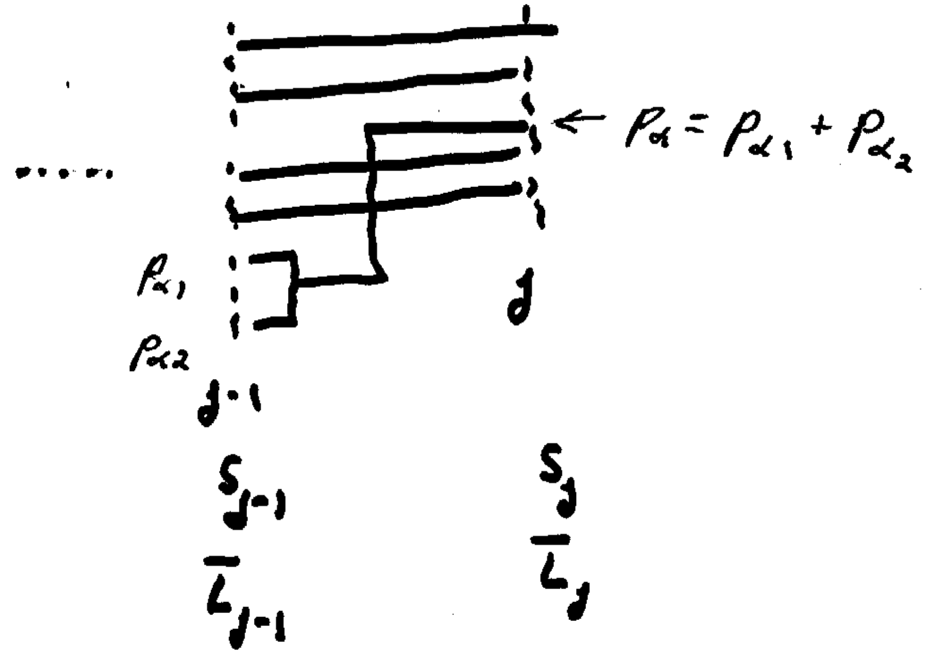
Proof of Optimality of Binary Huffman Codes.

(47)

with corrections



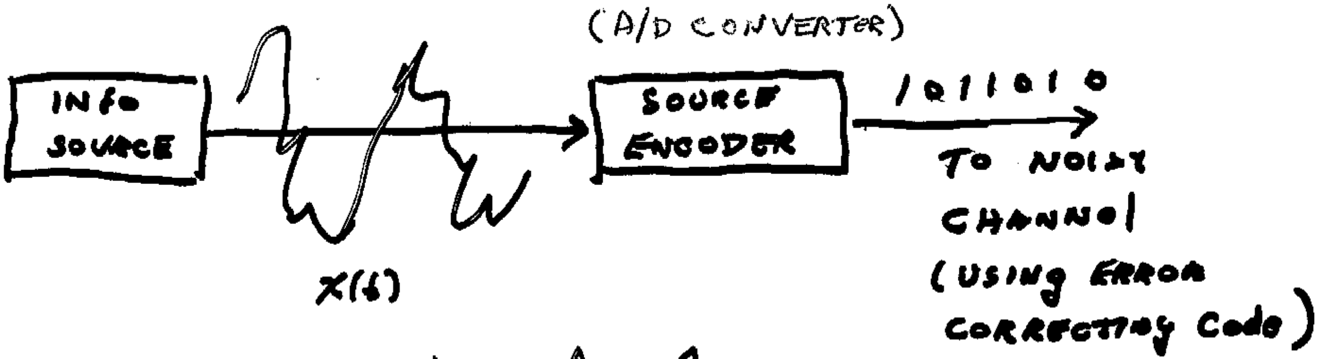
$$\bar{L}_{j-1} = \bar{L}_j + p_{\alpha_1} + p_{\alpha_2}$$ since code at $(j-1)$ is same as code at (j) except for two words that have length one more.

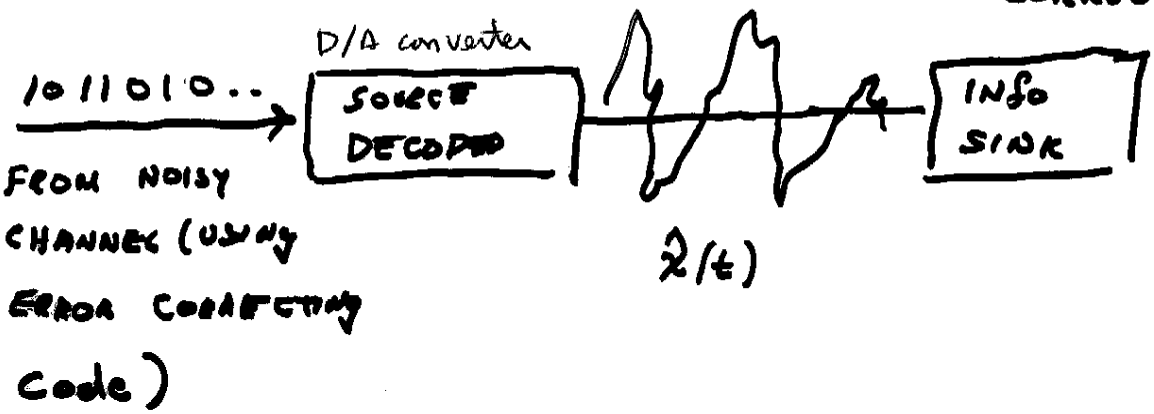
We now show that if code C_j is optimal then code C_{j-1} must also be optimal.

__Proof__ Suppose there were a better code at $(j-1)$. Call it's average length $\bar{L}'_{j-1} < \bar{L}_{j-1}$. But the two code words with probabilities p_{α_0} & p_{α_1} are identical in all but the last digit. Form a new code at j that has the identical prefix as the code word for p_α. This code will have average length $\bar{L}'_j = \bar{L}'_{j-1} - (p_{\alpha_1} + p_{\alpha_2})$ so that $\bar{L}'_{j} < \bar{L}_{j}$. But this can't be the case if C_j was optimal QED

CORRECTION

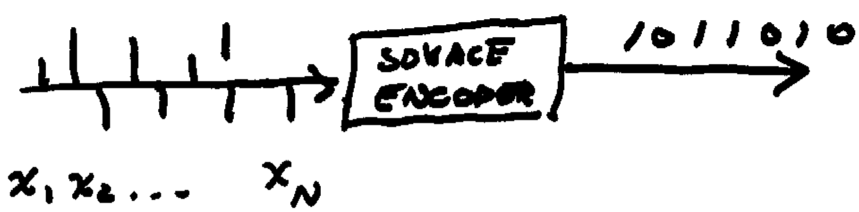
CODING WITH DISTORTION

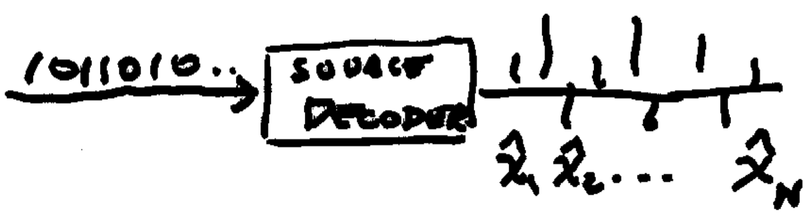




$$\varepsilon^2 = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} E\left(x(t) - \hat{x}(t)\right)^2 dt = \text{M.S.E.}$$

IF signals ARE bandlimited, ONE CAN SAMPLE AT NYQUIST RATE AND CONVERT CONTINUOUS-TIME PROBLEM TO DISCRETE-TIME PROBLEM. This SAMPLING IS PART of The A/D CONVERTER

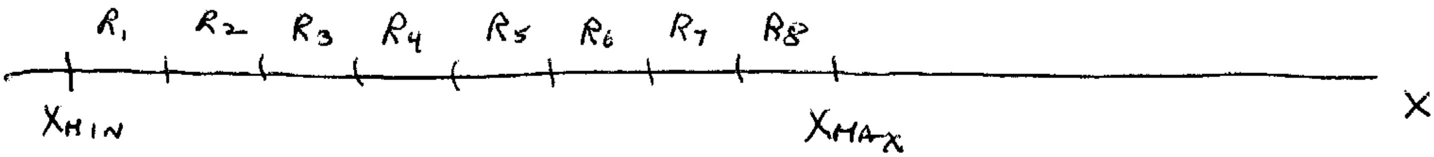




$$\varepsilon^2 = \lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} E\left(x_i - \hat{x}_i\right)^2$$

A/D Conversion and D/A Conversion

Assume a random variable X which falls into the range (X_{MIN}, X_{MAX}) is to be converted into k binary digits. Let $M = 2^k$. The usual A/D converter first subdivides the interval (X_{MIN}, X_{MAX}) into M equal sub-intervals of width $\Delta = (X_{MAX} - X_{MIN})/M$ as shown below for the case of $k = 3$ and $M = 8$. We call the i^{th} sub-interval, R_i.



Assume that if X falls in the region R_i ($x \in R_i$), then the D/A converter uses as an estimate of X, the value $\hat{X} = Y_i$ which is the center of the i^{th} region. Then the mean-squared error between X and $\hat{X}$ is

$$\varepsilon^2 \quad E[(X-\hat{X})^2] = \int_{X_{MIN}}^{X_{MAX}} (x-\hat{x})^2 f_X(x)\,dx$$

where $f_X(x)$ is the probability density function of the random variable X.

Let $f_{X|R_i}(x)$ be the conditional density function of X given that X falls in the region R_i. Then

$$\varepsilon^2 = \sum_{i=1}^{M} P[x \in R_i] \int_{x \in R_i} (x - Y_i)^2 f_{X|R_i}(x)\,dx$$

Note that

$$\sum_{i=1}^{M} P[x \in R_i] = 1$$

and

$$\int_{x \in R_i} f_{X|R_i}(x)\,dx = 1 \quad \text{for } i = 1, 2, \ldots, M$$

Now make the further assumption that k is large enough so that $f_{X|R_i}(x)$ is a constant over the region R_i. Then $f_{X|R_i}(x) = \frac{1}{\Delta}$ for all i,

and
$$\int_{x \in R_i} (x - Y_i) f_{X|R_i}(x)\,dx = \frac{1}{\Delta}\int_a^b \left(x - \left(\tfrac{b-a}{2}\right)\right)^2 dx = \frac{1}{\Delta}\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} (x-0)^2 dx$$

$$= \frac{1}{\Delta}\ \frac{2}{3}\cdot\left(\frac{\Delta}{2}\right)^3 = \frac{\Delta^2}{12}$$

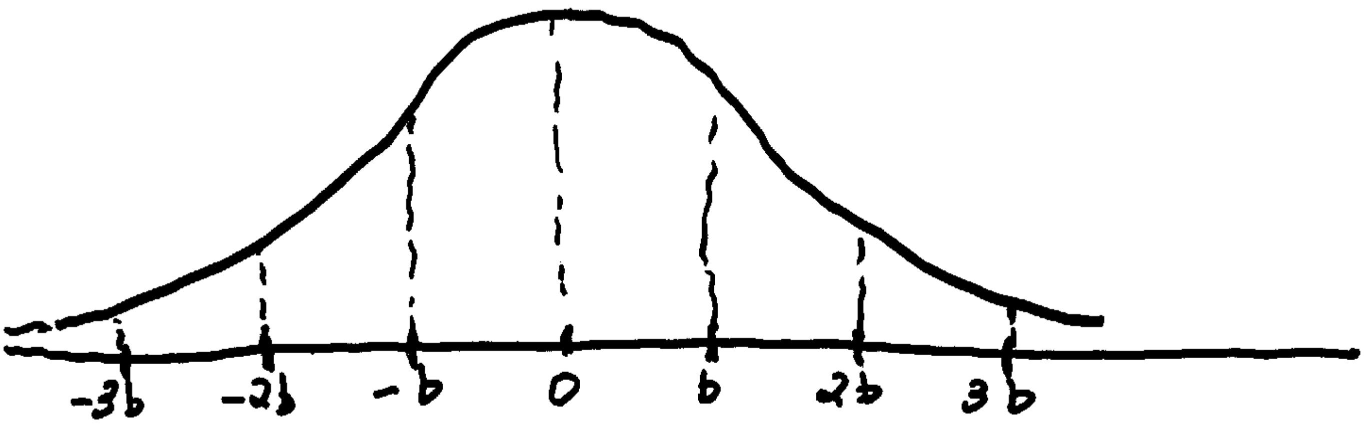
Then $$\varepsilon^2 = \sum_{i=1}^{M} P[x \in R_i]\cdot\frac{\Delta^2}{12} = \frac{\Delta^2}{12}$$

If X has variance σ_X^2, the signal-to-noise ratio of the A to D (& D to A) converter is often defined as $\left(\sigma_X^2 \Big/ \frac{\Delta^2}{12}\right)$

If X_{MIN} is equal to $-\infty$ and/or $X_{MAX} = +\infty$, then the last and first intervals can be infinite in extent. However $f_X(x)$ is usually small enough in those intervals so that the result is still approximately the same

SCALAR QUANTIZATION of (GAUSSIAN) samples.

USUAL SCALAR QUANTIZATION (3-BINARY DIGITS/SAMPLE)



ENCODER

$x \le -3b$	000	$0 < x < b$	100
$< x \le -2b$	001	$b < x \le 2b$	101
$-2b < x \le -b$	010	$2b < x \le 3b$	110
$-b < x \le 0$	011	$3b < x$	111

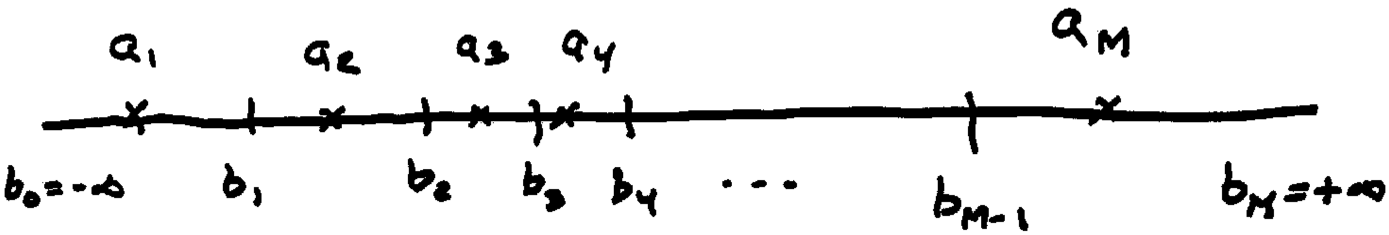
DECODER

000	$-3.5b$	100	$+.5b$
001	$-2.5b$	101	$+1.5b$
010	$-1.5b$	110	$+2.5b$
011	$-.5b$	111	$+3.5b$

OPTIMUM SCALAR QUANTIZER

$$b_{i-1} \le x < b_i \;\rightarrow\; \hat{x} = a_i \qquad i = 1, 2, \ldots, M$$

$$b_0 = -\infty, \quad b_M = +\infty$$



OPTIMIZE $\{b_i\}$ and $\{a_i\}$ to MINIMIZE ε^2

$$\varepsilon^2 = \sum_{i=1}^{M} \int_{b_{i-1}}^{b_i} (x - a_i)^2 f_X(x)\, dx$$

$$\frac{\partial \varepsilon^2}{\partial a_j} = 0 \qquad \frac{\partial \varepsilon^2}{\partial b_j} = 0 .$$

USE Leibnitz's Rule

$$\frac{\partial}{\partial t} \int_{a(t)}^{b(t)} f(x,t)\, dx = f(b(t),t)\, \frac{\partial b(t)}{\partial t} - f(a(t),t)\, \frac{\partial a(t)}{\partial t} + \int_{a(t)}^{b(t)} \frac{\partial}{\partial t} f(x,t)\, dt$$

$$\frac{\partial}{\partial b_j}\left(\sum_{i=1}^{M}\int_{b_{i-1}}^{b_i}(x-a_i)^2 f_X(x)\,dx\right)=$$

$$\frac{\partial}{\partial b_j}\int_{b_{j-1}}^{b_j}(x-a_j)^2 f_X(x)\,dx+\frac{\partial}{\partial b_j}\int_{b_j}^{b_{j+1}}(x-a_{j+1})^2 f_X(x)\,dx$$

$$=(b_j-a_j)^2 f_X(x)\Big|_{x=b_j}-(b_j-a_{j+1})^2 f_X(x)\Big|_{x=b_j}\overset{set}{=}0$$

$$\cancel{b_j^2}-2a_j b_j+a_j^2=\cancel{b_j^2}-2b_j a_{j+1}+a_{j+1}^2$$

$$2b_j(a_{j+1}-a_j)=a_{j+1}^2-a_j^2$$

$$\boxed{b_j=\frac{a_{j+1}+a_j}{2}}\quad (I)$$

$$\frac{\partial}{\partial a_j}\left(\sum_{i=1}^{M}\int_{b_{i-1}}^{b_i}(x-a_i)^2 f_X(x)\,dx\right)=-2\int_{b_{j-1}}^{b_j}(x-a_j)\,f_X(x)\,dx\overset{set}{=}0$$

$$a_j\int_{b_{j-1}}^{b_j}f_X(x)\,dx=\int_{b_{j-1}}^{b_j}x\,f_X(x)\,dx$$

$$\boxed{a_j=\frac{\int_{b_{j-1}}^{b_j}x f_X(x)\,dx}{\int_{b_{j-1}}^{b_j}f_X(x)\,dx}}\quad (II)$$

NOTE THAT THE $\{b_i\}$ CAN BE FOUND FROM (I) ONCE THE $\{a_i\}$ ARE KNOWN. (THE $\{b_i\}$ ARE THE MIDPOINTS OF THE $\{a_i\}$.)

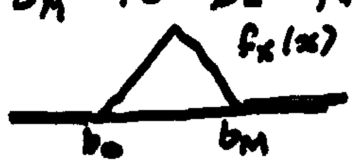
AND THE $\{a_i\}$ CAN BE SOLVED FROM (II) ONCE THE $\{b_i\}$ ARE KNOWN. (THE $\{a_i\}$ ARE THE CENTROIDS OF THE CORRESPONDING REGIONS.)

THUS ONE CAN USE A COMPUTER TO <u>ITERATIVELY</u> SOLVE FOR THE $\{a_i\}$ AND THE $\{b_i\}$.

1. ONE STARTS WITH AN INITIAL GUESS FOR THE $\{b_i\}$.
2. ONE USES (II) TO SOLVE FOR THE $\{a_i\}$.
3. ONE USE (I) TO SOLVE FOR THE $\{b_i\}$
4. ONE REPEATS STEPS 2 AND 3 UNTIL THE $\{a_i\}$ AND THE $\{b_i\}$ "STOP CHANGING."

COMMENTS

1. THIS WORKS FOR ANY $f_X(x)$
2. IF $f_X(x)$ ONLY HAS A <u>FINITE SUPPORT</u> ONE ADJUSTS b_0 & b_M TO BE THE LIMITS OF THE SUPPORT.



3. FOR A GAUSSIAN, ONE NEEDS TO KNOW

$$\int_\alpha^\beta f_X(x)\,dx \quad \text{and} \quad \int_\alpha^\beta x f_X(x)\,dx.$$ (TRUE FOR ANY $f_X(x)$)

$$\int_{\alpha}^{\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = Q(\beta) - Q(\alpha)$$

$$\int_{\alpha}^{\beta} x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \ldots$$ (integrate by parts) (or let $y = x^2$)

4. If $M = 2^a$ one could use "a" binary digits to represent the quantized value. However since the quantized values are not necessarily equally likely, one could use a Huffman code to use fewer binary digits (on the average)

5. After the $\{a_i\}$ and $\{b_i\}$ are known, one computes ε^2 from

$$\varepsilon^2 = \sum_{i=1}^{M} \int_{b_{i-1}}^{b_i} (x - a_i)^2 f_X(x)\,dx$$

6. For $M = 2$ and $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} x^2/\sigma^2}$ one can easily show that: $b_0 = -\infty$, $b_1 = 0$, $b_2 = +\infty$,

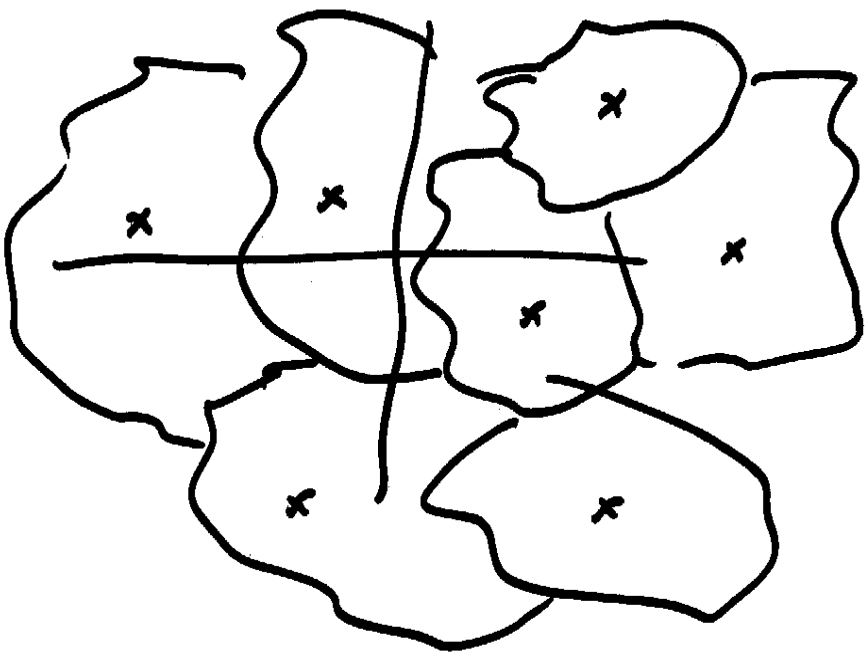
$$a_2 = -a_1 = \sqrt{\frac{2\sigma^2}{\pi}}$$

$$\varepsilon^2 = \left(1 - \frac{2}{\pi}\right)\sigma^2 = .3634\sigma^2$$

VECTOR QUANTIZATION

ONE CAN ACHIEVE A SMALLER ε^2 BY QUANTIZING SEVERAL SAMPLES AT A TIME.

WE WOULD THEN USE REGIONS IN A m-DIMENSIONAL SPACE



THE RATE-DISTORTION FORMULA TELLS US HOW SMALL ε^2 CAN BE AS $m \to \infty$.

FOR A GAUSSIAN WITH ONE BINARY DIGIT PER SAMPLE, $\varepsilon^2 \geq \frac{\sigma^2}{4} = (0.25)\sigma^2$

This follows from the RESULT ON The NEXT PAGE

DISCRETE-TIME GAUSSIAN SOURCE

Let source produce i.i.d Gaussian samples $x_1 x_2 \ldots$ where $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{x^2}{\sigma^2}}$

Let source encoder produce a sequence of binary digits at a rate of R binary digits/source symbol. In our previous terminology $R = k$

Let the source decoder produce the sequence $\hat{x}_1, \hat{x}_2, \ldots \hat{x}_i \ldots$ such that the mean-squared error between $\{x_i\}_i$ and $\{\hat{x}_i\}_i$ is ε^2.

Then one can prove that for any such system

$$R \geq \frac{1}{2}\log_2\left(\frac{\sigma^2}{\varepsilon^2}\right) \quad \text{for} \quad \varepsilon^2 \leq \sigma^2.$$

$(R = 0$ for $\varepsilon^2 \geq \sigma^2)$

This is an example of "RATE-DISTORTION THEORY"

Note that for $R = k = 1$,

$$1 \geq \frac{1}{2}\log_2\left(\frac{\sigma^2}{\varepsilon^2}\right)$$

$$2 \geq \log_2\left(\frac{\sigma^2}{\varepsilon^2}\right)$$

$$4 \geq \frac{\sigma^2}{\varepsilon^2}$$

OR

$$\varepsilon^2 \geq (1/4)\,\sigma^2$$

Reduced Fidelity Audio Compression

MP3 players use a form of audio compression called MPEG-1 Audio Layer 3. It takes advantage of a psycho-acoustic phenomena whereby a loud tone at one frequency "masks" the presence of softer tones at neighboring frequencies. Thus these softer neighboring tones need not be stored (or transmitted).

Compression efficiency of a audio compression scheme is normally described by the encoded bit rate (prior to the introduction of coding bits.) The CD has a bit rate of $(44.1 \times 10^3 \times 2 \times 16) = 1.41 \times 10^6$ bits/second. The term "44.1×10^3" is the sampling rate which is approximately the Nyquist frequency of the audio to be compressed. The term "2" comes from the fact that there are two channels in a stereo audio systems. The term "16" comes from the 16-bit (or $2^{16} = 65,536$ levels) A to D converter. (A slightly higher sampling rate 48×10^3 samples/second is used for a DAT recorder.)

Different standards are used in MP3 players. Several bit rates are specified in the MPEG-1, Layer 3 standard. These are 32, 40, 48, 56, 64, 80, 96, 112, 128, 144, 160, 192, 224, 256 and 320 kilobits/sec. The sampling rates allowed are 32, 44.1 and 48 kiloHz but the sampling rate of 44.1×10^3 Hz is almost always used.

The basic idea behind the scheme is as follows. A block of 576 time domain samples are converted into 576 frequency-domain samples using a DFT. The coefficents are then modified using psycho-acoustic principles. The processed coefficient are then converted into a bit stream using various schemes including Huffman encoding. The process is reversed at the receiver: bits → frequency-domain coefficents → time-domain samples.