29</td>
</tr>
<tr>
<td>10</td>
<td>New Age</td>
<td>30</td>
</tr>
<tr>
<td>11</td>
<td>Oldies</td>
<td>31</td>
</tr>
<tr>
<td>12</td>
<td>Other</td>
<td>32</td>
</tr>
<tr>
<td>13</td>
<td>Pop</td>
<td>33</td>
</tr>
<tr>
<td>14</td>
<td>R&amp;B</td>
<td>34</td>
</tr>
<tr>
<td>15</td>
<td>Rap</td>
<td>35</td>
</tr>
<tr>
<td>16</td>
<td>Reggae</td>
<td>36</td>
</tr>
<tr>
<td>17</td>
<td>Rock</td>
<td>37</td>
</tr>
<tr>
<td>18</td>
<td>Techno</td>
<td>38</td>
</tr>
<tr>
<td>19</td>
<td>Industrial</td>
<td>39</td>
</tr>
</tbody>
</table>

Source: [www.dv.co.yu/mpgscript/mpeghdr.htm](http://www.dv.co.yu/mpgscript/mpeghdr.htm)

### Additional Resources

<table>
<thead>
<tr>
<th>Organization</th>
<th>Web Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>American National Standards Institute (ANSI)</td>
<td><a href="http://wwwansi.org">wwwansi.org</a></td>
</tr>
<tr>
<td>Centre for Communication Interface Research (CCIR)</td>
<td><a href="http://www.ccir.org">www.ccir.org</a></td>
</tr>
<tr>
<td>Fraunhofer Gesellschaft</td>
<td><a href="http://www.iis.fhg.de/amm/techinf">www.iis.fhg.de/amm/techinf</a></td>
</tr>
<tr>
<td>ID3 Tag Specification</td>
<td><a href="http://wwwid3.org">wwwid3.org</a></td>
</tr>
<tr>
<td>International Standards Organization (ISO)</td>
<td><a href="http://wwwiso.ch">wwwiso.ch</a></td>
</tr>
<tr>
<td>Moving Picture Experts Group (MPEG)</td>
<td><a href="http://drogo.cselt.stet.it/mpeg">httpdrogo.cselt.stet.it/mpeg</a></td>
</tr>
</tbody>
</table>
Hard Disk Recording

The Internet and CDs aren’t the only sources for digital music. You may have old records that have never been released on CD that you would like convert to a digital format. You may have deteriorating tapes you would like to preserve. Or maybe you want to record live music, create audio books or record sound effects for a multimedia presentation.

Any type of analog audio can be preserved (and sometimes improved) by recording it in a digital format. Once the audio is in digital format, it’s easy to clean it up or add special effects. Digital audio also takes up less physical space than analog audio, even without compression.

The process of recording and storing the audio on a computer is called hard disk recording. Dedicated hard disk recorders costing thousands of dollars have been available to professional recording engineers for years. Now, with the right hardware and software, you can produce professional quality digital recordings on your computer for much less money.

Hard disk recording works much the same as tape recording. Audio is recorded in real-time from analog sources, such as records or cassette tapes or from digital sources, such as MiniDiscs or DATs. One hour of audio still takes 60 minutes to digitally record. But once audio is in a digital format, you are no longer limited to working in real-time. You can then use playlists to program continuous music in a fraction of the time it would take using a tape recorder.

You can record directly from a microphone or any other source fed into your sound card’s inputs. With the right sound card and software, anything you hear on your computer can also be recorded (for example, sounds from other programs or streaming audio from the Internet).
The recording capacity of your computer is limited only by the amount of free space on your hard disk. The total time will vary, depending on the sampling rate, resolution and number of channels. (See Chapter 11, *A Digital Audio Primer*, for more information on calculating file sizes.)

**Recording from Analog Sources**

Even though ripping and direct digital recording produce better quality sound, there are many cases where you have no choice but to use analog recording. You may have a record or tape you want to preserve, or you may want to record acoustical instruments or vocals, using a microphone.

Recording with a sound card in this manner starts out with an analog signal, so the signal will degrade slightly during the A/D conversion process. The result, however, will be a digital signal stored in a file that will not degrade, unlike records and analog tapes that will degrade a little bit each time they are played.

To record an analog source, it must be connected to your sound card’s line input with the right type of cable. (See Chapter 10, *Connecting Your PC to Your Stereo.*)

---

**MiniDiscs and DATs**

MiniDiscs use a lossy encoding scheme similar to MPEG Audio. The
If you have a component with a digital output (such as a DAT or MiniDisc player) and your sound card has a digital input, you can make a direct digital recording. This bypasses the A/D converter and analog circuits in the sound card and results in a better quality recording. Digital recording is still a real-time process, like analog recording.

Recording Software
Most operating systems include a basic recording program such as the Windows Sound Recorder, which is limited in functionality. If you plan on doing a lot of recording, you should purchase a specialized program like Cool Edit, Sound Forge or SoundEdit 16 (see Appendix B, What and Where to Buy). Specialized recording programs generally work better, have more features and can handle large files better.

Some sound cards come bundled with stripped-down versions of full-featured recording and editing software. These “lite” programs usually are adequate for basic recording and editing functions, such as removing silence or normalizing a file. The Sound Blaster Live comes bundled with Sound Forge XP (a good deal considering the software alone sells for about $50). Most bundled programs offer a reasonable cost upgrade to the full version.

Cool Edit
Cool Edit is a great sound recording and editing program, well liked by many users. You can use CoolEdit to remove the blank spaces at the beginning or end of a song, normalize audio files, remove noise from recordings and more. Cool Edit 96 is a “lite” version and offers the basic recording and editing features needed by most users. The professional version of Cool Edit adds more extensive capabilities, including multi-track recording.

Recording Directly to MP3
Most sound editing programs do not have the capability to record
directly to MP3, so the audio is normally first stored in a WAV file and then converted to MP3 in a separate step. A few of the all-in-one programs such as MusicMatch Jukebox, RealJukebox and the Macintosh version of AudioCatalyst can record directly to MP3.

Even though these programs can record directly to MP3, it’s often better to record to a WAV file first. This gives you the chance to clean up the file before it’s encoded. Also, more things can go wrong when you encode while recording because the computer has to work harder. Any interruptions can ruin an otherwise good recording.

**Sound Forge**
Sound Forge is a high-end professional sound recording and editing program. It reads and writes to almost any format and can process studio-quality audio files and optimize files headed for the Internet. Sound Forge XP is a “lite” version of Sound Forge and offers the basic recording and editing features needed by most users.

**Total Recorder**
Total Recorder is a universal sound recording tool that allows you to record digital and analog sounds from multiple sources. TotalRecorder can also record sound that is being played by other media programs, including live Internet broadcasts. You can use its scheduler to record a streaming audio program, even if you can’t be there when it is happening. It’s almost like having a VCR for the Internet.

**Sound Card Mixing Functions**
In addition to digital to analog conversion, sound cards work with software to perform the functions of a mixing console. On Windows systems, the Volume Control program provides the interface to the sound card’s mixer functions. Some recording programs and software bundled with sound cards provide their own mixer interface in place of the Volume Control program.

**Volume Control**
The Volume Control program provides controls for adjusting the volume and balance for each channel supported by the sound card. Separate screens are used for recording (input) and playback (output) controls.
Most sound editing and player programs have their own volume control that links to the built-in Volume Control. Changing the volume in these programs will cause one of the sliders in the Volume Control program to move (and vice versa).

Many sound card drivers replace the built-in Volume Control program and add or modify channels. Depending on which sound card is installed, the playback and recording controls may be labeled differently. The playback level screen is usually labeled “Volume Control,” but the sound card software may change it to “Play Control” or “Speaker Control.” The recording level screen is usually labeled “Recording Control” or “Record Control.” Regardless of the labels, the basic functions are the same.

**Playback Control**

The Playback Control screen provides level and balance controls for each input supported by the sound card. The inputs most people will use are CD Audio, Line-in, Mic and Wave. The Wave control is used to adjust the playback level of digital audio files (MP3, WAV, etc.). The

---

**Figure 30 - Sound Card Mixing Functions**

![Sound Card Mixing Functions Diagram]

- **Playback Levels**
  - Line-In
  - CD
  - Wave
  - Mic

- **Master Levels**
  - Preamp
  - Amp

- **Recording Levels**
  - Recording Software

- **Outputs**
  - Line-Out
  - Headphones
  - Audio File
Advanced button on some channels provides access to tone controls and special effects like 3D sound and reverberation.

**Playback Control**

Levels can be adjusted individually for each channel. The master volume slider sets the overall level sent to the line output and/or headphone jack. Check the Mute checkbox on any unused channels to reduce noise.

Channels can be displayed or hidden by clicking on the Options menu, then Properties, and checking the box next to each channel you want displayed. This makes the screen less cluttered by displaying only the channels you need. For example, if you are playing a CD and want use a microphone to make a voice-over announcement, you might want to display only the CD and microphone sliders.

**Recording Control**

The Recording Control provides separate level and balance controls for each input supported by the sound card. These sliders are independent of the playback control sliders and are used to set the levels of signals sent to the recording program. The playback levels only affect how loud the music is played and have no affect on the level of the signal recorded. It
is possible to have a source playing very loud, yet a recording level that is much too low.

When you launch the Volume Control, it defaults to the Playback Control screen. To adjust recording levels, select **Options** from the pull-down menu, then **Properties**, and then **Recording**. The screen with the recording controls will appear.

**Recording Control Properties**

![Recording Control Properties](image)

Below each slider is a checkbox (labeled “Select”) that activates that source. In the example below, Line-In is the only source selected, so only the signal from the sound card’s line input jack will be recorded.
Some sound card software adds a master volume control and recording level meter to the Recording Control screen. If your Recording Control screen doesn’t have a master recording level meter, you can use the meter in your recording program.

### SoundBlaster Live

The Recording Control of the SoundBlaster Live only allows you to select one device at a time, unlike most other sound cards, which allow you to select multiple devices. To record a mix of sound from multiple sources, select the “What U Hear” control. Then go to the Playback Control and select the sources you want to record.

### Recording Levels

It’s important to set the input levels as high as possible to obtain a good signal-to-noise ratio and maximum dynamic range. However, they should not be set so high that the meter stays constantly in the yellow level, except for occasional peaks. If the meter reaches the red level, the recording level is approaching the maximum, or clipping level. You should avoid clipping at all costs because it causes extreme distortion.

Programs like Cool Edit and Sound Forge have their own level meters that are more precise than the simple meter in the Recording Control.
These meters are usually labeled in dB, with 0dB equal to the maximum level. Levels below the maximum are shown in negative dB and the lowest possible level is referred to as infinity.

When using this type of meter to set recording levels, make sure the peaks average around –6 dB and don’t exceed –3 dB. This will normally give you enough headroom to avoid clipping, while maintaining a good signal to noise ratio. These meters usually have a peak reading marker to show the highest level measured during a recording session.

The Recording Process

Although hard disk recording is similar to tape recording, there are several additional steps. Below is an outline of the basic process. (See Chapter 20, Software Tutorials, for instructions on recording with Cool Edit 96.)

Setting Audio Parameters

Before you start recording, you must specify the sampling rate, resolution and number of channels. For CD-quality audio, you would choose 16-bit stereo at 44.1 kHz. For voice or background music, you could choose mono to save space. You could also choose 8-bits or a lower sampling rate for voice or music that does not need to be high quality.
Selecting the Source

The source, whether internal or external, needs to be selected from the Recording Control screen of the Volume Control program (Windows). For an external source, select **Line In**, **Mic** or **S/PDIF-In**. For an internal source, you would normally select **Wave** (or **Wav/Direct Sound**). If you have a Sound Blaster Live, you can select **What U Hear** to record everything that passes through the sound card.

Make sure to only select the channel for the source you want to record, otherwise noise on other selected channels may get mixed in with the recording. It’s also a good idea to zero the level sliders on unused channels to prevent noise leakage.

Setting the Recording Level

Just like making a tape recording, you should set as high a level as possible without the signal being clipped. If the level is too low, the signal will not take advantage of the full dynamic range and will be noisy. If the level is too high the signal will be clipped and distorted.

To set the correct level, start the source playing, and watch the level meters. Adjust the level control sliders so the peaks stay below the red area (about -3 dB).

Skip forward to the loudest part of the song to make sure those peaks are not too high. Once you are satisfied with the level, rewind or reset the source and pause it at the beginning of the track.

**Tip**

If you need to switch back and forth between playback and recording controls, instead of going back and forth through the pull-down menus, you can launch two instances of the Volume Control program and set one to display the playback controls and the other to display the recording controls.

To do this, click on the Volume Control icon to launch one instance of it, then click on the icon again to launch a second instance. Both instances will default to playback controls, so switch one of them to display the recording controls. You can now switch between them by clicking on their icons in the Windows Task Bar.
Recording

Now you are ready to record. In quick succession, click on the Record button in the recording program, then press or click (if the source is another program) the Play button on the source. If the level meters start moving and you don’t hear anything, the playback level control is probably muted. To monitor the sound while recording, go to the Playback control and uncheck the mute buttons for Wave or Direct Sound and adjust the volume sliders.

When the playback is complete, stop the source and also click the stop button in the recording program. Zoom out and view the whole recording. This will show you quickly whether or not there are any gaps or clipping. You may notice a section of silence at the beginning and end of the recording. This can be removed using the trim feature of the recording software (see Chapter 17, Editing Sound Files). You may want to keep a half second of silence at either end. This is a matter of preference, and the appropriate amount of silence will vary depending on the song.

File Formats

Before you save the file, you must specify the format (e.g., WAV, AU, AIFF, etc.). Some recording programs require that you select the format before recording. Choose PCM WAV with a sampling rate of 44.1 kHz and 16-bit resolution if you plan to convert the file to MP3 format.

Reducing Background Noise

Noise can be introduced into an audio signal in many places and can ruin an otherwise good recording. A good test, before recording on a
When recording from vinyl records, clean the record and make sure your turntable’s stylus and cartridge are in good shape. When recording from tape, make sure the heads are clean and demagnetized. In either case, use good quality shielded cables to reduce noise from electrical interference. Set the highest possible recording level, without clipping, to help mask noise and maximize the dynamic range.

The quality of your sound card will have a big effect on the quality of your recordings. When you record through a sound card, the A/D conversion process adds distortion from quantization errors and electrical noise can be picked up from other components in the computer.

Many lower priced sound cards are poorly shielded, which makes them more susceptible to noise. Some lower priced sound cards also have low resolution A/D converters, which will introduce more distortion from quantization errors.

If you plan to make analog recordings, consider using a better sound card such as the Creative Sound Blaster Live or the Turtle Beach Montego II, or use an external A/D converter, such as Opcode’s.
SONICport. (See Appendix B, *What and Where to Buy*, for more information).

**Record and Tape Preservation**

Many people have collections of vinyl records or tapes they would like to preserve. Often these recordings are not available in digital format (or they may not be replaceable in any format). Digitizing a recording (recording it in a digital format) is an excellent way to preserve the music and avoid wear and tear each time the record or tape is played.

Many recording programs can also be used to remove clicks, pops and background noise. Most sound editing programs have special filters for removing unwanted noise. A few programs, such as Adaptec’s Spin Doctor, are specifically geared towards recording and cleaning up audio from vinyl records and tapes. You can also use Cool Edit and Sound Forge to remove noise, although Spin Doctor is somewhat easier to use.

Spin Doctor allows you to set separate levels for noise filtering and pop removal. Cool Edit lets you select a section of silence so it can develop a noise profile to apply to the rest of the file. Cool Edit allows you to save noise profiles and use them to remove noise from similar files. Sound Forge is slightly more difficult to use but works well.

Whichever program you use, you must be careful to avoid removing audible parts of the music along with the noise. It’s a good idea to try several settings and listen to the result before you remove any noise permanently.
Table Error! Bookmark not defined. - Sound Recording and Editing Software

<table>
<thead>
<tr>
<th>Program</th>
<th>Manufacturer</th>
<th>Retail Price</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cool Edit 96</td>
<td>Syntrillium</td>
<td>$50.00</td>
<td>Great low-cost sound recorder and editor (shareware)</td>
</tr>
<tr>
<td>Cool Edit Pro</td>
<td>Syntrillium</td>
<td>$399.00</td>
<td>Professional version of Cool Edit</td>
</tr>
<tr>
<td>GoldWave</td>
<td>Chris Craig</td>
<td>$40.00</td>
<td>Nice low-cost sound recorder and editor (shareware)</td>
</tr>
<tr>
<td>MPEG Tape Deck</td>
<td>Marcus Kuenzel</td>
<td>Free</td>
<td>Records and encodes audio from external sources to MPEG format</td>
</tr>
<tr>
<td>Peak LE (Mac)</td>
<td>Bias, Inc.</td>
<td>$99.00</td>
<td>Good low-cost sound recorder and editor</td>
</tr>
<tr>
<td>SoundEdit 16 (Mac)</td>
<td>Macromedia</td>
<td>$349.95</td>
<td>High-end, professional-quality sound recorder and editor</td>
</tr>
<tr>
<td>Sound Forge</td>
<td>Sonic Foundry</td>
<td>$499.95</td>
<td>High-end, professional sound recording and editing</td>
</tr>
<tr>
<td>Sound Forge XP</td>
<td>Sonic Foundry</td>
<td>$59.95</td>
<td>Lite version of Sound Forge (bundled with Sound Blaster Live)</td>
</tr>
<tr>
<td>Total Recorder</td>
<td>High Criteria</td>
<td>$11.95</td>
<td>Records audio from any source, including other programs</td>
</tr>
</tbody>
</table>

Additional Resources

<table>
<thead>
<tr>
<th>Site</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio Cafe</td>
<td><a href="http://www.audiocafe.com">www.audiocafe.com</a></td>
</tr>
<tr>
<td>Home Recording</td>
<td><a href="http://www.homerecording.com">www.homerecording.com</a></td>
</tr>
<tr>
<td>Mix Online</td>
<td><a href="http://www.mixonline.com">www.mixonline.com</a></td>
</tr>
</tbody>
</table>

Digital Audio Extraction

Digital audio extraction (DAE), commonly referred to as ripping, is the process of copying audio data directly from a CD. Because it bypasses the sound card, ripping normally results in a perfect copy with no introduction of noise or loss of fidelity.

Ripping is the fastest way to get songs from a CD onto your computer. Since it is a digital copying process, the speed of ripping is limited only by the performance of your hardware and software, unlike recording, which is always a real-time process.
When you record a CD through a sound card, the digital audio data is converted to analog, then resampled and converted back to digital. While the signal is in analog form, it can pick up noise from the interior of the computer. When the signal is converted back to digital, quantization distortion will introduced by the A/D converter in the sound card.

When you record a four-minute song from a CD, it will always take at least four minutes to record, whether you use a tape recorder, sound card, or any other recording method. However, with a fast CD-ROM drive, the same song can be ripped in less than 30 seconds.

**Ripping Software**

Specialized software is required for ripping. All-in-one programs, such as MusicMatch Jukebox and RealJukebox, include rippers, encoders, players and playlist managers in one program. A few manufacturers, like Plextor, include ripping software with their CD-ROM drives. Audiograbber (also used in AudioCatalyst) is one of the best ripping programs available. MusicMatch also has a very capable ripper.
Will Your Drive Rip?

The performance of your CD-ROM drive is the single biggest factor in the success of ripping. Not all drives support digital audio extraction, and most manufacturers do not include digital audio extraction performance in their specifications. This may be because manufacturers worry that advertising this capability would make them subject to the royalty requirements for digital recording devices marketed for consumer use.

The Web sites listed in Table 21 have compiled lists of CD-ROM drives that support digital audio extraction. This information is provided by end-users and usually includes the fastest ripping speed and the software used. These lists are by no means complete, but if your drive is listed, the information may be useful for troubleshooting.

<table>
<thead>
<tr>
<th>Site</th>
<th>Web Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD Digital Audio Extraction Page</td>
<td><a href="http://www.tardis.ed.ac.uk/~psyche/cdda/">www.tardis.ed.ac.uk/~psyche/cdda/</a></td>
</tr>
</tbody>
</table>

Figure 32 - Ripping Audio From a CD
To determine the model number of your CD-ROM drive (in Windows 95 or later), click the Start button, and select Settings, then Control Panel. Double-click on the System icon, then select the Device Manager tab. Click on the plus (+) symbol to the left of the CD-ROM icon. This will display a second CD-ROM icon labeled with the manufacturer and model number of your drive.

Keep in mind that even if a CD-ROM drive supports ripping, it doesn’t mean the drive will work well on your system. Many variables affect ripping performance, including processor speed, hard disk fragmentation, the type of ripping software, and the CD-ROM access method used (e.g., ASPI vs. MSCDEX). Other factors, such as bad cables, software incompatibility and incorrect configuration settings, can prevent even a good CD-ROM drive from ripping.
If your drive is listed as being able to rip but doesn’t work properly, your software or system configuration may be at fault. If your drive is listed as not being able to rip, it doesn’t necessarily mean it can’t. It just means that someone else with the same model drive tried ripping with it and was unsuccessful. If you can’t get your drive to rip after trying a few different settings, you are probably better off replacing it rather than struggling to get it to work.

If your CD-ROM rips but frequently has problems, it’s well worth the cost of a new drive that works more reliably. When I started ripping my CD collection, I probably wasted 20 to 30 hours messing with a marginal CD-ROM drive. Once I installed a new drive, 95% of my ripping problems went away.

**Jitter**

Pops or breaks in the sound of a ripped audio track are often caused by jitter. Jitter is caused by the inability of many CD-ROM drives to accurately seek a specific sector (also called a frame) on an audio CD.

CDs were originally designed for audio and then adapted to computer data. In an audio CD player, once the laser read head is in position, the data is read in a continuous stream. The head does not have to jump to a new position while playing, it simply follows the spiral track. Computers read information from CDs in blocks, rather than in a continuous stream.
Programs that extract CD audio must first read a block of sectors, then write this data to the hard disk. The drive must then seek the beginning of the next block of sectors. The Red Book CD audio specification states that a CD player only needs to be accurate to within $1/75^{\text{th}}$ of a second. Because of this $1/75^{\text{th}}$ of a second tolerance, when a program is extracting CD audio, it can’t be sure that the sector returned by the drive is the exact one it requested.

**Figure 34 - Jitter Correction**

Most ripping software has settings for jitter correction (sometimes referred to as synchronization or error correction) that can correct this problem. With jitter correction enabled, the ripping software reads sectors in blocks, and overlaps the reads by a specified number of sectors. It can then compare the blocks and discard the sectors that overlap.

Jitter correction slows down ripping because it takes more time to read the overlapping data. Older drives tend to have more jitter problems than newer models. Drives with poor seeking accuracy may rip unreliably or not at all, even with jitter correction. Some newer drives, such as the Plextor models, perform jitter correction internally.

**Seek Errors**

Not all CD-ROM drives can produce identical files extracting the same track at different times. This doesn’t necessarily mean that the track was ripped inaccurately. Because of the $1/75^{\text{th}}$ second positioning inaccuracy, the CD-ROM drive can’t tell exactly where the recording started on a track. Jitter correction can’t help since there is no previous block to synchronize with.
If you extract the same track twice, there’s a chance it will start at a slightly different position. The extracted track may contain slightly more or less silence at the beginning. This changes the binary data slightly but doesn’t affect how a track sounds. This type of inconsistency is more common with lower-cost CD-ROM drives. Many of the better drives can rip the same tracks multiple times and consistently produce identical files, even when different ripping software is used.

**Jitter Correction Settings**

The settings of your ripping software, especially those for jitter correction, play a big part in ripping performance. It pays to experiment with several settings and compare the files to see which one works best. If you have one of the Plextor drives, always rip with jitter correction disabled, since the drive handles this internally. Otherwise, the ripping will take longer than necessary because of the extra time required for the overlapping reads of the software jitter correction.

**MusicMatch**

MusicMatch is very good at optimizing itself for jitter correction (MusicMatch uses the term “error correction”) because it tests the CD-ROM drive and automatically configures itself. You can toggle jitter correction on and off but you shouldn’t have to tweak any of the advanced settings. If your drive will not rip, you can try changing a few of MusicMatch’s advanced settings, but there’s a good chance the problem lies elsewhere.

**AudioCatalyst**

AudioCatalyst (and Audiograbber) use the term synchronization in place of jitter correction. AudioCatalyst provides more control over the ripping process than MusicMatch, but you must manually configure it. Buffered and unbuffered burst settings provide no jitter correction. Dynamic sync varies the amount of correction depending on the data. Fixed width sync provides progressive levels of jitter correction from 1 to 10.

If you need to use a higher level than 4 or 5 you should consider replacing your CD-ROM drive. AudioCatalyst supports both ASPI and MSCDEX ripping, but ASPI is much better. MSCDEX should only be used if there is no other choice. For step-by-step instructions for ripping with AudioCatalyst, see Chapter 20, *Software Tutorials*. 
Successful Ripping

Once you have determined optimum settings for ripping on your system and have successfully ripped at least one CD, you generally should not have to change them. Although, in some cases, changes in your computer’s configuration may affect ripping performance and require reconfiguration of the ripping program.

### Tips for Successful Ripping

- Use a CD-ROM drive that supports digital audio extraction and has good seeking accuracy.
- Make sure your system is not running unnecessary programs or processes when it’s ripping.
- Keep your hard drive defragmented.
- Make sure your CD-ROM drivers are up-to-date.
- Use the ASPI method of accessing the CD-ROM.
- Use jitter correction (sometimes called error correction or synchronization) unless your drive handles this directly.
- Disable auto insert notification for your CD-ROM drive.
- Don’t waste your time with a marginal CD-ROM drive or one that does not rip at 4X or better speed.
- Test with a few tracks from different CDs and listen to the WAV files to verify the quality before ripping your entire CD collection.

One of the first things to do if you have trouble either ripping or encoding is to exit all other programs. On Windows systems you can press the Ctrl, Alt and Del keys simultaneously to display the Windows Task Manager and see a list of all programs that are in memory. See Chapter 18, *Recording Your Own CDs* for tips on optimizing your system.

Programs running in the background may cause problems because they put more of a load on the system’s processor and tie up memory that otherwise would be available. Idle programs, like a wordprocessor or a spreadsheet, normally will not cause problems, but they still use
memory. Any program that puts a load on the processor or writes frequently to the hard disk can also cause problems.

**Ripping Speed**

A good CD-ROM drive should rip reliably at 4X or better. This means that a four-minute song should take less than one minute to rip. Only a handful of drives, such as those made by Plextor, can rip reliably at their full rated speed. If you plan to do a lot of ripping, the Plextor drives are certainly worth the $200 or so you will spend for the drive plus a SCSI controller.

You may notice that some tracks rip faster or slower than others. This is normal for CAV drives (see Chapter 8, *Choosing the Right Hardware*) because CDs are read from the inside out and CAV drives spin at a constant RPM. CAV drives will rip slower on the inner (lower number) tracks and faster on the outer (higher numbered) tracks. CLV drives should rip at similar rates on all tracks, but sometimes they have problems on the inner tracks. This is because they must increase the RPM to maintain the linear velocity, and vibration at the higher speeds can cause seek errors.

If you have a slow CD-ROM drive, you may be tempted to work on other tasks while it’s ripping, but you risk ruining the ripped file. This is another reason why it’s worthwhile it to get a fast SCSI CD-ROM like the Plextor 12/20 Plex, which can rip a full CD in under five minutes. And, if you do run other programs while it’s ripping, it will have fewer problems than an IDE CD-ROM drive because the SCSI interface puts less of a load on the system’s processor.

On a fast system, you might be able to get away with working in a spreadsheet or wordprocessor program while you are ripping, but you

---

**Tip**

If you have enough disk space, keep the WAV files until you are satisfied with the sound quality of the MP3 file. You may listen to a few seconds of a WAV file and think that it is good, but later on, when you listen to the MP3, you may find problems further into the file.
should still be careful, because a track can appear to rip successfully, but the WAV file will be full of errors and sound horrible.

**Disabling Auto Insert Notification**

The auto insert notification feature of Windows senses whenever a CD is inserted into the CD-ROM drive and automatically executes the instructions in the autorun.inf file or starts playing it if it is an audio CD. Auto insert notification can interfere with ripping and recording CDs, so it should be disabled.

To disable auto insert notification, click the Start button, select Settings, then Control Panel. Double-click on the System icon, then select the Device Manager tab. Click on the plus (+) symbol to the left of the CD-ROM icon. This will display a second CD-ROM icon. Highlight the second CD-ROM icon and click Properties. Select the Settings tab and uncheck the Auto insert notification box.

**Rip Offset**

Some CD-ROM drives have problems recording the first or last tracks of CDs because the drive inaccurately reads the start and end times by a few frames (sectors). AudioCatalyst and MusicMatch allow you to enter
an offset value to compensate for this. Generally a value of 10 to 30 will work. Higher numbers may cut into the audio data.

**Ripping Under Windows NT**

To rip under Windows NT you must have an ASPI driver installed. If you have a SCSI controller, ASPI is probably installed on your system. ASPI was originally developed for SCSI devices but also works with some IDE CD-ROMs.

ASPI versions 4.01 through 4.53 will not work with most IDE drives. To see which version of ASPI is installed on your system, find the file wnaspi32.dll in the \Windows(or \WinNT)\system32 directory and right-click on it. Select Properties, then select the Version tab and click on Product Version.

**Ripping to WAV Files or Direct to MP3?**

Ripping usually produces a WAV file, although some software can rip and create an MP3 file in one operation. The advantage of ripping to a WAV file is that it can be edited to adjust the volume or to trim off silence. A WAV file can also be used to encode several MP3 files at different bit-rates without the need for the original CD. Ripping directly to an MP3 file is a bit riskier than ripping to a WAV file because it is more taxing on your system and there are more things that can go wrong.

**File Names**

Many rippers can use information from the CDDB to automatically name files in a way that makes sense. For example, Billy Idol-White Wedding.WAV is easier to remember than Trk01.WAV. Encoders usually keep the same name and just add the .MP3 extension.

Some rippers provide you with full control over file names and folders, although a few only allow you to specify the folder where the files will be stored. Unless you specify otherwise, or if the information can’t be obtained from the CDDB, most rippers will create file names based on the CD identifier and/or track number (e.g., 5b9a97_Trk01a.WAV).

See Chapter 9, *Organizing and Playing Music*, for more information on file names and the CDDB.
Verifying the Quality

Before ripping your entire music collection, you should rip a few tracks to WAV files and listen to them to verify that the results are satisfactory. Most problems with WAV files that weren’t ripped properly are fairly obvious, even when played through computer speakers. Clicks and pops or phasing noises are signs that the audio data got scrambled during the ripping process.

Once you have successfully ripped a few WAV files and are satisfied that your configuration is optimal, you may want to try ripping a few test tracks directly to MP3 (if your software supports this). If the tracks ripped directly to MP3 sound OK, then you can probably rip your whole collection this way.

Error Indicators

A good ripping program will warn you of any errors during ripping and include tools for determining if the rip was successful or not.

MusicMatch uses a color-coded indicator to tell you whether or not the track was ripped successfully. A green indicator means the track was ripped with no errors, yellow means there were some errors (but the entire track was ripped), and red means the process was aborted.

AudioCatalyst warns of “possible speed problems” and outright failures but does not verify if the rip was successful. It does calculate a checksum for each ripped track, but this does not necessarily mean the rip was successful. Unless you have a checksum from the same track that was ripped successfully to compare, there’s no way to tell if the checksum is valid. Even then, the checksum could be different because of a seek error at the beginning of the track, which would not affect the quality.

With a good CD-ROM drive, if you rip the same track more than once, the checksums and the WAV files should always be identical. Ripping the same track twice and comparing the files or checksums is a good way to test the accuracy of a CD drive and the effects of different ripping settings.

AudioCatalyst has a file comparison feature for this purpose. If the WAV files or checksums for the same track do not match exactly, then at least one of them has errors. Sometimes the WAV files may be
different, but still usable if the differences are just in the first few or last few bytes. The only way to be sure is to listen to the files.

Analog Ripping

Many CD-ROM drives are simply incapable of ripping. Some CDs may be scratched or otherwise damaged to the point where they cannot be ripped. If your CD-ROM doesn’t support ripping, or you have CDs that are so badly scratched they it won’t rip, you may have no other choice than to record them via analog (see Chapter 14, *Hard Disk Recording*).

Some programs refer to this as “analog ripping,” but that term is really a misnomer. Analog ripping is the same as any other analog recording process except that the source starts out as digital and is converted to analog, and then converted back to digital in the sound card.

MusicMatch will automatically switch to analog recording if it detects too many errors during ripping. AudioCatalyst and RealJukebox will also let you record CDs via analog. You can accomplish the same thing by playing the CD and using a program like CoolEdit or Sound Forge to record it, but it’s easier to use an all-in-one program.

**Additional Resources**

<table>
<thead>
<tr>
<th>Organization</th>
<th>Web Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDDDB</td>
<td><a href="http://www.cddb.com">www.cddb.com</a></td>
</tr>
<tr>
<td>CD Digital Audio Extraction Page</td>
<td><a href="http://www.tardis.ed.ac.uk/~psyche/cdda/">www.tardis.ed.ac.uk/~psyche/cdda/</a></td>
</tr>
<tr>
<td>CD Page</td>
<td><a href="http://www.cdpage.com">www.cdpage.com</a></td>
</tr>
</tbody>
</table>

**Figure 35 - Analog “Ripping” a Four-Minute Song**
Making Your Own MP3 Files

Since most music by major artists is not yet available in MP3 or other downloadable formats, you may want to create MP3 files from music you already own. Keep in mind that while it’s legal to create MP3 files from music you have purchased, it’s not legal to give these to your friends or post them on a Web site without permission (see Chapter 5, Digital Music and Copyright Law).

Figure 36 - Options for Creating MP3 Files

To create an MP3 file from a record, tape or other external source, you first record it and save it as a PCM WAV file. Then you convert it to MP3 with an encoding program. To create an MP3 file from an audio CD, you digitally extract (rip) the audio data to a WAV file and then convert (encode) it to MP3. Some programs like MusicMatch and RealJukebox can rip or record and encode in parallel.
Encoding

Encoding is the process of converting uncompressed digital audio into a compressed format such as MP3. Ripping (see previous chapter) just copies the audio from a CD to a WAV file (or AIFF file on a Mac). The underlying format remains PCM. Encoding actually converts the file to a new format. When you play the encoded file, it must first be decoded before being processed by the D/A converter in your sound card or portable player.

Compared to ripping, there are fewer things that can go wrong with encoding. However, there are more parameters that affect the file size and sound quality. The goal of ripping is to create a perfect copy every time. The goal of encoding is to create files with the best possible sound quality, given the constraints of file size or available bandwidth.

Encoding is a game of trade-offs between speed, file size and sound quality. MPEG provides programmers with a lot of flexibility as to how they implement a particular standard. The result is that there can be substantial differences in sound quality between different encoders, even for files encoded at the same bit-rate. The differences between encoders are usually more apparent in the sound quality they achieve at lower bit-rates. At higher bit-rates (above 256 kbps), it’s difficult to tell the difference between encoders.

Encoding audio is much more time consuming and processor-intensive than decoding it. A fast 486 computer might be adequate for playing encoded audio but would be unbearably slow for encoding it. A fast Pentium class system is needed to encode MP3 files in real-time. On a Pentium II 266Mhz system, a four-minute song can be encoded in about a minute. On a 486 system, this could take 15 minutes or more. Some newer formats, such as MPEG AAC, take even longer to encode because the algorithms used to achieve higher compression are more complex and require more processing time.

Xing Technologies makes a very fast encoder that is used by both MusicMatch and AudioCatalyst. The Xing encoder used with the VBR option produces files with a good balance of sound quality and file size. The Fraunhofer Institute was heavily involved in developing the standard for MP3 and makes an excellent MP3 encoder that many people believe produces superior sound quality at lower bit-rates. Unfortunately, it’s very expensive to license, so it’s found primarily in
higher priced ($100+) products. It’s also much slower than the Xing encoder.
Recording or Ripping Directly to MP3

Some programs, such as AudioCatalyst, MusicMatch and RealJukebox, allow you to rip and encode (or record and encode) in a single process.

Bit-rates

As mentioned in Chapter 11, *A Digital Audio Primer*, the term bit-rate refers to how many bits (1s and 0s) are used each second to represent a digital signal, and the bit-rate correlates directly to the size and sound quality of an MP3 file. MP3 files can be encoded at anywhere from 8 kbps to 320 kbps. Lower bit-rates result in smaller files, with reduced sound quality, and higher bit-rates result in better sound quality, but larger files. Table 22 shows the formula for calculating the file size for four minutes (240 seconds) of audio encoded at 128 kbps.

| Table Error! Bookmark not defined. - Calculating File Sizes for Encoded Audio |
|-----------------------------------------|-----------------------------------|-----------------------------------|---------------------------|
| (Bit-rate x Length) / (Bits/Byte x 1KB) / 1024 = File Size                  |
| (128,000 x 240) / (8 x 1024) / 1024 = 3.8MB                                  |

Sound Quality

Because sound quality is subjective and varies depending on the encoding program, the bit-rate required to obtain a certain quality level will vary. At higher bit-rates (193 kbps+) on typical home stereo systems, many people will not be able to tell the difference between the encoded file and the original (although audiophiles with good ears and expensive systems often can).
**Table**  Comparable Sound Quality for MP3 at Different Bit-rates

<table>
<thead>
<tr>
<th>Bit-rate</th>
<th>Mode</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 kbps</td>
<td>Mono</td>
<td>Telephone</td>
</tr>
<tr>
<td>16 kbps</td>
<td>Mono</td>
<td>Short-wave Radio</td>
</tr>
<tr>
<td>32 kbps</td>
<td>Mono</td>
<td>AM Radio</td>
</tr>
<tr>
<td>64 kbps</td>
<td>Stereo</td>
<td>FM Radio</td>
</tr>
<tr>
<td>128 kbps</td>
<td>Stereo</td>
<td>Near CD</td>
</tr>
<tr>
<td>256 kbps</td>
<td>Stereo</td>
<td>Equal to CD</td>
</tr>
</tbody>
</table>

**Other Encoders**

Many freeware MP3 encoders are available that are quite capable. Blade and Lame are two that are fairly popular. BladeEnc is a freeware and supports multiple platforms, including Linux and Solaris. Lame is an open source encoder that is available only as source code.

MP3 at 128 kbps is sometimes referred to as CD-quality. But anyone who refers to 128 kbps as CD-quality either has poor hearing or has never done a side-by-side comparison. There are so many other factors, such as the type of music and the encoding algorithm that affect sound quality, it’s best not to infer that a specific bit-rate is equal to a certain quality level. Fortunately, mainstream MP3 encoders like Xing continue to improve, and if MPEG AAC encoders become available for a reasonable price, CD-quality sound at 128 kbps will become a reality for the average user.

Hard disk capacity is so cheap these days that it doesn’t make sense to use lower encoding rates just to save disk space. Today you can purchase a 16GB hard disk for under $200. With variable bit-rate encoding at the normal/high setting, this disk would hold over 3,000 songs. That’s more songs than in many people’s CD collections (especially if you count only the songs they actually listen to).

Lower bit-rates may produce files that sound fine with your present stereo system, but someday you may own a higher-end system where you can tell the difference. It’s less work to use higher rates now than to
have to recreate the files at a later date. Lower rates do make sense for music intended for a portable player like the Rio (where storage space is at a premium) or for voice recordings (where high frequency response is less critical).

**Constant Bit-rate Encoding**

Constant bit-rate (CBR) encoding uses the same number of bits each second to record a section of silence as it does to record a complex passage of music. This is like taking out the trash every day regardless of how full it is.

An advantage of constant bit-rate encoding is that it will always produce a predictable file size. The file size can be determined by multiplying the bit-rate by the length of the song in seconds (see Table 22). A disadvantage of CBR is that bits will be wasted on simple or quiet passages, when they would have been better used for more complex passages.

**Variable Bit-rate Encoding**

Variable bit-rate (VBR) encoding uses more or less bits per second, depending on the complexity of the signal. The encoder takes bits away from where they are needed least and puts them where they needed most—in the more complex sections. This is like emptying the trash only when it’s full. In general, VBR will produce significantly better sound quality than CBR at a similar rate.

With VBR, the file size will vary depending on the complexity of the music. A slow song consisting mostly of vocals and simple rhythms will result in a smaller, yet higher quality file than if it was encoded at a comparable constant bit-rate. Rock, jazz or other more complex music usually requires more bits, which results in larger files. VBR also produces files with a more constant signal-to-noise ratio than CBR.

A disadvantage of VBR is that many portable players will not report song lengths and elapsed times properly. VBR is also difficult to stream over a network because most streaming protocols allocate fixed bandwidth for each channel. VBR uses slightly more processing power than CBR during both encoding and playback, but this should not be a problem on systems with G3 or Pentium II (or better) processors.
VBR can be set to produce several levels of quality ranging from low to high. Naturally, the higher settings also produce larger files. AudioCatalyst and MusicMatch use different terms to describe their settings for VBR, which seems odd because they both use the same encoder. MusicMatch provides settings from 1 to 100 in increments of 5 to 10 and AudioCatalyst provides a range of settings from low to high. Table 24 shows similar VBR settings for both programs, along with comparable constant bit-rates.

<table>
<thead>
<tr>
<th>AudioCatalyst VBR Setting</th>
<th>Music Match VBR Setting</th>
<th>Comparable Constant Bit-rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>25</td>
<td>96</td>
</tr>
<tr>
<td>Normal/Low</td>
<td>40</td>
<td>112</td>
</tr>
<tr>
<td>Normal</td>
<td>50</td>
<td>128</td>
</tr>
<tr>
<td>Normal/High</td>
<td>75</td>
<td>160</td>
</tr>
<tr>
<td>High</td>
<td>100</td>
<td>192</td>
</tr>
</tbody>
</table>

*Average bit-rates for VBR will vary according to the type of music.

**Selecting the Best Bit-rate**

Even though the optimum bit-rate depends on many factors, the generalizations below will help new users better understand the options. Advanced users should experiment with different rates to find what works best for them.

With MP3, fair quality sound can usually be achieved with VBR set to **normal**, or with CBR at 128 kbps. Music with complex passages and wide stereo separation will benefit from higher rates. All types of music will benefit from VBR. Rock music usually sounds fine with CBR at 128 kbps, or with VBR set to **normal**. For jazz or classical music, it’s better to use a CBR of at least 192 kbps or VBR set to **normal/high**.

If you listen to a lot of classical or complex music, you should use even higher bit-rates. For critical listening, or archiving music with MP3, I recommend using either a CBR of at least 256 kbps or VBR set to high quality. For AAC, I recommend using a bit-rate of 128 kbps with the “Main” profile. At these rates, most people can’t tell the difference between the encoded file and the original source.
Lower bit-rates make sense if you will be listening to your music on a portable player. A 64MB player can hold about 74 minutes of audio at 128 kbps. At 80 kbps, it could hold almost two hours. Because the earphones and speakers on most portable units limit the quality of the sound, bit-rates much higher than 96 kbps would be wasted. Table 25 shows how many four-minute songs would fit on various types of media at different bit-rates.

<table>
<thead>
<tr>
<th>Bit-rate</th>
<th>4-min. Songs/32MB RAM</th>
<th>4-min. Songs/64MB RAM</th>
<th>4-min. Songs/CD (650MB)</th>
<th>Hours/CD (650MB)</th>
<th>4-min. Songs/Gigabyte</th>
</tr>
</thead>
<tbody>
<tr>
<td>64 kbps</td>
<td>17</td>
<td>34</td>
<td>355</td>
<td>24</td>
<td>546</td>
</tr>
<tr>
<td>80 kbps</td>
<td>14</td>
<td>28</td>
<td>284</td>
<td>19</td>
<td>437</td>
</tr>
<tr>
<td>96 kbps</td>
<td>11</td>
<td>22</td>
<td>237</td>
<td>16</td>
<td>364</td>
</tr>
<tr>
<td>128 kbps</td>
<td>9</td>
<td>18</td>
<td>177</td>
<td>12</td>
<td>273</td>
</tr>
<tr>
<td>160 kbps</td>
<td>7</td>
<td>14</td>
<td>142</td>
<td>9</td>
<td>218</td>
</tr>
<tr>
<td>192 kbps</td>
<td>6</td>
<td>12</td>
<td>118</td>
<td>8</td>
<td>182</td>
</tr>
<tr>
<td>256 kbps</td>
<td>4</td>
<td>8</td>
<td>89</td>
<td>6</td>
<td>137</td>
</tr>
</tbody>
</table>

Some programs can export the files listed in a playlist to a portable player in one batch process. MusicMatch can export playlists to the Rio and has an option to convert the files to a lower bit-rate before they are exported. This feature is handy because you don’t have to maintain two versions of the same song at different bit-rates.

**Verifying the Results**

It’s important to listen to the first couple of songs after they are encoded before you take the time to process your entire music collection. If the MP3 file doesn’t sound right, rip the same track to a WAV file and listen to it. If the WAV file sounds okay, the problem most likely happened during the encoding process. If the WAV file sounds funny, then you may need to adjust the jitter correction settings in the ripping software.

**A/B Comparisons**

The best way compare sound quality of the original and the encoded file is to do an A/B comparison, where you can play both simultaneously, and quickly switch back and forth.
An A/B comparison between a computer and external stereo system works well because you can sync up the two sources and rapidly switch back and forth. This is important because the “acoustic memory” of most people is very short lived—especially for subtle differences in sound. Another good test is to compare the sound of the WAV file to the MP3 file and switch back and forth by using two instances of Winamp or one of the DJ mixer programs listed in Chapter 7, *Choosing the Right Software*.

Some inexpensive CD-ROM drives use poor quality D/A converters and can cause a CD to actually sound worse than MP3 files encoded from it. If the source is a CD, it’s best to play it on an audio CD player and listen to it and the MP3 file through a good set of headphones or stereo speakers, and use an A/B switch to change between the two sources.

Table 26 lists a few of the more popular MP3 encoders. Check out the MP3 sites listed in Appendix A, *Interesting Web Sites*, for information on other encoders.

<table>
<thead>
<tr>
<th>Encoder</th>
<th>Developer</th>
<th>Web Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>BladeEnc</td>
<td>Tord Janssen</td>
<td><a href="http://home8.swipnet.se/~w-82625">http://home8.swipnet.se/~w-82625</a></td>
</tr>
<tr>
<td>MP3 Producer</td>
<td>Fraunhofer/Opticom</td>
<td><a href="http://www.opticom.de">www.opticom.de</a></td>
</tr>
<tr>
<td>MPecker (Mac)</td>
<td>Rafael Luebbert</td>
<td><a href="http://www.anime.net/~go/mpeckers.html">www.anime.net/~go/mpeckers.html</a></td>
</tr>
<tr>
<td>Lame</td>
<td>N/A (Open-Source)</td>
<td><a href="http://www.sulaco.org/mp3">www.sulaco.org/mp3</a></td>
</tr>
<tr>
<td>Xing</td>
<td>Xing Technologies</td>
<td><a href="http://www.xingtech.com">www.xingtech.com</a></td>
</tr>
</tbody>
</table>
Editing Sound Files

Some WAV files will benefit from a bit of clean-up before they are encoded. The most common forms of clean-up are trimming silence from the ends of songs, removing unwanted noise and normalizing the volume so all songs will play at similar levels. More sophisticated users may want to add fades, equalization or dynamic range compression. These effects change the nature of the music, so they should be used sparingly unless you are mixing a recording of your own music.

Files must be uncompressed (PCM WAV or AIFF format) before they can be edited. If the song is already in MP3 format, you can convert it to WAV format, edit it, and then convert it back to MP3. Each time you do this you will lose fidelity. Files encoded at higher (192 kbps+) bit-rates will lose less fidelity during the decoding/re-encoding cycle than those encoded at lower bit-rates.

Note: A few programs, such as MP3 Trim, allow you to edit MP3 files directly. These programs are typically limited to just trimming silence and normalizing the volume.

Sound Editing Software

Cool Edit and Sound Forge are popular programs that can record and edit sound files. Both programs are available in professional and “lite” versions. The lite versions include the basic features for trimming silence, adding fades, normalizing the volume and removing noise. The professional versions offer high-end capabilities for recording engineers and are overkill for most users.

AudioCatalyst and MusicMatch versions 4.0 and higher can normalize the volume automatically when the file is ripped. AudioCatalyst can also automatically trim silence from the ends of each track, and MusicMatch 4.0 can automatically apply fade-ins and fade-outs. For more extensive editing, you should use specialized sound editing software, such as CoolEdit or Sound Forge.

Normalization

Many CDs do not use the full 96 dB dynamic range that’s available. This can result in songs from some CDs playing much louder than others, even at the same volume setting. Normalization corrects this by scanning the uncompressed audio file to determine the peak or average level and proportionally increasing or reducing the levels throughout the file to obtain the desired volume level.

Prerecorded CDs that were digitally remastered from analog tapes are more likely to require normalization than CDs that were originally digitally mastered, but there is no hard and fast rule. Normalization is often needed for WAV files created from records and tapes.

Some rippers and most sound editing programs include a normalization feature. The Audiograbber ripper has the most flexible normalization feature of all the programs I’ve used, although some people might find it too complicated. The normalization features in Sound Forge XP and Cool Edit 96 are very basic and require working with each file.
individually. The professional versions of Sound Forge and Cool Edit are much better, but they are expensive and neither is as flexible as Audiograbber.

AudioCatalyst, which is based on Audiograbber, has a simple normalization feature that allows you to normalize all files to a set level, or to normalize only the files where the peak level is lower or higher than the thresholds you specify. MusicMatch Version 4 has a very limited normalization feature that adjusts all tracks to a single level.

Most normalizers allow you to specify a percentage of the maximum possible level for the highest peak. The maximum level may be referred to as 1, 100% or 0 dB, depending on the software. A setting of 50% (or .5) would be the same as –6 dB, because each doubling or halving of the signal level represents a change of 6 dB. A value of 100% (0 dB) will normalize the volume so it covers the full dynamic range, so the highest peak will be at the maximum level. Values above 100% should not be used because this will cause clipping wherever the level exceeds 0 dB.

Generally, all songs on a prerecorded CD will be recorded at about the same level, so you can assume that if the level needs to be adjusted for one song, the same adjustment will be needed for all the songs.

**Normalization Settings in AudioCatalyst**

If you have a CD collection that contains a variety of music, there’s a good chance that some songs will sound louder than others even when normalized to the same level. This occurs because the average volume may be different even though the peak levels are similar. Other factors such as differences in frequency content (especially with higher frequencies) and recorded distortion (electric guitar effects, synthesizers, etc.) will affect the apparent loudness of a song.

Listening is the best way to judge the appropriate level, but it takes time to listen to every song and normalize each one individually. An approach that works for most CDs is to normalize all songs lower than 91% or higher than 98% to a 97% level. Table 27 shows the results of normalizing four songs using these settings.
If you want more control over normalization, Audiograbber provides advanced settings based on either average or peak levels, along with an option for dynamic range compression. Normalizing based on average levels will make the playback levels more consistent. However, raising the average level can easily cause clipping.

Audiograbber can be set to automatically apply dynamic range compression all of the time, or only when it’s needed to avoid clipping. Audiograbber can also be set to not compress songs that are already highly compressed. (See Chapter 11, *A Digital Audio Primer*, for an explanation of the difference between dynamic range compression and “compaction” type compression.)

If you have MP3 files that have not been normalized, you can use a player like Winamp to store preset equalization settings for each song. With the “auto” mode of its equalizer enabled, Winamp will read the setting for the song when it’s played and adjust the level accordingly. It’s still better to normalize the WAV file before you encode it, because most portable and dual/mode CD players can’t compensate for individual songs recorded at different levels.

### Transitions Between Songs

Professional DJs and anyone who records tapes for their own listening pleasure understand the importance of having a smooth flow of music. Playlists eliminate the need to swap records and CDs, but the problem of transitions between songs still exists.

Transitions between some types of songs need to be handled differently. For background music it’s OK to have a few seconds of silence between songs, but for dance music it’s usually better to crossfade between songs with no silence in between.

### Trimming Silence

Many songs have a few seconds of silence at either end. Trimming off this silence will make the files smaller and allow for a more continuous flow of music. In most cases, you should leave anywhere from ¼ to ½ second of silence at the end of each song, unless you are making a continuous mix of dance music, in which case you’ll want no silence.

Many songs have excessively long intros or trailing sections of music or vocals that can be removed. If you remove one of these sections you should add a fade-in or fade-out so the song does not start or stop abruptly.
It usually sounds bad to have one song end abruptly and the next one start immediately, except for dance mixes where the songs have the same number of beats per minute. Fade-ins and fade-outs can be applied to the ends of the songs to provide a smoother transition, similar to the way a DJ would use a crossfade.

Fades can be created with a sound editing program, but it’s important to remember that these will be permanently stored in the file. Winamp has a crossfade plug-in that works without modifying the file. It works fairly well if you play most of your music on a computer, but if you use a portable or dual-mode CD player, you will need to create the fades by editing the WAV files. (See the tutorial for Cool Edit 96 in Chapter 20, for instructions on trimming silence and adding fades to a WAV file.)

Playing a song that slowly fades in immediately after a song that slowly fades out may result in too long of a lull in the music. One way to avoid the lull is to steepen the fade-in and fade-out slopes. Most editors give you several ways to control the “envelope” of the slope. Usually, you highlight the section of the file where the fade is needed, then you either graphically adjust the slope of the fade, or specify the starting and ending volume levels. Most sound editing programs include preset fade envelopes and the ability to define and save custom envelopes.

If you splice files together to make a continuous dance mix, you may have a hard time getting the beats to match exactly. One way around this is to use a sound editor to reduce or increase the tempo of one file to match the other and then splice them together so the beats match. The tempo adjustment feature stretches or compresses the length of a song, which effectively changes the tempo. DJ mixer programs like those by VisioSonic have built-in features for matching tempos of songs.

Optimizing Audio for the Web

Sound editing software can also be used to convert digital audio to different formats and to optimize audio files for use on the Web as downloadable music or streaming audio.

Internet access is slow for many people. Currently, most people connect to the Internet via a 33.6 kbps or slower analog modem. Even with a 56k modem, users are lucky to achieve connection speeds of more than 48 kbps. If you want the broadest possible audience for your Web site, it’s best to assume a “lowest common denominator” connection speed of 28.8 kbps.

Downloadable Formats

Compressed formats like MP3 are a good choice for just about any type of downloadable music. Uncompressed formats can be used for very short clips but should not be used for
full-length songs. The advantage of using an uncompressed format like PCM Wave is that most Web browsers will be able to play it without special software.

You can choose from many different formats when adding downloadable music to your Web site, but you should stick with popular formats as much as possible. Otherwise, you risk losing users who may not want to install yet another player to support some proprietary format. Table 28 lists some of the more common formats for downloadable music.

<table>
<thead>
<tr>
<th>Type</th>
<th>Extensions</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP3</td>
<td>.mp3</td>
<td>MPEG Layer-III</td>
</tr>
<tr>
<td>MS Audio</td>
<td>.asf, .wma</td>
<td>Proprietary (Microsoft)</td>
</tr>
<tr>
<td>QuickTime</td>
<td>.qt</td>
<td>Proprietary (Apple)</td>
</tr>
<tr>
<td>RealAudio</td>
<td>.ra, .ram</td>
<td>Proprietary (Real)</td>
</tr>
<tr>
<td>WAV</td>
<td>.wav</td>
<td>PCM*</td>
</tr>
</tbody>
</table>

"WAV files can use other formats besides PCM.

Small Is Beautiful

Your goal with Web audio is to create the highest quality sound file, at the smallest possible size, in the most commonly readable format. You can reduce file sizes (and bandwidth requirements) of both compressed and uncompressed audio. Even if you plan on using a compressed format like MP3, it still makes sense to tweak the uncompressed audio file to make it smaller before it is encoded. This will result in a smaller encoded file, as well.

The type of material, and desired sound quality are the two main factors to consider in optimizing an audio file. For example, for sound effects and voice, the sampling rate (which determines the frequency response) and resolution don’t need to be as high as required for music. Table 29 shows different combinations of sampling rates, resolution and channels that are appropriate for various types of uncompressed audio.

Stereo or Mono?

Is stereo necessary for the type of audio you are using? Certainly it is, if you are working with CD-quality music. For short clips and voice, using mono will cut the file size in half. Mono is also fine for many sound effects and background music.

16 Bits or 8 bits?

You can reduce the resolution from 16 to 8 bits and cut the file size in half again, but the signal will have more distortion from quantization errors (because it cannot be recorded as precisely with fewer bits). The difference between 8-bit and 16-bit resolution will be more noticeable in complex music with a wide dynamic range. For voice and sound effects, 8-bit resolution is usually adequate.
### Sampling Rate

CD audio is sampled at 44.1 kHz and can reproduce frequencies up to 20 kHz. Most people can’t hear frequencies above 16 kHz. For music on the Web, you could use a sampling rate of 22.5 kHz, and many people will not notice any difference when using typical computer speakers. For higher quality music, a 32 kHz sampling rate can be used in place of 44.1 kHz and many people will not be able to tell the difference, even with a good speaker system. For voice, you can reduce the sampling rate to 11.25 kHz and it will usually sound fine.

### Streaming Audio

Streaming audio is optimized by the streaming server and is usually compressed to deliver a higher bit-rate over slow Internet connections. Some streaming systems, such as RealNetwork’s SureStream technology, automatically optimize the bit-rate of each stream to the speed of the user’s connection. Other systems may need to use a different streaming server for each bit-rate.

For streaming audio to work well, the speed of your Internet connection must be greater than the bit-rate of the sound file. The Internet is designed to send data in scattered bursts. Good audio playback requires audio data to be delivered continuously, at a constant rate.

To allow for network congestion, the bit-rate should be no more than two thirds the available bandwidth. For instance, 128 kbps is considered the minimum bit-rate for good quality MP3 files, but this is much higher than the bandwidth that any analog modem can deliver. A bit-rate of 15 to 20 Kbps would be more appropriate for a 28.8 or 33.6 modem.

### Table: Streaming Media Systems

<table>
<thead>
<tr>
<th>Type</th>
<th>Primary Format</th>
<th>Developer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows Media</td>
<td>Active Streaming Format (ASF)</td>
<td>Microsoft</td>
</tr>
<tr>
<td>Technologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Icecast (open source)</td>
<td>MP3</td>
<td>The Icecast Team</td>
</tr>
<tr>
<td>QuickTime</td>
<td>QuickTime</td>
<td>Apple Computer</td>
</tr>
<tr>
<td>RealSystem</td>
<td>RealAudio</td>
<td>RealNetworks</td>
</tr>
<tr>
<td>SHOUTcast</td>
<td>MP3</td>
<td>Nullsoft</td>
</tr>
</tbody>
</table>
To listen to streaming audio at 128 Kbps, even a dual channel ISDN connection would be just barely enough. A higher speed connection like a cable modem or an ADSL (Asynchronous Digital Subscriber Line) connection would be required.

For short promotional clips of music, you should consider offering more than one format so your site will appeal to a wider group of users. It is becoming more common to find sites offering audio clips in multiple formats, such as MS Audio, RealAudio and streaming MP3.

The major streaming media systems support multiple formats, including MP3. But if you only offer streaming audio in a proprietary format like RealAudio or ASF, you risk losing users who may prefer to use a player like Winamp to listen to streaming audio. Table 30 lists some of the more common streaming media systems.

For more information on integrating audio into a Web site check out the book *Audio on the Web* by Jeff Patterson and Ryan Melcher, listed in Appendix C, Recommended Reading.

**Sound Editing Utilities**

Programs for directly editing MP3 files, processing WAV files and creating HTML interfaces for MP3 CDs are described below.

**MP3 Trim and Wave Trim**

MP3 Trim and Wave Trim ([www.jps.net/kyunghi/mp3encod.htm](http://www.jps.net/kyunghi/mp3encod.htm)) by Jean Nicolle are handy for editing large batches of MP3 and WAV files. Both programs are available in shareware and professional versions.

MP3 Trim allows you to edit MP3 files without decoding them to WAV format. It can detect and remove digital silence and truncated frames to recover wasted space. Without MP3 Trim, you would have to decode the MP3 file to WAV format, trim the silence with a program like CoolEdit, and then convert it back to MP3.

Wave Trim scans the first and last 10 seconds of a WAV file and removes the digital silence and any bits of audio left over from other tracks. Wave Trim also can normalize the volume, and accept command-line parameters for batch processing. Batch processing allows you to process a lot of songs without opening and editing individual WAV files.

**MPEG DJ GoWave**

MPEG DJ GoWave from Xaudio ([www.xaudio.com](http://www.xaudio.com)) is a program for decoding MP3 files into WAV format. This is only necessary if you need to normalize the level or otherwise edit the sound, and you do not have the original CD or WAV file.
MP3 Prepare

MP3 Prepare (http://aryhma.pspt.fi/download/mp3prepare.html) is a program for making HTML-based user interfaces for MP3 CDs. It can automatically create an HTML page that lists all albums on the CD and a playlist that includes all the songs. MP3 Prepare can also create separate HTML pages for each album, and include cover art, song titles and play times. MP3 Prepare features a range of fully customizable graphics and HTML templates, and includes MP3 Browser, which is an Explorer-like interface for browsing MP3 files on your hard disk. MP3 Browser can also display and edit ID3 information and generate playlists.

Ray Gun

Ray Gun from Arboretum Systems (www.arboretum.com) is a shareware program for removing noise from WAV files and adjusting signal levels. It can also be used to record audio. Ray Gun is available in both Windows and Macintosh versions and can be used as a plug-in for recording programs like CoolEdit and Sound Forge that support DirectX.

Recording Your Own CDs

It’s great to have hundreds of hours of music stored on your PC, but how do you take all that music on a trip? And what happens if the 17 gigabyte hard drive holding your entire MP3 collection crashes? You could use a tape drive to backup your MP3 files and make them portable, but tape is a linear media—which is very slow for accessing individual files. A better solution is to use a CD Recorder and “burn” your own CDs.

CDs are a good solution for storing MP3 files and backing-up moderate amounts of data. CDs are more portable than tapes and support random access, so any file can be accessed quickly regardless of where it’s located on the CD. Another advantage is that most CDs can be read by virtually any CD-ROM drive, unlike tapes, which come in dozens of different incompatible formats.

CD Construction

Prerecorded CDs are created by a pressing process similar to the process used to create vinyl records. A pattern of pits and lands (raised areas) that correspond to the 1s and 0s of binary data is pressed into the disc. The difference in reflectivity between the pits and the lands is sensed by the laser in the audio CD player or CD-ROM drive and converted to a digital signal.

Recordable CDs

Recordable CDs (CD-Rs) are “burned” with a CD recorder (also called a CD writer). A blank CD-R disc contains a pre-grooved spiral track that guides the recorder’s laser as it burns a microscopic series of holes called pits in a layer of organic dye. The pattern of pits and lands (the unburned part) encodes the information in the same manner as the pattern stamped on a pre-recorded CD. Once recorded, data on a CD-R disc can’t be erased.
Rewritable CDs
Rewritable CDs (CD-RWs) are similar to CD-Rs but can be erased and re-recorded thousands of times. Most audio CD players and many older CD-ROM drives can’t read CD-RWs. This is because the amount of laser light reflected from the recording layer of a CD-RW is much lower than that of a CD-R.

Instead of burning a pit in the recording layer like a CD-R, the recording laser in a CD-RW recorder causes a phase change (from crystalline to amorphous state) in the recording layer. The different states act as the pits and lands do on a CD-R disc and are detected by the difference in the way they refract light.

CD-RW drives cost slightly more than CD-R drives but can also burn CD-R discs. As of this writing, you can purchase a 4X CD-R drive for under $100 and a 2X CD-RW drive for under $150.

CD Recorders
CDs can be created on computers with CD-R or CD-RW drives, on dedicated CD recorders that are designed to work with home audio equipment, and with stand-alone recorders designed for mass duplication. CD-R and CD-RW drives used with computers can also function as CD-ROM drives, although they tend to be slower because of the heavier laser mechanism required for recording.

Why Record Your Own CDs?
CD recording is a convenient way to create standard audio CDs or MP3 CDs with custom mixes of music. If you’re in a band, CD recorders provide a low-cost way to create demo CDs of your songs.

Write Speeds
The write speed of a CD recorder determines the time it will take to record a CD. Write speeds are measured in the same “X” units (2X, 4X, 8X, etc.) that are used to measure a CD-ROM drive’s read speeds.
The combined read/write speed of CD-R recorders is specified by the read speed, followed by the write speed, for example, 12X/4X. CD-RW recorders often have different write speeds for CD-RW and CD-R discs. The speed of a CD-RW drive is specified by [CD-ROM read speed]/[CD-R write speed]/[CD-RW write speed], for example, 20X/4X/2X.

Recording a full audio CD at 1X speed takes at least 74 minutes plus the time required to locate files on the hard drive and about 2 minutes to write the table of contents. Recording a CD at 2X takes at least 37 minutes, 4X takes at least 18 minutes, and 8X takes at least 9 minutes.

**CD Standards and File Systems**

Standards define the way different types of information, such as audio, video or data, are stored on a CD. Currently, there are more than 10 different standards for CDs. Many of these, such as CD-I (Compact Disc-Interactive), were designed for use with proprietary players that combined audio and text or graphics data, and never caught on. The three main formats currently used for audio and data CDs are named for the color of the standards books that describe them.

**Red Book Audio**

The *Red Book* standard was the original format developed for storing music on CDs. This standard is also referred to as CD-DA (Compact Disc-Digital Audio). Audio CDs have the advantage of being playable almost anywhere, though the capacity is limited to 74 minutes of music (approximately 18 four-minute songs).

Audio CDs contain only the digital data for the music, plus a table of contents (TOC) with track numbers and the starting position and lengths of each track. The TOC does not contain any information about the artist, album or song titles, although this information can be obtained from the CDDB. (See Chapter 9, *Organizing and Playing Music.*)

**Data CDs**

The *Yellow Book* standard defines how data is stored on prerecorded CD-ROMS. The *Orange Book* standard is similar to the *Yellow Book* and defines the format for CD-Rs and CD-RWs. MP3 files and other compressed audio formats are simply data files, so they are stored on *Yellow Book* (pre-recorded) or *Orange Book* (CD-R and CD-RW) CDs.

Data CDs are limited to 650MB, due to the overhead of increased error correction. When used with a compressed format such as MP3, data CDs can hold many times more audio than *Red Book* CDs. These MP3/data CDs can only be played on PCs with CD-ROM drives or special dual-mode CD players with built-in MP3 decoders, such as the AudioReQuest (see Chapter 8, *Choosing the Right Hardware*).

**CD File Systems**

File systems describe how data files are stored and retrieved by different computer operating systems. The most common CD file systems are ISO-9660, Joliet, and HFS. ISO-9660 only supports file names up to eight characters long, plus an optional one to three character extension (filename.ext). ISO-9660 can be read by computers running DOS, Macintosh, OS-2, UNIX and all versions of Windows. Joliet (developed by Microsoft) supports file names up to 64 characters, but is only supported by Windows 95/98/NT and recent versions of Linux. HFS is only supported by the Macintosh.
A number of other CD-ROM and file systems standards exist that are beyond the scope of this book. For more information, consult the resources listed at the end of this chapter.

## CD Media

Blank CD-R and CD-RW discs are available from many different manufacturers. Depending on the type of dye and reflecting layer, each brand will be some combination of gold, silver, green and blue. Different brands are certified for 2X, 4X and 8X recording.

Many CD recorders will work better with some brands of media than others, and some audio CD players and CD-ROM drives will read some brands but not others. It’s a good idea to try several brands in both your recorder and player before buying large quantities.

Both CD-R and CD-RW discs are more sensitive to heat, humidity and direct sunlight than pre-recorded CDs, so don’t leave them in a hot car. The data on any type of CD is closer to the metal side of a CD than the plastic side, so be careful when labeling not to scratch it. Use a permanent felt-tipped pen or a circular CD label to avoid damaging the CD.

A good product for removing scratches from CDs is Wipeout ($14.99 – [http://activemarketplace.com](http://activemarketplace.com)) If the CD is scratched but sounds OK, the error correction is working. Be careful when polishing a CD, because polishing can actually introduce more errors if performed improperly.

## Media Cost

CD-RW discs are more expensive than CD-R discs, but not by much if you shop around. CD-R discs can be purchased for anywhere from 25 cents to $2 each, and CD-RW discs can be purchased for anywhere from $1.50 to $7.50. Prices vary widely depending on the quality, quantity and whether or not jewel cases are included. Higher cost does not necessarily equate with higher quality. I’ve had just as good luck with lower-cost brands as I’ve had with more expensive brands.

## Audio vs. Data CD-R Media

Two kinds of blank CD-R media are available. One type is intended for data, the other for audio. The discs for audio cost more than discs for data. The Audio Home Recording Act specifies that royalties must be paid on blank CDs marketed for home audio recording. These royalties are placed into a pool to compensate for the loss of royalties from illegal CD duplicating.

Many standalone CD recorders check for a special code present on blank audio CDs and will refuse to work with data CDs. CD recorders used on PCs aren’t required to look for this special code and can use the lower cost data CDs for audio as well as data. Blank CD-R discs labeled for audio usually cost $4 or more each.
CD Recording Software

Two types of programs can be used for to record CDs: stand-alone programs, such as Adaptec’s Easy CD Creator (for PCs) or Toast (for Macs), and packet writing programs, such as Adaptec’s Direct CD program. It’s best to use a stand-alone program for creating audio and MP3 CDs.

With stand-alone CD recording software, all setup, layout creation and recording is done through the program’s interface. Packet writing software is more like a driver that makes your CD recorder act like a floppy drive. This allows you to copy files or drag and drop them to the drive from Windows Explorer or Finder (Mac).

A packet-written disc can only be read on the system on which it was created until it is finalized (see below). Even then, some CD-ROM drives have problems reading packet-written discs. Packet writing can waste up to several hundred MB of the capacity of a disc, and is most useful with CD-RW discs used for data backup.

Before You Record

Successfully recording a CD requires a constant, uninterrupted stream of high-speed data. CD recording software places the data to be written in a small area of memory called a buffer. The CD recorder can draw data from this buffer at a constant rate while the software is busy reading files from the hard drive.

Mouse movement, network and Internet activity, virus scanners, screen savers, or anything that requires the processor’s attention can interfere with keeping the buffer full of data. A buffer underrun occurs if the CD recorder empties the buffer before it’s finished writing a track. This ruins the disc, creating what is called a “coaster.” A smoothly running system is essential to keeping the buffer filled.

Update Your Software, Drivers and Firmware

Check the Web site of the manufacturer of your CD recording software for downloadable updates or patches. A patch contains only changes to the software and will update the software without the need to reinstall it.

Also check the Web site of the manufacturer of your CD recorder to make sure you have the latest firmware. Firmware is the internal program that controls the CD recorder hardware. You’ll need a special program and instructions on how to upload the firmware to the recorder. If you have a SCSI CD recorder, check the Web site of the manufacturer of your SCSI controller to make sure you have the most current drivers.

Stabilize Your System

Before you attempt to record a CD, clean up your system by following the steps listed below.

- Delete any unnecessary files, empty the recycle bin and defragment (defrag) your hard disk. Also, make sure you have plenty of free space on your hard disk for temporary files.

- Disable auto insert notification for your CD recorder by right clicking on it in the Windows Device Manager, under the CDROM folder, and then clear the check box
for Auto Insert Notification under the settings tab. (See Chapter 15, Digital Audio Extraction, for instructions on disabling auto insert notification.)

- Exit all other programs and disable virus scanners and screen savers. Remove any non-essential programs from your Startup folder, and restart your computer.

If you performed the steps listed above and still have trouble recording, you may need to disable other programs or background processes. If you have a network card, you may need to temporarily disable it in the Windows Device Manager (see the Easy CD Creator tutorial in Chapter 20, Software Tutorials).
To see if any other programs are running in the background, press Ctrl-Alt-Del to start the Task Manager. If any programs other than Explorer and Systray are listed, you can terminate them by highlighting each one and selecting End Task. (You may need to do this twice for some programs.) You can also disable these programs using the Msconfig program that is included with Windows 98.

**Test Your Hardware**

After installing the CD recording software, be sure to run the system tests to characterize the performance of your hard disk and CD recorder. This helps the software optimize the data flow from the hard disk to the CD writer and determines the maximum write speed.

**Recording Options**

The following examples are taken from Adaptec’s Easy CD Creator program. Although the information applies to all CD recording software, some of the terms may vary. (See the Easy CD Creator tutorial in Chapter 20 for step-by-step instructions on burning a CD.)

A layout describes which data or audio tracks are to be recorded on the blank CD, and in what order. When you start recording, the recording software searches your hard drive and reads the data as it is writing to the CD. This only requires a small (less than 50MB) amount of temporary space on your hard drive but can cause problems if your hard drive is slow or fragmented.

An image file contains all of the data that will be recorded on the CD, exactly as it will be written, in one large file. This makes for more reliable recording, especially at high speeds, but requires as much free space on your hard drive as the CD you are recording (up to 750MB). Using an image file can significantly increase your success in creating CDs, especially at higher speeds.

To make an image file, create your layout and select Create CD Image File rather than Record CD. Although this adds an extra step, the extra time is minimal because the data files specified in the layout have to be read anyway. If you create multiple copies from an image file, the overall time will be less because the second and subsequent copies don’t have the overhead of gathering the data.

**Multi-Session CDs**

CDs can be multi-session, which allows you to mix audio and data on the same CD. This way you can have the CD audio and MP3 versions of your songs on the same CD. The audio tracks must be stored in the first session. To allow this, choose Close Session and Leave Disc Open when recording. Then you can record a second session with MP3 files.

You can record more than one data session on a CD, but normally you can view only the files added in the last session, unless you specifically select another session to be active. A better way to access data recorded in multiple sessions is to import the previous session into each new session when recording. To do this, right-click on the icon for the CD in the lower-left quadrant of the Data CD Layout screen and select Properties. In the Settings tab of the CD Layout Properties screen, select Automatically import previous session. To finalize a CD, choose Close Disc when you record the last session.
Be aware that making a multi-session disc takes about 23MB extra for the first session, and 16MB for each additional session. Take this into account when calculating how many files you’ll be able to fit on a CD.

**Track-at-Once vs. Disc-at-Once Recording**

Track-at-once allows the CD recorder to write one track at a time and turn off its laser while reading data for the next track. This allows the recording software as much time as it needs to read the data for the next track and can be more forgiving than disc-at-once recording. When recording audio CDs, track-at-once places a two second gap between each song.

Disc-at-once keeps the laser on for the entire recording and eliminates the requirement for the two second gap between audio tracks. Disc-at-once is also more demanding of your system because of the need for uninterrupted data. Some plants that perform CD duplication services require disc-at-once masters.
Test Writes

Most CD recording software can perform a test write to determine if the recording process is likely to succeed. This test goes through the entire process of locating, reading and sending data to the CD recorder, with the recording laser turned off. A test write takes the same amount of time as actually recording the disc. If the test write fails, you can troubleshoot the problem without wasting a disc.
Don’t Touch

Once you start recording, don’t touch your system or attempt to run any other programs. This is a good time to take a break and get away from your computer.

It’s a good idea to do a test write the first few times you record, especially at higher recording speeds. Prior to recording, you’ll usually have a choice of Test Only, Test and Create, and Create. Test and Create will automatically start recording the CD if the test phase is successful. Later, you can select Create when you are sure of your system’s performance.

Recording Audio CDs

If you want to play your custom CD in a standard audio CD player, you must create a Red Book format CD. Each song needs to be in a 16-bit stereo 44.1 kHz uncompressed WAV format. If you only have MP3 files, many programs, such as MusicMatch and Winamp, have an option to decode them into WAV files (see the MusicMatch and Winamp tutorials in Chapter 20).

Once you have the songs in WAV format, select the Audio tab of the CD layout. Browse for the WAV files in the upper pane of the window, then drag and drop them on to the lower pane. After you have added all of the files, you can rename or reorder them. Make sure to select the option to close the session or to close the disc. Otherwise, the CD will only be playable in the CD recorder and not in audio CD players.

Figure 39 - Recording Audio CDs from MP3 Files

Recording MP3 CDs

MP3 files are just data files as far as a CD recorder is concerned, so you will need to record them to a data format CD. Before you record the CD, your songs must already be encoded as MP3 files. (See Chapter 16, Making Your Own MP3 Files.)

Figure 40 - Recording MP3 CDs

To begin, select the Data tab of the CD layout window, then select File, then CD Layout Properties, then select the Data Settings tab. If the CD will only be used on Windows 95/98/NT machines, choose the Joliet format to allow long filenames. If you want to use the CD on a Mac or other system, choose the ISO-9660 format.
If you choose the ISO-9660 format and you have used any long filenames for your MP3 files, the CD recording software will truncate the names to conform to the 8+3 character limitation of ISO-9660. Sometimes this creates names that are difficult to decipher. To avoid this problem, rename the files before burning the CD.

Add your MP3 files using the explorer-like interface by dragging and dropping them into the lower pane of the window. Rename and reorder the files as needed, and click on the Record button. Since these are data CDs, you don’t need to close the CD to use it. On the Advanced tab of the CD Creation Setup screen, choose Close Session and Leave Disc Open if you want to add more files later. Otherwise, choose Close Disc to finalize the CD.

**Duplicating CDs**

There are varying opinions on the legality of making CD-R copies of prerecorded music. The RIAA maintains that it’s illegal to burn copies of a prerecorded CD, even for your own noncommercial use. However, the Doctrine of Fair Use (see Chapter 5, *Digital Music and Copyright Law*) can be interpreted as to allow it. In any case, it’s illegal to sell or give away CDs containing copyrighted music without authorization.

The Adaptec Easy CD Creator program includes a utility called *CD Copier Deluxe* for duplicating an entire CD. The files or audio tracks are read the same as if you were creating a CD layout manually. This requires a large amount of temporary file space on your hard drive and may not work with some CDs.
If you have a Plextor CD writer and a Plextor SCSI CD-ROM drive, you can use the Plextor Discdupe utility. Discdupe copies the entire CD “bit by bit” and can duplicate any type of CD, even non-PC formats. If your system is fast enough, the “on the fly” option doesn’t require much hard disk space. This option is more sensitive other processes running at the same time, so avoid the temptation to do anything else on the system while the CD is recording.