**Chapter 6**

**Basics of Digital Audio**

**Digitizing Sound**

* Microphone produces analog signal
* Computer deals with digital signal

**Sampling Audio**

**Analog Audio**

Most natural phenomena around us are continuous; they are continuous transitions between two different states. Sound is not exception to this rule i.e. sound also constantly varies. Continuously varying signals are represented by analog signal.

Signal is a continuous function f in the time domain. For value *y=f(t)*, the argument *t* of the function *f* represents time. If we graph *f,* it is called wave. (see the following diagram)

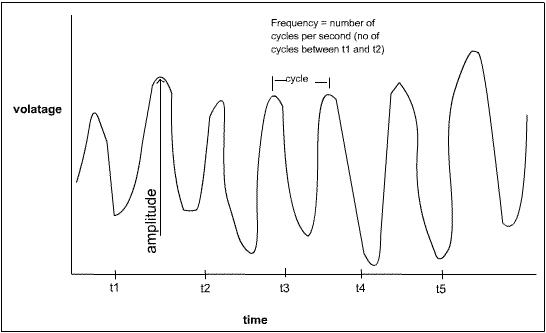


Fig 1 analog signal

A wave has three characteristics:

* Amplitude
* Frequency, and
* Phase

*Amplitude:* is the intensity of signal. This is can be determined by looking at the height of signal.If amplitude increases, the sound becomes louder. Amplitude measures the how high or low the voltage of the signal is at a given point of time.

*Frequency:* is the number of times the wave cycle is repeated. This can be determined bycounting the number of cycles in given time interval. Frequency is related with pitchness of the sound. Increased frequencyhigh pitch.

*Phase:* related to the wave’s appearance.

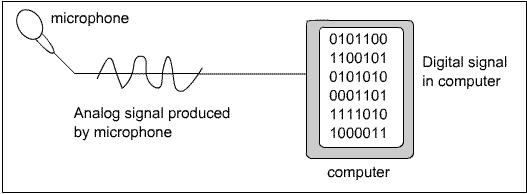


Fig 2 recording sound and the need for digitization

When sound is recorded using microphone, the microphone changes the sound into analog representation of the sound. In computer, we can’t deal with analog things. This makes it necessary to change analog audio into digital audio. How? Read the next topic.

**Analog to Digital Conversion**

Converting an analog audio to digital audio requires that the analog signal is *sampled*. Sampling is the process of taking periodic measurements of the continuous signal. Samples are taken at regular time interval, i.e. every T seconds. This is called *sampling frequency/sampling rate*. Digitized audio is sampled audio. Many times each second, the analog signal is sampled. How often these samples are taken is referred to as *sampling rate*. The amount of information stored about each sample is referred to as *sample size*.

Analog signal is represented by *amplitude* and *frequency*. Converting these waves to digital information is referred to as digitizing**.** The challenge is to convert the analog waves to numbers (digital information).

In digital form, the measure of amplitude (the 7 point scale - vertically) is represented with binary numbers (bottom of graph). The more numbers on the scale the better the quality of the sample, but more bits will be needed to represent that sample. The graph below only shows 3-bits being used for each sample, but in reality either 8 or 16-bits will be used to create all the levels of amplitude on a scale. (Music CDs use 16-bits for each sample).

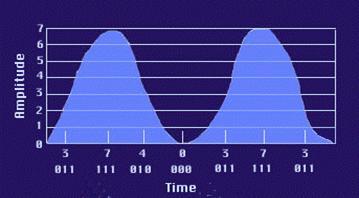


Fig 3 quantization of samples

In digital form, the measure of *frequency* is referred to as how often the sample is taken. In the graph below the sample has been taken 7 times (reading across). Frequency is talked about in terms of Kilohertz (KHz).

Hertz (Hz) = number of cycles per second KHz = 1000Hz

MHz = 1000 KHz

Music CDs use a frequency of 44.1 KHz. A frequency of 22 KHz for example, would mean that the sample was taken less often.

Sampling means measuring the value of the signal at a given time period. The samples are then quantized. **Quantization** is rounding the value of each sample to the nearest amplitude number in the graph. For example, if amplitude of a specific sample is 5.6, this should be rounded either up to 6 or down to 5. This is called quantization. Quantization is assigning a value (from a set) to a sample. The quantized values are changed to binary pattern. The binary patterns are stored in computer.

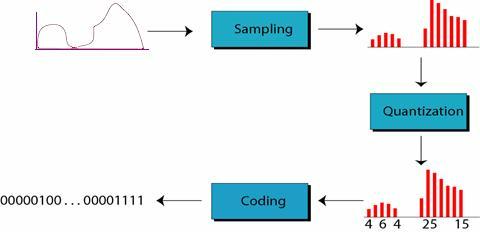


Fig 4 digitization process (sampling, quantization, and coding)

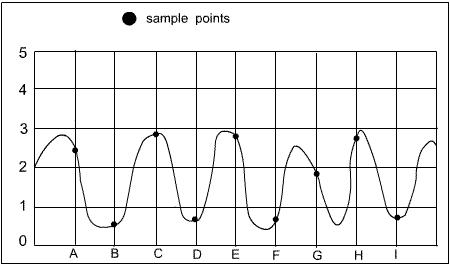


Fig 5 Sampling and quantization

Example:

The sampling points in the above diagram are A, B, C, D, E, F, H, and I.

The value of sample at point A falls between 2 and 3, may be 2.6. This value should be represented by the nearest number. We will round the sample value to 3. Then this three is converted into binary and stored inside computer.

Similarly, the values of other sampling points are: B=1

C=3

D=1

E=3

F=1

G=2

H=3

I=1

The values of most sample points are quantized. After quantization, we convert sample values into binary digits.

***Sample Rate***

A sample is a single measurement of amplitude. The sample rate is the number of these measurements taken every second. In order to accurately represent all of the frequencies in a recording that fall within the range of human perception, generally accepted as 20Hz–20KHz, we must choose a sample rate high enough to represent all of these frequencies. At first consideration, one might choose a sample rate of 20 KHz since this is identical to the highest frequency. This will not work, however, because every cycle of a waveform has both a positive and negative amplitude and it is the rate of alternation between positive and negative amplitudes that determines frequency. Therefore, we need at least two samples for every cycle resulting in a sample rate of at least 40 KHz.

**Sampling Theorem**

*Sampling frequency/rate* is very important in order to accurately reproduce a digital version of ananalog waveform.

*Nyquist’s Theorem:*

The Sampling frequency for a signal must *be* at least twice the highest frequency component in the signal.

Sample rate = 2 x highest frequency

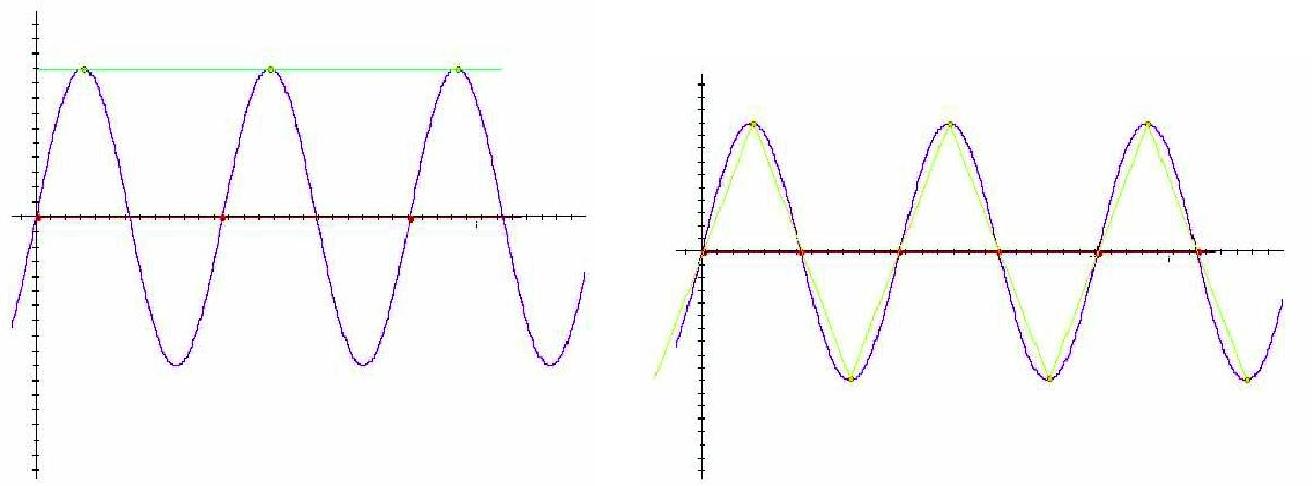
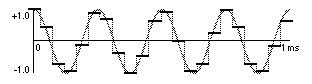


Fig 5 Sampling at signal frequency and at twice Nyquist frequency

When the sampling rate is lower than or equal to the Nyquist rate, the condition is defined as under sampling. It is impossible to rebuild the original signal according to the sampling theorem when such sampling rate is used.

**Aliasing**

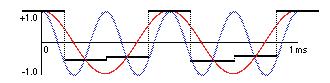
What exactly happens to frequencies that lie above the Nyquist frequency? First, we’ll look at a frequency that was sampled accurately:



In this case, there are more than two samples for every cycle, and the measurement is a good approximation of the original wave. we will get back the same signal we put in later on when converting it into analog.

**Remember:** speakers can play only analog sound. You have to convert back digital audio toanalog when you play it.

If we undersample the signal, though, we will get a very different result:



In this diagram, the blue wave (the one with short cycles) is the original frequency. The red wave (the one with lower frequency) is the aliased frequency produced from an insufficient number of samples. This frequency, which was in all likelihood a high partial in a complex timbre, has folded over and is now below the Nyquist frequency. For example, a 11KHz frequency sampled at 18KHz would produce an alias frequency of 7KHz. This will alter the timbre of the recording in an unacceptable way.

Under sampling causes frequency components that are higher than half of the sampling frequency to overlap with the lower frequency components. As a result, the higher frequency components roll into the reconstructed signal and cause distortion of the signal. This type of signal distortion is called aliasing.

**Common Sampling Rates**

* 8KHz: used for telephone
* 11.025 KHz: Speech audio
* 22.05 KHz: Low Grade Audio (WWW Audio, AM Radio)
* 44.1 KHz: CD Quality audio

**Sample Resolution/Sample Size**

Each sample can only be measured to a certain degree of accuracy. The accuracy is dependent on the number of bits used to represent the amplitude, which is also known as the sample resolution.

How do we store each sample value (*quantized value*)?

* ***8 Bit Value*** *(0-255)*
* ***16 Bit Value*** (Integer) (0-65535)

The amount of memory required to store t seconds long sample is as follows:

* If we use 8 bit resolution, mono recording memory = f\*t\*8\*1
* If we use 8 bit resolution, stereo recording memory = f\*t\*8\*2
* If we use 16 bit resolution, and mono recording memory = f\*t\*16\*1
* If we use 16 bit resolution, and stereo recording

memory =f\* t\*16\*2 where f is sampling frequency, and

t is time duration in seconds

Examples:

Abebe sampled audio for 10 seconds. How much storage space is required if

1. 22.05 KHz sampling rate is used, and 8 bit resolution with mono recording?
2. 44.1 KHz sampling rate is used, and 8 bit resolution with mono recording?
3. 44.1 KHz sampling rate is used, 16 bit resolution with stereo recording?
4. 11.025 KHz sampling rate, 16 bit resolution with stereo recording?

Solution:

1. m=22050\*8\*10\*1

m= 1764000bits=220500bytes=220.5KB

1. m=44100\*8\*10\*1

m= 3528000 bits=441000butes=441KB

1. m=44100\*16\*10\*2

m= 14112000 bits= 1764000 bytes= 1764KB

1. m=11025\*16\*10\*2

m= 3528000 bits= 441000 bytes= 441KB

**Implications of Sample Rate and Bit Size**

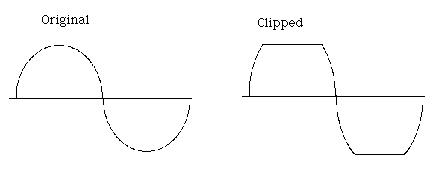
* Affects Quality of Audio
* Affects Size of Data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **File Type** | **44.1 KHz** | **22.05 KHz** |  | **11.025 KHz** |
|  |  |  |  |  |
| 16 Bit Stereo | 10.1 Mb | 5.05 Mb |  | 2.52 Mb |
| 16 Bit Mono | 5.05 Mb | 2.52 Mb |  | 1.26 Mb |
| 8 Bit Mono | 2.52 Mb | 1.26 Mb |  | 630 Kb |

Table Memory required for 1 minute of digital audio

**Clipping**

Both analog and digital media have an upper limit beyond which they can no longer accurately represent amplitude. Analog clipping varies in quality depending on the medium. The upper amplitudes are being altered, distorting the waveform and changing the timbre, but the alterations are slightly different. Digital clipping, in contrast, is always the same. Once an amplitude of 1111111111111111 (the maximum value in a 16 bit resolution) is reached, no higher amplitudes can be represented. The result is not the smooth, rounded flattening of analog clipping, but a harsh slicing of off the top of the waveform, and an unpleasant timbral result.



***An Ideal Recording***

We should all strive for an ideal recording. First, don’t ignore the analog stage of the process. Use a good microphone, careful microphone placement, high quality cables, and a reliable analog-to-digital converter. Strive for a hot (high levels), clean signal.

Second, when you sample, try to get the maximum signal level as close to zero as possible without clipping. That way you maximize the inherent signal-to-noise ratio of the medium. Third, avoid conversions to analog and back if possible. You may need to convert the signal to run it through an analog mixer or through the analog inputs of a digital effects processor. Each time you do this, though, you add the noise in the analog signal to the subsequent digital re-conversion.