Digital Technology Basics

The basic of digital audio recording
Digital audio recording is inherently more complex than analogue recording. The reason for this is because, sound itself is an analogue or linear phenomena.

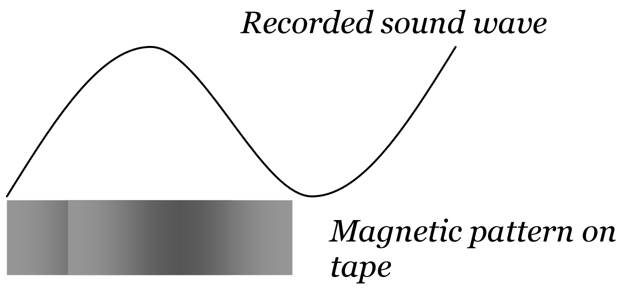
Changes in amplitude can be graphed along as a continuous straight line.

An intermediate value can always be inserted in between two adjacent values. For example, between 2 and 3 there is 2.5 and between 2.5 and 3 there is 2.75... etc so on into indefinite small increments.

In a digital system, however, only whole numbers can be used. The increments follow a decidedly step like pattern. There is no such thing as an intermediate value. eg.. between 2 and 3 there is nothing. It is either 2 or 3 – that's it.

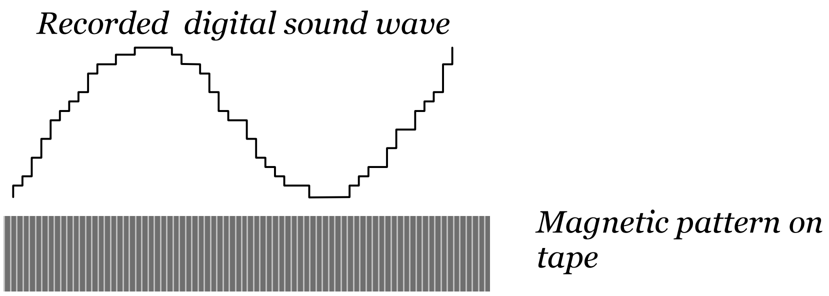
What is analog recording?
During an analog recording, sound wave picked up microphone is being converted to electric voltage, amplified to an appropriate level and fed to a recording head which in turn will produce a magnetic field that cuts across the magnetic tape passing underneath the recording head.

The magnetic particles on the tape consequently align themselves with this magnetic field and as this field varies in sympathy with the continually varying signal source from the microphone, the net result is a recorded piece of tape with magnetic particles of varying alignment spread over the length of the recording – an analog recording is born.



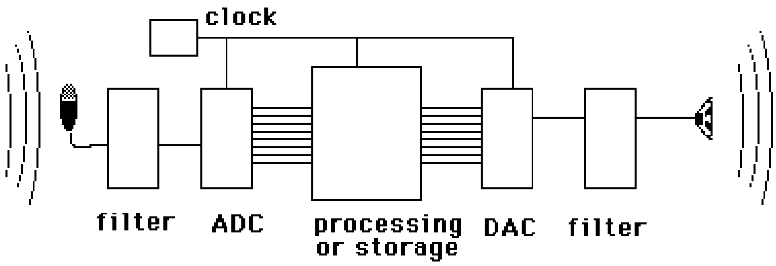
What is digital recording?

In digital recording, the signal to be recorded is convert to numbers of “0” and “1” (or digital). This “0” and “1” is then recorded on magnetic tape, CD, hard disk or other storage medium.



Comparison between the two

Analog Audio	Digital Audio
Obvious generation lost	No generation lost
Noise added during copying	No noise added during copying
No perfect copy can be made	Perfect copy can be made
Can only be stored on limited analog medium	Can be stored on large number of digital medium
Cannot be manipulated by computer	Can be manipulated by computer



The digital signal chain (basic diagram)

Numbering systems

Any value can be used as the base of a numbering system. But, converting from one numbering system to another can be fairly difficult.

You do not need to perform any mathematical calculations to understand digital audio.

The word DIGITAL implies the use of numbers. So, any digital process including digital audio recording is by definition a mathematical process of some kind.

The numbering system we are all used to working with is the decimal system. It is based on a value of ten (10). Decmal simply means based on ten.
In the decimal numbering system there are 10 available digits, ranging from zero to nine (0 – 9) and all decimal numbers must be made up of some combination of these ten digits.

In digital electronics, the binary system is normally used.
Binary means, 'based on two'. So in the binary numbering system there are just two digits, 0 and 1.

Every digital circuit has a fixed number of digits. If the circuitry is setup for 4 digits, all four places must be filled with either a 0 or a 1. No space can be left blank.

Each digital in a binary number is called a BIT or a Binary Digit.
Eight bits make up a BYTE.

Eg... 0101 0011

Four bits form a NYBBLE.
Eg... 0101

Basic Concepts Behind the Binary System
To understand binary numbers, begin by recalling elementary school math. When we first learned about numbers, we were taught that, in the decimal system, things are organized into columns :

H | T | O
1 | 9 | 3

such that "H" is the hundreds column, "T" is the tens column, and "O" is the ones column. So the number "193" is 1-hundreds plus 9-tens plus 3-ones. Years later, we learned that the ones column meant 10^0, the tens column meant 10^1, the hundreds column 10^2 and so on, such that the number 193 is really $\{(1x10^2)+(9x10^1)+(3x10^0)\}$.

$$10^2|10^1|10^0$$
$$1\ |\ 9\ \ |\ 3$$

As you know, the decimal system uses the digits 0-9 to represent numbers. If we wanted to put a larger number in column 10^n (e.g., 10), we would have to multiply 10 x 10^n, which would give $10^{(n+1)}$, and be carried a column to the left.

For example, putting ten in the 10^0 column is impossible, so we put a 1 in the 10^1 column, and a 0 in the 10^0 column, thus using two columns.

Twelve would be 12 x 10^0, or $10^0(10+2)$, or $10^1+2\text{x}10^0$, which also uses an additional column to the left (12).

The binary system works under the exact same principles as the decimal system, only it operates in base 2 rather than base 10. In other words, instead of columns being

$10^2|10^1|10^0$ they are $2^2|2^1|2^0$

Instead of using the digits 0-9, we only use 0-1 (again, if we used anything larger it would be like multiplying 2 x 2^n and getting 2^n+1, which would not fit in the 2^n column. Therefore, it would shift you one column to the left. For example, "3" in binary cannot be put into one column. The first column we fill is the right-most column, which is 2^0, or 1. Since 3>1, we need to use an extra column we fill is the right most column, which is 2 0, or 1. Since 3 1, we need to use an extra column to the left, and indicate it as "11" in binary
$(1 \text{ x } 2^1) + (1 \text{ x } 2^0)$.

Converting Binary To Decimal

Try converting these numbers from binary to decimal:

> *10*
> *111*
> *10101*
> *11110*

Remember:

2^4	2^3	2^2	2^1	2^0	
			1	0	$= (1x2^1)+(0x2^0) = 2+0=2$
		1	1	1	$= (1x2^2)+(1x2^1)+(1x2^0)=4+2+1=7$
1	0	1	0	1	$=(1x2^4)+(0x2^3)+(1x2^2)+(0x2^1)+(1x2^0)$ $= 16+0+4+0+1$ $= 21$
1	1	1	1	0	$= (1x2^4)+(1x2^3)+(1x2^2)+(1x2^1)+(0x2^0)$ $= 16+8+4+2+0$ $= 30$

Converting Decimal to Binary

Converting from decimal to binary notation is slightly more difficult conceptually, but can easily be done once you know how through the use of algorithms. Begin by thinking of a few examples.

We can easily see that the number 3= 2+1. and that this is equivalent to $(1x2^1)+(1x2^0)$. This translates into putting a "1" in the 2^1 column and a "1" in the 2^0 column, to get "11". Almost as intuitive is the number 5: it is obviously 4+1, which is the same as saying $[(2^2) +1]$, or 2^2+1. This can also be written as $[(1x2^2)+(1x2^0)]$.

Looking at this in columns,

2^2	2^1	2^0
1	0	1

or 101.

What we're doing here is finding the largest power of two within the number ($2^2=4$ is the largest power of 2 in 5), subtracting that from the number (5-4=1), and finding the largest power of 2 in the remainder ($2^0=1$ is the largest power of 2 in 1). Then we just put this into columns. This process continues until we have a remainder of 0. Let's take a look at how it works.

We know that:
$2^0=1$
$2^1=2$

$2^2=4$
$2^3=8$
$2^4=16$
$2^5=32$
$2^6=64$
$2^7=128$

and so on. To convert the decimal number 75 to binary, we would find the largest power of 2 less than 75, which is 64. Thus, we would put a 1 in the 2^6 column, and subtract 64 from 75, giving us 11. The largest power of 2 in 11 is 8, or 2^3. Put 1 in the 2^3 column, and 0 in 2^4 and 2^5.

Subtract 8 from 11 to get 3. Put 1 in the 2^1 column, 0 in 2^2, and subtract 2 from 3. We're left with 1, which goes in 2^0, and we subtract one to get zero.

Analogue sound into digital bits

In digital audio recording the instantaneous level of the analogue sound is repeatedly sampled.

The measured amplitude at that instant, is converted into the nearest digital value. Or in other words, the waveform is converted into a series of discrete numbers.

The string of numbers, representing the recorded waveform can be electronically stored by the digital circuitry and the playback will work in just the opposite way.

Each number in the sequence is converted back into a analogue voltage.
During this playback, a filter can be used to smooth off the sharp edges of the digital 'steps'.

Once the analogue signal has been converted into digital form, another problem arises.

Principles of True Modulation

Modulation and Digital Storage Techniques

If we are just using a computer memory, there is no problems in storing the data in digital form. But, musical recordings take up a lot of memory – thus making the system very expensive. A specialized media, such as the CD can sometimes be designed to store digital data directly. Eg... on a CD, a pit indicates a '0' and an island indicates a '1'.

But of course the data on a CD is modulated and encoded in a fairly complex manner.

Now, tape recording is very handy, efficient and relatively cheap. It can be easily recorded and played back.
But, unfortunately, a 1 or a 0 cannot be recorded onto a strip of magnetic tape, unless it is put through some sort of modulation.

One of the simplest approaches to recording digital data is the FSK.

This method is used by many inexpensive computers to store programs and data onto ordinary cassette tapes. But again it is an analogue recording method. The data is converted into analogue form before it is recorded (as tones).

FSK is not normally used for music recording. Because it is actually an analogue method for digital signals.

We are interested in a digital recording method for analogue signals.

In practical digital recording systems 'true modulation' is used.

MODULATION involves a constant waveform known as a carrier signal - usually at a frequency well above the audible range.

In the modulation process, some of the carrier signal is varied in step with the audible signal to be recorded (program signal).

To playback the recorded or modulated signal, we use de-modulation.

This process is the exact opposite of modulation and will involve the carrier signal to be deleted from the modulated signal, leaving only the program signal.

In digital recording systems, the carrier signal is always in the form of a pulse wave. A pulse wave switches between two discrete analogue voltages.

And because the transition time between the two stages is extremely short, the pulse wave switches instantly between states.

A pulse wave is used because it can easily be recognised and treated by other analogue or digital circuitry.

The modulation method used I most digital sound recorders is PCM.

Pulse Code Modulation
In PCM, the PCM circuitry looks at each new sample of the digitized program signal and converts the sample value into a pulse chain of a specific length and in de-modulation, these pulse chains are converted back into the appropriate sample values.

There are several factors to be considered in any recording system. These are as follows:

> *Frequency bandwidth*
> *Signal to noise ratio*
> *Linearity*

The frequency bandwidth is the range of program frequencies that can be recorded.

The signal to noise ratio is the difference between the nominal operating level and the noise floor.

Linearity is an expression for signal accuracy. Eg.. if the signal is graphed for various conditions, the straighter the lie, the better the systems linearity would be.

All three of these factors are important for both analogue and digital recordings.

Sampling frequency
The instantaneous amplitude or level of the analogue signal to be recorded is measured (sampled) many times per second. Each sample taken is then converted into an appropriate digital value. The rate at which this is done is known as the Sampling frequency or the Sampling Rate. Increasing the sampling frequency *will increase how much data* will be recorded for a recording of a given length. This will decrease the tape economy of the recording system.

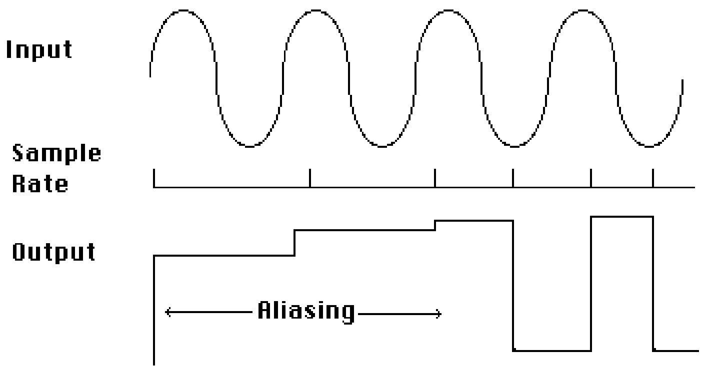
Increasing the sampling frequency of a digital recorder is the same as using a higher tape speed on an analogue tape recorder.

On the other hand, using too low sampling frequency will deteriorate the quality of the recording. This is like using slower tape speed on an analogue recorder.

However, the sampling frequency must be at least twice the highest signal frequency to be recorded. Which is known as twice the Nyquist frequency.

Also, if the sampling frequency is too low, a problem known as **aliasing** is likely to occur.

Aliassing occurs when the circuitry cannot properly decode the samples from the too high signal frequency – thus being translated a noise and distortion and even sometimes a low frequency which is not part of the original signal.



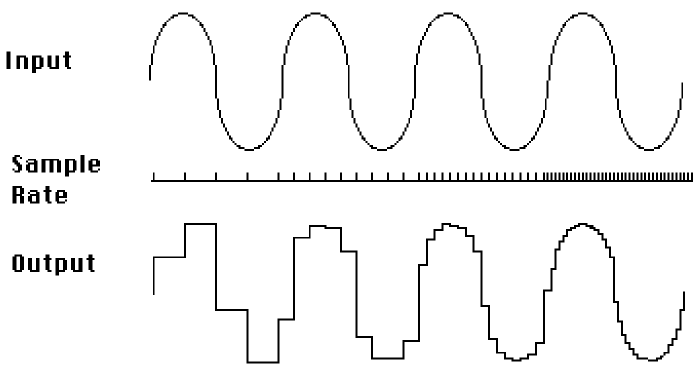
Effects of low sample rates

So, proper selection .of the sampling frequency is very important.

A CD has a sampling frequency of 44.1KHz; this will reproduce the whole frequency spectrum for high fidelity giving little extra headroom to the system, without increasing the overall cost of the system.
But many tape recorders are rated for a frequency response that only goes up to 15 kHz.

So if we set a 15 kHz limit on the recorded signal, the sampling frequency can be 30KHz, and to prevent the aliasing problems we need to block off the signal content above the maximum recordable limit.

A low pass filter is generally used to limit the original signal before it is converted into digital form for recording.



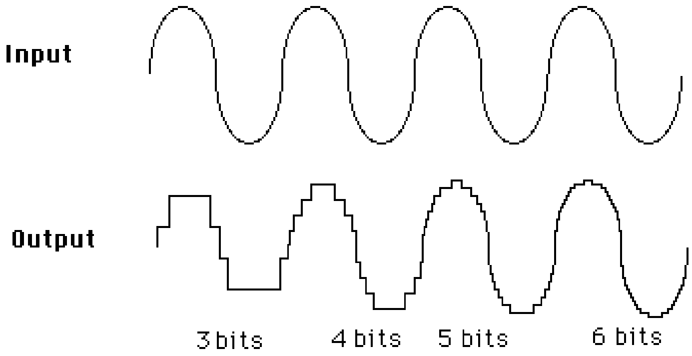
Effect of increasing sample rate

Resolution
Another very important specification for a digital audio recording system is the Resolution.

This is the number of bits used to represent the amplitude of each individual sample of the audio signal to be recorded.

Obviously the higher the number of bits used, the greater the amount of detail the system can record.

The minimal acceptable resolution is considered to be 14 bits and most modern digital recording equipment has a resolution of 14 bits, 16 bits and higher.



Word size

Quantisation
This is the process of breaking up the signal into discrete amplitude steps. The number of quantisation bits is a measurement of the resolution of the digital recorder. This is a important specification as the number of quantisation bits is directly related to the SNR of the recorder.

Eg.. a 16 bit system will have a SNR of 96dB

1 bit = 6 dB 16bit = 96dB

$DR = 20\log 2^n$

$(DR=20 \log (65536/1))$

$DR = 20 \log 2^{16}$

DR = 96dB

The process of quantisation introduces some noise of its own into the recorded signal, known as quantisation noise. This occurs because the digital recording system does not permit in between values. Each sample must have a discrete whole number value. So, if the amplitude of the audio signal being sampled happens to fall between two adjacent steps, the conversion circuitry must round the value either up or down. Either way some inaccuracy is introduced into the recorded signal, which will give a distortion like effect. This can be reduced by adding a small amount of analogue noise to the signal being digitized. White noise is used for this purpose and successfully masks all the effects of the quantisation noise. This added noise signal is called DITHER.

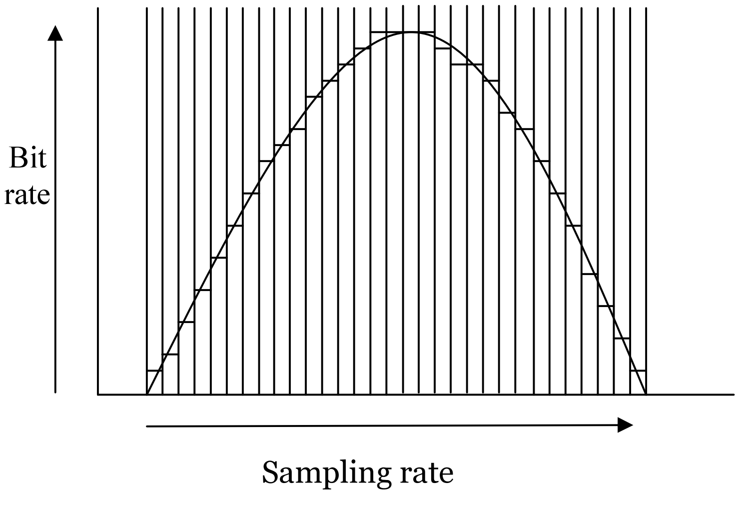
The sampling process in brief
The process of converting analog signal to digital signal is called *sampling*.

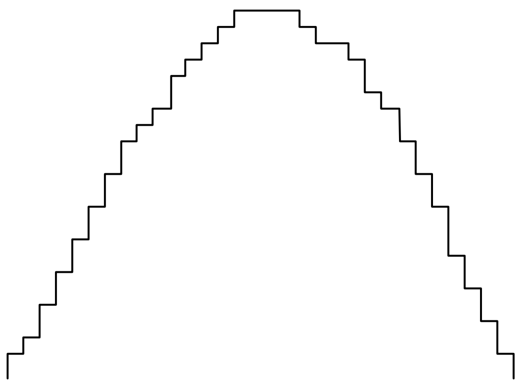
1. During sampling, "snap shots" of the analog signal is being taken at regular interval of 48,000 or 44,100 shots per second. The interval of the "snap shots" being taken is called "*sampling rate*".

2. In the case of compact disc, 44,100 "samples" are taken in each second. Each sample is then assigned a digital value depend on its average amplitude within the sample. This process is called "*quantization*".

   Increase in sampling rate and number of digits in describing the sample will increase the audio quality. However, that will also mean more storage space and faster processor is needed which results in expensive and less affordable hardware.

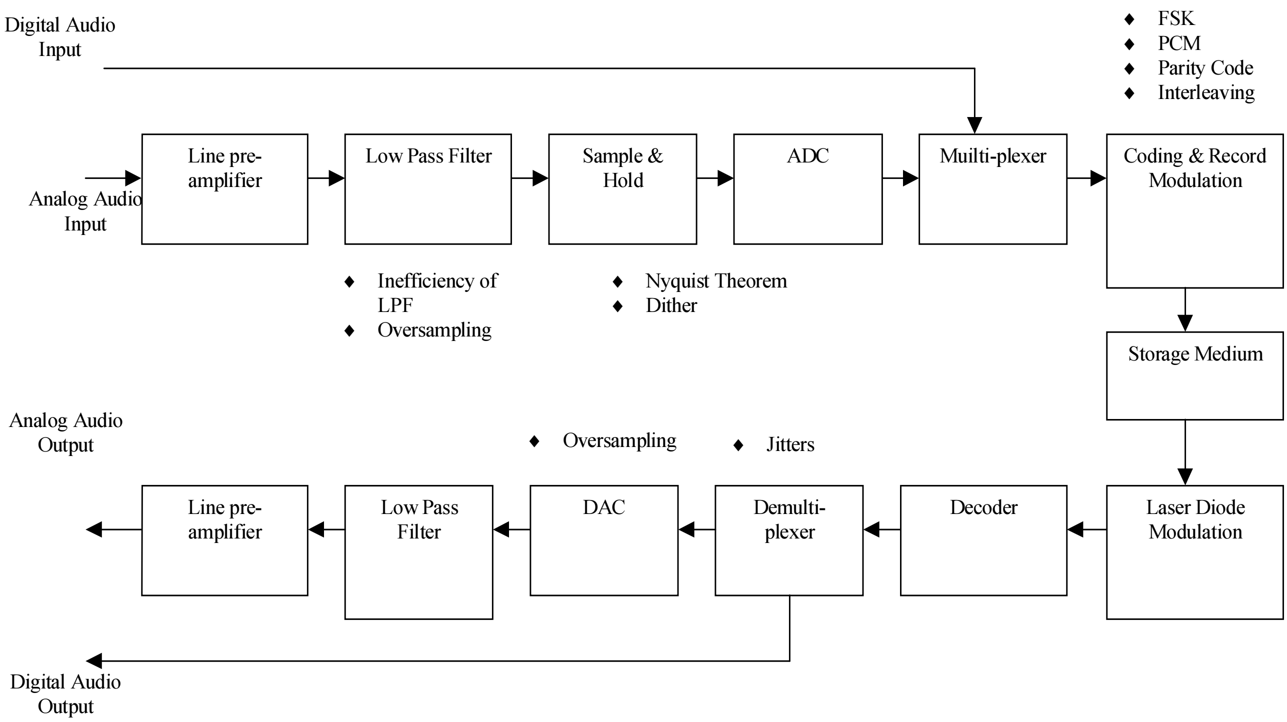
3. The resultant digital data is then encoded with error correction code to facilitate the storage and retrieval process.





Resultant digital audio with quantization noise/errors

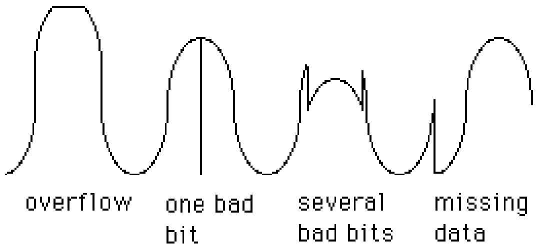
Digital and analog signal path



Error Correction Codes

Error Correction

Even with these techniques, the bits are going to be physically very small, and it must be assumed that some will be lost in the process. A single bit can be very important (suppose it represents the sign of a large number!), so there has to be a way of recovering lost data. Error correction is really two problems; how to detect an error, and what to do about it.



Effects of data errors

The most common error detection method is **parity** computation. An extra bit is added to each number which indicates whether the number is even or odd. When the data is read off the tape, if the parity bit is inappropriate, something has gone wrong. This works well enough for telephone conversations and the like, but does not detect serious errors very well.

In digital recording, large chunks of data are often wiped out by a tape dropout or a scratch on the disk. Catching these problems with parity would be a matter of luck. To help deal with large scale data loss, some mathematical computation is run on the numbers, and the result is merged with the data from time to time. This is known as a Cyclical Redundancy Check Code or **CRCC**. If a mistake turns up in this number, an error has occurred since the last correct CRCC was received.

Once an error is detected, the system must deal gracefully with the problem. To make this possible, the data is recorded in a complex order. Instead of word two following word one, as you might expect, the data is interleaved, following a pattern like:

words 1,5,9,13,17,21,25,29,2,6,10,14,18,22,26,30,3,7,15,19,27 etc.

With this scheme, you could lose eight words, but they would represent several isolated parts of the data stream, rather than a large continuous chunk of waveform. When a CRC indicates a problem, the signal can be fixed. For minor errors, the CRCC can be used to replace the missing numbers exactly. If the problem is more extensive, the system can use the previous and following words to reconstruct a passable imitation of the missing one. One of the factors that makes up the price difference in various digital systems is the sophistication available to reconstruct missing data.

The Benefits of Being Digital

You may be wondering about the point of all of this, if it turns out that a digital system is more complex than the equivalent analog circuit. Digital circuits are complex, but very few of the components must be precise; most of the circuitry merely responds to the presence or absence of current. Improving performance is usually only a matter of increasing the word size or the sample rate, which is achieved by duplicating elements of the circuit. It is possible to build analog circuits that match digital performance levels, but they are very expensive and require constant maintenance. The bottom line is that good digital systems are cheaper than good analog systems.

Digital devices usually require less maintenance than analog equipment. The electrical characteristics of most circuit elements change with time and temperature, and minor changes slowly degrade the performance of analog circuits. Digital components either work or don't, and it is much easier to find a chip that has failed entirely than one that is merely 10% off spec. Many analog systems are mechanical in nature, and simple wear can soon cause problems. Digital systems have few moving parts, and such parts are usually designed so that a little vibration or speed variation is not important.

In addition, digitally encoded information is more durable than analog information, again because circuits are responding only to the presence or absence of something rather than to the precise characteristics of anything. As you have seen, it is possible to design digital systems so that they can actually reconstruct missing or incorrect data. You can hear every little imperfection on an LP, but minor damage is not audible with a CD.

The aspect of digital sound that is most exciting to the electronic musician is that any numbers can be converted into sound, whether they originated at a microphone or not. This opens up the possibility of creating sounds that have never existed before, and of controlling those sounds with a precision that is simply not possible with any other technique.

Sampled sound processing

There are times when we want to store a sound briefly without resorting to tape or some other permanent medium. An example is the simulation of reverberation, where there are various delays of the sound as it reflects off distant walls. There is some propagation delay in electronic circuits, but that delay is of the order of five microseconds per mile, and is not very useful. The response delay of some circuits is longer, but is by nature frequency dependent.

S/PDIF

History

Since the early 80's, a step towards digital audio has been set by the introduction of the Compact Disc player. In the beginning, those signals stayed inside the set, and were converted to analog signals before leaving the cabinet. A new trend is to keep signals into the digital domain as long as possible, because this is the only way to keep the signal quality. To make this possible different devices must be able communicate with one another within the digital domain. Several interfaces exist to perform such tasks, from which one has grown to the audio standard worldwide: IEC958 1989-03 (consumer Part) from the EBU. In Japan an equivalent EIAJ CP-340 1987-9 is standard.

Characteristics

Standard IEC958 "Digital audio interface" from EBU (European Broadcasting Union) details:

- Audio format : linear 16 bit default, up to 24 bit expandable
- Allowed sampling frequencies (F*s*) of the audio:
  - 44.1kHz from CD
  - 48 kHz from DAT
  - 32 kHz from DSR
- One way communication: from a transmitter to a receiver.
- Control information:
  - V (validity) bit : indicates if audio sample is valid.
  - U (user) bit : user free coding i.e. running time song, track number.
  - C (channel status) bit : emphasis, sampling rate and copy permit.
  - P (parity) bit : error detection bit to check for good reception.
- Coding format: biphase mark except the headers (preambles), for sync purposes.
- Bandwidth occupation : 100kHz up to 6Mhz (no DC!)
- Signal bitrate is 2.8Mhz (Fs=44.1kHz), 2Mhz (Fs=32kHz) and 3.1Mhz (Fs=48kHz).

Physical connection:

- Cable: 75ohm +/-5% (l<10m) or 75ohm +/-35% (l>10m)
- Line driver:
  - Zout: 75Ω +/-20% (100kHz .. 6Mhz)
  - Vout: 0.4Vpp .. 0.6Vpp, <0.05Vdc (75ohm terminated)
- Line receiver:
  - Zin: 75Ω +/-5%
  - Vin: 0.2Vpp .. 0.6Vpp

The Interface

IEC958 is a newer standard which supersedes AES/EBU and also S-PDIF. The S/PDIF interface (IEC-958) is a 'consumer' version of the AES/EBU-interface. The two formats are quite compatible with each other, differing only in the **subcode information** and **connector**. The professional format subcode contains ASCII strings for source and destination identification, whereas the commercial format carries the SCMS.

Here is a short comparison table of AES/EBU and S/PDIF interfaces:

	AES/EBU	**S/PDIF (IEC-958)**
Cabling	110 Ω shielded twisted pair	75 Ω coaxial or fiber
Connector	3-pin XLR	RCA (or BNC)
Signal level	3..10V	0.5..1V
Modulation	biphase-mark-code	biphase-mark-code
Subcode information	ASCII ID text	SCMS copy protection info
Max. Resolution	24 bits	20 bits (24 bit optional)

Both S/PDIF and AES/EBU can, and do transfer 24 bit words. In AES/EBU, the last 4 bits have a defined usage, so if anyone puts audio in there, it has to go to something that doesn't expect the standard specifies. But in S/PDIF, there's nothing that says what you have to use the bits for, so filling them all up with audio is acceptable. Typical S/PDIF equipment's only use 16 or 20 bit resolutions. While many equipment's use more than 16 bits in internal processing, it's not unusual for the output to be limited to 16 bits.

Note on HDR-2 (2 pin header) interface used in some PC products:

Many modern PC CD-ROM drives and some soundcards (SB32, AWE32, etc.) have a two pin digital output connector in the back of the drive and they sometimes call that interface S/PDIF. Unfortunately the electrical signal which comes from it is not exactly what is described in S/PDIF specifications. The data format is exactly the same, but the signal is TTL level (5Vpp) signal instead of the normal 1Vpp signal. The output level might be selected to make the interfacing to other digital electronics easy when signal is travelling inside the computer (the normal output driver system and input amplifiers can be avoided). The downside of this is that you need to build some electronics to make the signal from the CD-ROM drive to match what normal S/PDIF equipment expects.

Multi channel audio and S/PDIF
IEC958 was named IEC60958 at 1998. IEC60958 (The S/PDIF) can carry normal audio and IEC61937 data streams. IEC61937 data streams can contain multichannel sound like MPEG2, AC3 or DTS. When IEC61937 datastreams are transferred, the bits which normally carry audio samples are replaced with the data bits from the data stream and the headers of the S/PDIF signal. Channel-status information contains one bit (but 1) which tells if the data in S/PDIF frame is digital audio or some other data (DTS, AC3, MPEG audio etc.). This bit will tell normal digital audio equipment that they don't try to play back this data as they were audio samples. (would sound really horrible if this happens for some reason). The equipment which can handle both normal audio and IEC61937 just look at those header bits to determine what to do with the received data.

Cabling details
S/PDIF (IEC-958) uses 75Ω coaxial cable and RCA connectors. 75Ω coaxial cable is inexpensive, because it is the same cable as used in video transmission (you can buy a video cable with RCA connectors to connect you S/PDIF equipment together).
AES/EBU-interface uses the well-known symmetrical connections with transformer isolation and an output impedance of 110Ω. The signal-level of this interface is reasonably higher than in the consumer version (3...10 volts).

Because AES/EBU digital audio signals are transmitted at high, video-like frequencies (at around 6MHz) and should be handled very differently than standard analog audio lines. Commonly used XLR-3 microphone cables have various impedance ratings (30Ωto 90Ω typical) and exhibit poor digital transmission performance. The result is signal drop out and reduced cable lengths due to severe impedance mis-matching (VSWR) between AES/EBU 110Ω equipment.

There also an optical version of S/PDIF interface which is usually called TOSLINK®, because it uses Toslink optical components. The transmission media is 1 mm plastic fiber and the signals are transmitted using visible light (red transmitting LED). The optical signals have exactly the same format as the electrical S/PDIF signals, they are just converted to light signals (light On/Off).

What can make difference in the sound of digital signal?
There are two things, which can cause differences between the sound of digital interfaces:

1. Jitter (clock phase noise)

This really only affects sound of the signal going directly to a DAC. If you're running into a computer, the computer is effectively going to be reclocking everything. It applies also to CD-recorders, DAT tape decks and similar devices. Even modern DACs have typically a small buffer and reclocking circuitry, so the jitter is not so big problem nowadays that it used to be.

2. Errors

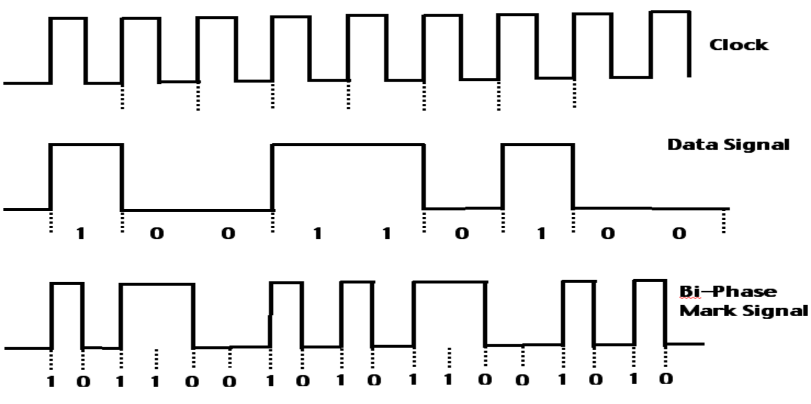
This usually causes very significant changes in the sound, often loud popping noises but occasionally less offensive effects. Any data loss or errors in either are a sign of a very broken link which is probably intermittently dropping out altogether.

S/PDIF signals
The signal on the digital output of a CD-player looks like almost perfect sine-wave, with an amplitude of 500 mVp-p and a frequency of almost 3 MHz.

For each sample, two 32-bit words are transmitted, which results in a bit-rate of:
2.8224 Mbit/s (44.1 kHz samplingrate, CD, DAT)
3.072 Mbit/s (48 kHz sampling rate, DAT)
2.48 Mbit/s (32 kHz sampling rate, for satellite purposes)

The output impedance is standard 75Ω, so ordinary coaxial cable designed for video applications can be used. The minimal input level of S/PDIF interface is 200 mVtt which allows some cable losses. There is no real need for special quality cable as long as the cable is made of 75Ω coaxial cable (a good video accessory cable works also as good S/PDIF cable).

The Coding Format
The digital signal is coded using the 'biphase-mark-code' (BMC), which is a kind of phase-modulation. In this system, two zero-crossings of the signal mean a logical 1 and one zero-crossing means a logical 0.



The frequency of the clock is twice the bit rate. Every bit of the original data is represented as two logical states, which, together, form a cell. The length of a cell ('time-slot') is equal to the length of a databit. The logical level at the start of a bit is always inverted to the level at the end of the previous bit. The level at the end of a bit is equal (a 0 transmitted) or inverted (a 1 transmitted) to the start of that bit.
The first 4 bits of a 32-bit word (bits 0 through 3) form a preamble which takes care of synchronization. This sync-pattern doesn't actually carry any data, but only equals four databits

in length. It also doesn't use the BMC, so bit patterns which include more than two 0's or 1's in a row can occur (in fact, they always do).

Digital recorders

There are many types of digital tape recorder using various (non compatible) formats. They all use the principles of digitizing sound covered in the digital audio easy, and they all face the same challenge: how to get a high enough frequency response to record the massive amounts of data audio needs. There are two fundamental approaches- many tracks, or a very high tape speed.

The machines that use the multi track systems are very expensive, (they need up to four tracks per audio channel) but are very reliable, and with some models the tape can be edited with a razor blade. High speed machines get the speed from a rotating head, just like video recorders. In fact the first of these were modified VTRs, and the new budget multitracks use standard consumer type video transports and tape.

Incidentally, it is this relationship of digital recorders to video that accounts for the funny sample rate of 44.1 khz. A single video frame has 490 lines, each recorded as a diagonal stripe across the tape. It turns out you can stuff 3 stereo samples in one of these lines. At 30 frames per second, you get 44,100 samples. Of course these numbers are based on black and white television. Color television runs a tad slower, so if you use a color VCR for recording, the sample rate winds up at 44,056.

The most important formats found today are Sony multitrack, Sony PCM, DAT, ADAT, and Tascam DA-88.

Studio Multitracks

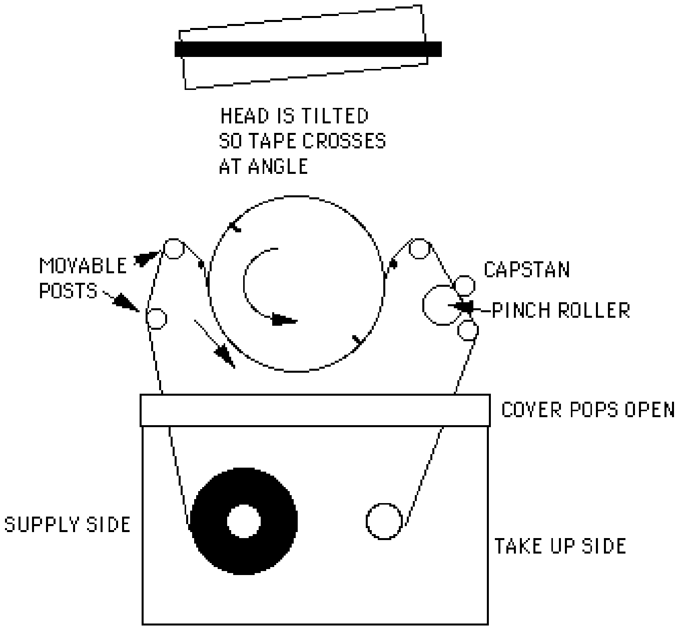
Sony DASH Multitracks are the workhorses of pro studios. Similar machines are made by Studer and a few others. They cost about as much as a house but are unsurpassed in sound and reliability. DASH stands for Digital Audio, Stationary Head. The quarter inch format uses 8 digital tracks to record a stereo signal. Some versions of these can be edited with a razor blade. The half inch machines record on up to 48 digital tracks, and depending on tape speed, can give up to 48 channels of audio.

PCM

Sony PCM systems are accessories to standard video decks. The 1630 and similar models use 3/4 in videocassettes, and are found in pro situations, especially CD mastering facilities. The F-1 was a consumer version, designed to be used with Betamax video decks (they work fine with VHS decks and many tapes were made this way). F-1 is no longer made, but systems survive, especially in electronic music studios. F-1 recordings have a 44,056 kHz sample rate, a fact that causes problems (like a pitch change) when the data is transferred to a newer medium.

DAT

DAT recorders record on a very narrow, slowly moving tape. They achieve the bandwidth necessary for this trick with a rotating head (Actually two heads on a rotating cylinder). Individual bits of data take up a microscopic area of tape; therefore the tape must be treated very gently, and never touched by human hands. The tape is normally hidden inside the plastic cassette out of reach: when it is inserted into the DAT recorder, the case is opened by the mechanism and threaded around the head spool. This is known as loading.



The head will spin whenever tape is loaded, and the tape is always contacting the head, even in fast forward and rewind. If you don't run the tape for a period of ten minutes or so, most decks unload to prevent head wear. This means an extra delay when you press play. The tape travels at 8.15 mm per second, but the head rotation of 2000 rpm gives an effective speed of 3.15 meters per second. (124 ips).

DATA ENCODING

Even at this speed the data has to be processed heavily to allow error free recovery at the other end. Some bit patterns, such as 00010000, would give a very narrow blip in the playback signal that is especially hard to detect. To avoid these, the usual 8 bit data words are recorded as selected 10 bit words, with the difficult ones left out. (This is called ETM for Eight to Ten Modulation. CDs use Eight to Fourteen Modulation.)

Each track (with a rotary head the tracks run across the tape at an angle) is just under an inch long. This is enough for 56448 bits after ETM demodulation. These are divided into 196 blocks of 288 bits. 256 bits of a block are used for data, the others for synchronization and error detection. Within a track, 128 blocks are for audio data and 16 for sub code data (IDs, Time and so forth); the rest are used to precisely control the way the tape moves across the heads.

A single 16 bit sample occupies 0.0003 inches. Naturally, any kind of a hole or dropout on the tape is going to have disastrous consequences. To prevent this, the data is scattered around the tape, a technique known as interleaving. Within each block, data bytes are mixed with parity bytes and error correcting bytes for other blocks. A total is kept of the data and parity, and if they don't match, that block is marked as containing an error. If the damaged region isn't too big (less than 22 blocks) the error correction bytes can be used to completely reconstruct the data. Within a somewhat larger region (74 blocks) interpolation can keep the music going. With more damage than this, the machine usually shuts off.

Error Indicators

Many decks have a light that flashes when errors are detected. Others have a hidden feature that gives some statistic, like errors per second. Errors are inevitable. Even a calibration tape will have two or three per second, and fresh recordings will read in the 20s or 30s. It is a good idea with a new machine to make a recording on a fresh tape and note the error rate. Then put the tape away. When doubt arises, play the reference tape and see what the error rate is. If it has gone up sharply, clean the heads. If you didn't happen to make a reference tape, you can tell the heads need cleaning when no tape will play properly. If a single tape stops playing, the tape is damaged or simply worn out. They can go surprisingly fast. In fact, enough tapes are unusable right out of the box that most engineers use two DATs when making critical recordings.

Head Cleaning on DATS

Proper cleaning of a DAT head requires disassembly of the machine and should be done by a qualified technician. Manufacturers used to include abrasive type head cleaners with the machines, but these should be used sparingly (if at all). Only clean when you know it's necessary.

About Time:

The time from the beginning of tape is recorded in every track of the tape as part of the subcode. If you start recording in the middle of a tape, the machine reads the time at that spot and keeps recording time subcode appropriately. If the tape is blank, this won't work. Therefore, every inch of a DAT tape should have something recorded on it. If you want silent spots on the tape, record zeros; either with the REC MUTE or by turning the input down. To find the end of recorded space, simply hit FAST FORWARD. You will be left cued up to the first blank. (Some decks have a special End Search function to do this.)

Start IDs:

Start IDs are also recorded in the subcode data of the tape. They run for about 9 seconds (so they are easy to find in fast motion) and contain a digital code indicating a start point.

On most decks, a start ID is recorded every time you hit the RECORD button and new one will be added if you Pause. Some decks feature an AUTO mode, where a Start is added if the signal stops for a second or so and restarts. Generally, you can record a start ID anywhere you want one, even while playing back. They are also easily removed.

Program Numbers

A program number is another chunk of data recorded in the subcode. These are the numbers searched for during Previous Play or AMS operations Most of the time the deck includes a program number with each start ID, but not always. (For instance, there will be no program numbers when you add IDs during play.) To sort out program numbers, most DATs have a RENUMBER function.

Skip IDs

Skip IDs instruct the deck to find the next start ID. You may write them in Record or Play mode.

End ID

An End ID marks that spot as the end of tape. The deck will not play or fast forward past the End ID. Adding them is just like adding START IDs. You should avoid having more than one of these on the same tape.

SCMS

The Serial Copy Management System is designed to enforce copyright protection of some kinds of material. Decks that follow the SCMS standard are locked out of digital recording under certain circumstances. This is accomplished by means of a two bit code known as ID6 and a byte called category code, which identifies the source of a digital input. The topic is confused by the fact that DAT machines made before SCMS was adopted may follow anther scheme called the "DAT Conference Specification" or ignore this issue entirely. The Topic is confused even further by the provisions that "professional" gear is exempt from SCMS.

The basic tenet of SCMS is that copyrighted material may be digitally copied only once. Copyright materials distributed on digital media (CDs, prerecorded DAT, Digital Broadcast) will have the ID6 set to 11 (binary). An SCMS deck is supposed to reset the ID6 to 10 if it was 11, and refuse to copy a digital signal coded 10. If ID6 is 00 unlimited copying is allowed.

The conference system was a little simpler. Such decks would not make digital copies of anything with a sample rate of 44.1 or anything with a an ID6 other than 00. If the source is analog, a conference type machine will mark the tape copyable (ID6=00) and an SCMS unit will mark it copy once (ID6=11). An SCMS machine recognizes an external A/D converter as analog and treats it the same way.

Loopholes

This state of affairs is not popular with the people who record for artistic reasons, as it is dreadfully inconvenient to be unable to copy your tapes. Various ways around SCMS have been found:

- You can buy quasi-legal boxes that will allow you to copy tapes regardless of the ID setting.
- Many decks can have the copy protection disabled by removing a jumper wire inside.
- You can always copy via analog connections, and the loss of quality is not noticeable.

SCMS and the Pros

Any DAT in professional use is exempt from SCMS, so machines designed for the pro market may dispense with it. In fact, it is hard to find a "consumer" DAT recorder right now, as the format was a flop in the audio market. However, all machines seem to follow SCMS sometimes.

The definition of a professional deck apparently lies in the type of digital input signal it will accept. Currently there are three standard types:

AES/EBU

This is a pro format, and copying is unlimited. It uses XLR type connectors. When this is used, Start IDs and the like get lost.