Waveform Coding Modulation

Pulse Code Modulation

Pulse Code Modulation, Delta Modulation and their variants can all be described using the block diagram of Figure 1.

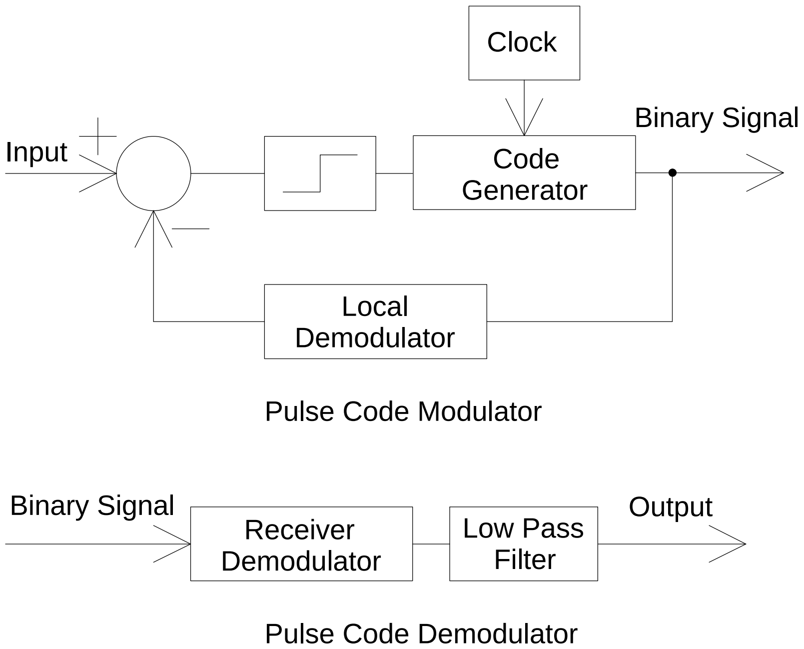

Figure 1. PCM Block Diagram.

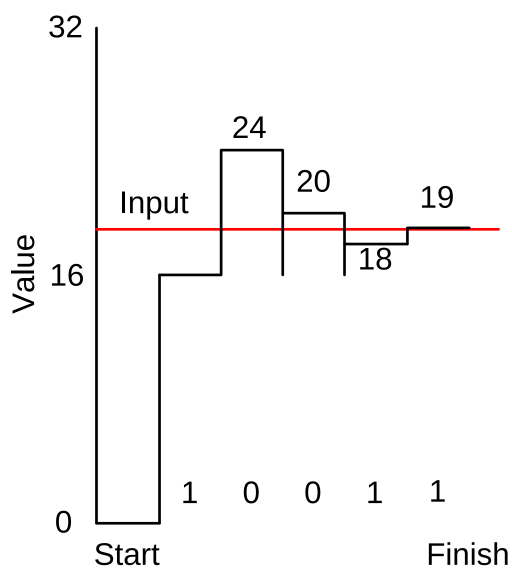

Figure 2. Timing waveforms for PCM.

The operation of a typical Pulse Code Modulator is illustrated in Figure 2, where an input value of 19 is evaluated with a 5-bit (ie 32 level) PCM coder. At the start of the conversion interval, the code generator makes a guess of the value of the input, by checking if the input is greater or smaller than the most significant bit.

According to information theory, the most information about an event is obtained if the outcome from a test or question is equally likely to be one or the other. The testing for the most significant bit will thus result in the quickest way to determine the input voltage. If the input voltage is more than the most significant bit, that bit is kept at one, otherwise it is set to zero and the voltage is removed. The second most significant bit is then tested and kept if needed and removed if not needed. After 5 tests we do this know that the input is closest to 19 units. For an 8-bit ADC or Pulse Code Modulator, 8 tests are required to evaluate the input voltage to a precision of 1 in 256, for a 10 bit PCM coder, 10 tests are required to evaluate the input voltage to a precision of 1 in 1024.

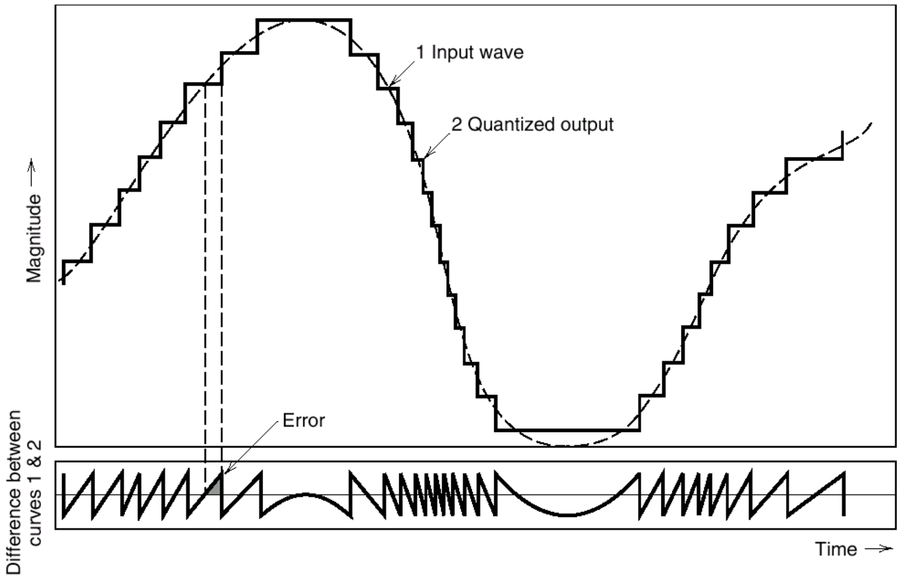

Figure 3. PCM Waveforms (Bennett 1948, Haykin Fig 3.11)

Figure 3 shows how an input waveform is approximated to during the digitising process and how the quantisation error is the difference between the analogue input and quantised output waveforms.

If a waveform like a sinewave or a speech signal is applied to the input of the Pulse Code Modulator, then if the peak amplitude of the input signal is less than the maximum possible value, then a constant quantisation noise is obtained and the signal to noise ratio is proportional to the input signal power. When the peak input signal amplitude is larger than the maximum permissible amplitude, clipping will occur resulting in a large distortion, which increases rapidly with increasing input signal power. Figure 4 is a plot of the calculated SNR for simulated speech versus the number of bits in a PCM system. The quantum levels are assumed constant, so that as the number of bits increases, the maximum permissible amplitude increases as the number of bits raised to the power of two. It can be seen that increasing the number of bits by one increases the peak SNR by 3 dB.

SNR of PCM

The SNR of a PCM system can be calculated under ideal conditions. In practice the SNR will be worse than these ideal conditions. However this formula gives a good indication of the trends. The number of quantum levels can be obtained as:

$$q = \mu^{\nu} \qquad \text{Eqn. 1}$$

Where

μ - Number of discrete levels in transmission (binary μ =2, for 64 QAM, μ= 64)
ν - Number of Digits per quantum level ie 8 bit for telephony , 16 bit for soundcard
q - Number of quantum levels

For example 7 bit PCM has $q = 2^7 = 128$ levels.

The bandwidth required to transmit a pulse of width τ is:

$$BW_T = \frac{1}{2\tau} \qquad \text{so that :} \qquad \tau = \frac{1}{2BW_T} \qquad \text{Eqn. 2}$$

For every sample there are ν digits to be transmitted, so that:

$$T_s = \frac{1}{F_s} = \nu\tau \qquad \text{Eqn. 3}$$

From Nyquist, we have the minimum sampling rate as:

$$F_s = 2W \qquad \text{Eqn. 4}$$

Where W is the audio bandwidth. Substituting Equation 4 into 2 gives:

$$\tau = \frac{1}{2\nu W} \qquad \text{and} \qquad BW_T = \nu W \qquad \text{Eqn. 5}$$

Using this together with equation 2 into equation 1 gives:

$$q = \mu^{\nu} = \mu^{BW_T/W} \qquad \text{Eqn. 6}$$

Giving an exponential relationship between the number of quantum levels and the bandwidth required. This expression can also be rewritten as:

$$BW_T = W \log_{\mu}(q) \qquad \text{Eqn. 7}$$

Consider the quantisation noise, shown in figure 3. If the spacing between the quantum levels is α and the amplitude is normalised such that the maximum amplitude is ±1 then:

$$q\alpha = 2$$

If the system is sampled fast enough, the noise does not exceed the limits, so that the probability of a particular noise amplitude ε occurring has a uniform probability in the region ±½ α. Since the area under the probability curve is 1:

$$p(\varepsilon) = \frac{1}{\alpha} \qquad \text{Eqn. 8}$$

$$\overline{\varepsilon}^2 = \int_{-\infty}^{\infty} \varepsilon^2 p(\varepsilon) d\varepsilon = \int_{-\frac{\alpha}{2}}^{\frac{\alpha}{2}} \varepsilon^2 \frac{1}{\alpha} d\varepsilon = \frac{\alpha^2}{12} = \frac{1}{3q^2} \qquad \text{Eqn. 9}$$

Combining this with equation 6 gives:

$$SNR = 3q^2 \overline{X}^2 = 3\mu^{2BW_t/W} \overline{X}^2 \qquad \text{Eqn. 10}$$

This gives an exponential exchange of Bandwidth and SNR, the best of any system, so that going from a 10 bit PCM system to an 11 bit system increases the bandwidth required by 10% but increases the SNR by 6 dB. For instance FM can be shown to have a square law relationship, doubling the bandwidth gives a 6 dB increase in SNR. This is the prime reason why the communication systems are becoming digital systems.

For a sinewave modulation system, $\overline{X}^2 = \frac{1}{2}$. And since $\frac{BW_T}{W} = \nu = N$, where N is the number of bits for an Analogue to Digital Converter, the same as ν in the above derivation, the SNR of a ADC is given as:

SNR= 1.76 + 6.02 N dB Where $1.76 = 10\log_{10}(\frac{3}{2})$.

Actual communication systems will have a poorer performance, since the sampling frequency is > 2W, the error is sometimes greater than half a quantum level and $\overline{X}^2$ is much less than 0.5. However the SNR versus bandwidth will still show a similar trend to equation 10. For example simulated speech with a 3.5 kHz bandwidth and an 8 kHz sampling frequency has:

$SNR \cong 0.6x2^{1.6BW_T/W}$ for a binary transmission system. Eqn. 11

The performance of PCM for simulated speech is shown in figure 4. It consists of two regions: At low signal level, a region of constant quantisation noise where the SNR increases linearly with input power and a region of overload, where the SNR drops rapidly with increasing power. The value from equation 11 corresponds to the maximum value of each of these curves.

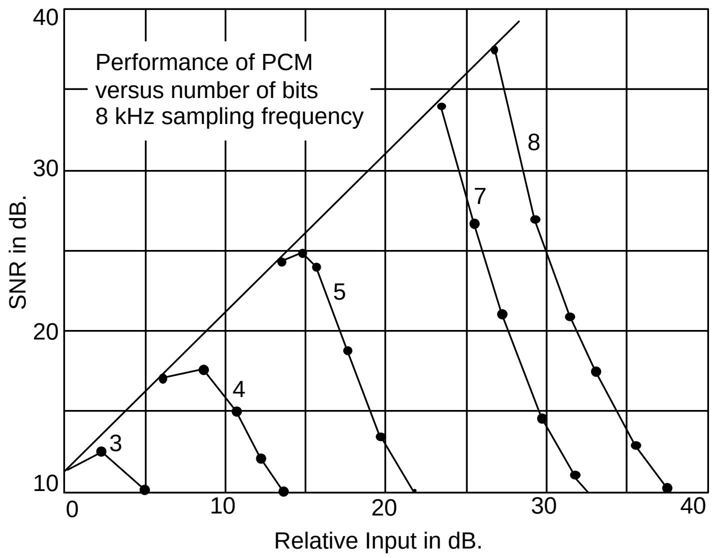

Figure 4 Performance of PCM

For Telephony systems, a 3.5 KHz audio bandwidth is required and an 8 kHz sampling frequency is used. For an 8 bit system, a 64 kBPS data stream is required and this results in a 37 dB SNR, which is good enough for the average listener not to be able to pick the difference between the decoded PCM waveform and the uncoded waveform. A 5 bit PCM waveform is still acceptable and intelligible, but the coding is obvious.

Delta Pulse Code Modulation

In Delta Pulse Code Modulation (ΔPCM), the difference between the present sample and the previous sample is sent. This has the advantage that even with a small number of bits, a large DC number can be sent, as this is an accumulation of many samples. ΔPCM is no longer amplitude limited but is slope limited, since the change between samples is limited. Since speech has a spectrum that decreases with frequency, it is well matched to a slope limited system like ΔPCM. A 7 bit ΔPCM system performs as well on speech as an 8 bit PCM system. The main problem with ΔPCM is that a transmission error will cause a permanent DC error in the demodulated output. Since the ear is not sensitive to DC, this is not a great problem.

One of the problems in transmitting PCM is that it is necessary to know the location of the start of the word, ie which bit is the most significant bit. A frame synchronisation scheme is required. This uses additional data bits, with in the simplest systems, one bit being used for synchronisation. In addition the precision of the most significant bit must be greater than half the size of the least significant bit. For an 18 bit PCM system, the accuracy of the most significant bit must thus be better than 1:524 288, a difficult task to achieve.

Delta Modulation

Delta Modulation is a system single bit ΔPCM system operating at a much higher sampling frequency than required by the Nyquist Sampling Theorem. The operation of the system can be illustrated using the waveforms of Figure 4.

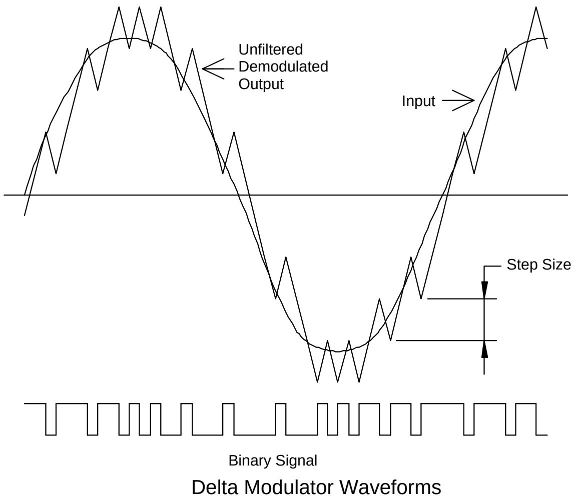

Figure 4. Delta Modulation Waveforms.

At each sampling instant, the input and the output are compared. If the output is smaller than the input, a binary output is generated, which increases the output. At each sampling instant, negative feedback is thus applied to reduce the difference between the output and the input.

A higher sampling rate for the same step size of the output, results a larger peak signal being able to be tracked. Since the quantisation noise is determined by the step size and is thus the same, a higher sampling frequency will thus result in a higher SNR.

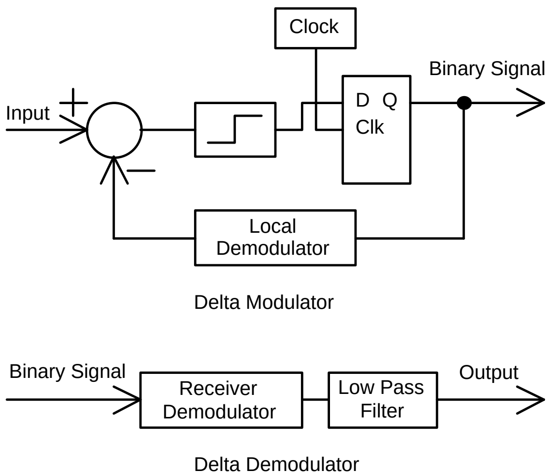

Figure 5. Delta Modulator Block diagram

Figure 5 shows the block diagram of a Delta Modulator. The Code generator fro Figure 1 is now replaced by a simple Flip-flop. The local demodulator is often a simple first order RC low pass filter network. The hardware requirements for a delta modulator is thus very simple. Since the output of a Delta Modulator is a simple continuous data stream, no word synchronisation is required. Since the output has only one step size for incrementing or decrementing the output, the accuracy requirements like those on the most significant bit do not exist. Delta Modulation has thus several advantages.

The SNR of Delta Modulation for simulated speech is plotted in Figure 6. By comparing this figure with the corresponding Figure 3 for PCM, it can easily be seen that the performance of Delta Modulation is worse than that of PCM.

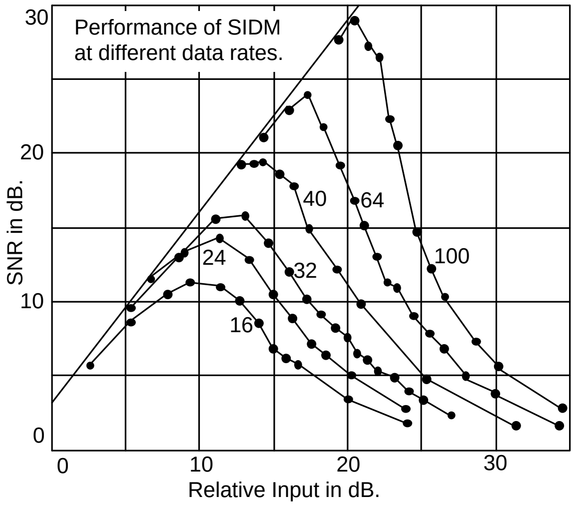

Figure 6. Performance of Delta Modulation.

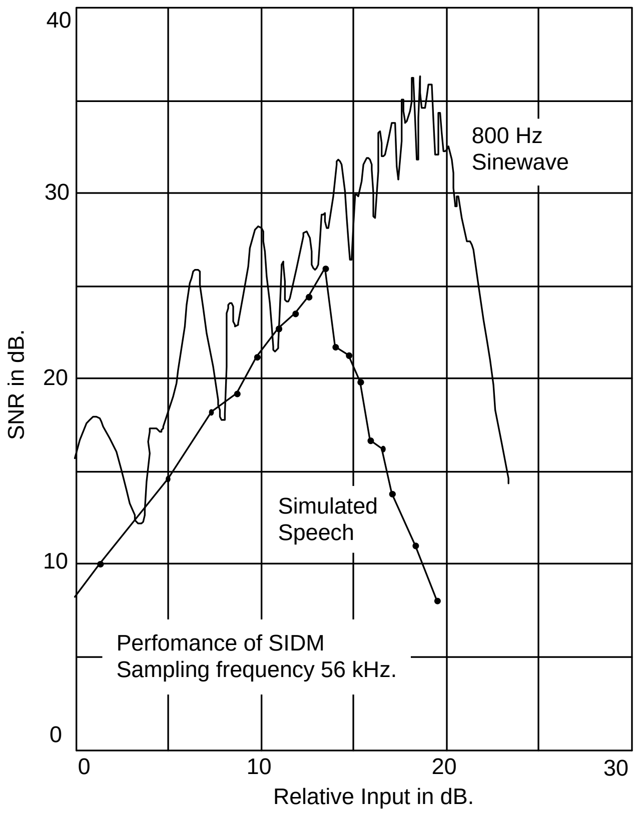

Figure 7. Performance of Delta Modulation for Speech and Sinewave Inputs.

It should be noted that for a data rate of 40 kSPS or less, corresponding to a 5 bit or less PCM system with an 8 kSPS sampling rate, the difference in performance between PCM and Delta Modulation is minimal and not worth the additional hardware complexity and word synchronisation requirements. For a 64 kBPS data transmission rate PCM and ΔPCM performs far better.

It is very important that any computer simulation of a system and testing of an actual system is done using an input for which the system is to be designed. This is clearly illustrated using Figure 7, which shows the performance of a Delta Modulator when subject to a simulated speech input and a sinewave input. The performance of the Delta Modulator for sinewave inputs varies dramatically over a small change of input level as the generated patterns give a good fit to a sinewave, with very few harmonics in the audio bandwidth or a poor fit with a high level of harmonic distortion. It is interesting to see that the measured performance published by other workers on Delta Modulation, who use sinewaves for testing follows the peaks of the plot in Figure 7. The performance using speech is very different in two ways. Firstly because the signal is more random and its power level varies with each syllable, the SNR for speech is less than that for sinewaves. Secondly the input level at which the peak SNR occurs about 5 dB less than that for sinewaves. This is important for the design of companding, which is discussed later.

Delta Sigma Modulation

Delta Sigma Modulation (ΔΣM or DSDM), also known as Sigma Delta Modulation, is a system where the local demodulator of Figure 5 is placed after the summer, rather than in the feedback loop, as shown in Figure 8.

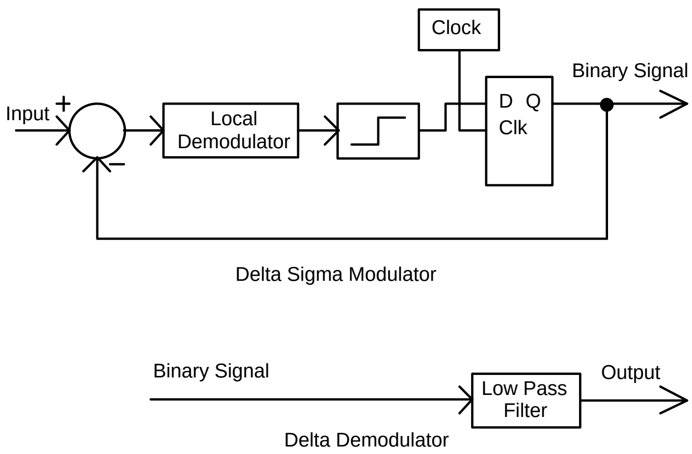

Figure 8. Block diagram of Delta Sigma Modulation.

This has the advantage that the demodulation process is simply a low pass filter. Note that even though the demodulation process for ΔΣM and PWM, PPM and PAM is the same, ΔΣM is a true digital system, where the signal is quantised and the others are analogue systems, where the information is carried in the transitions between the levels. ΔΣM must have a differencing circuit that is capable of handling the large amplitude and high bandwidth digital signal. Because of its simplicity in modulation and demodulation, ΔΣM is often used in precision audio analogue to digital and digital to analogue converters, like those used in computer soundcards.

Double Integration Delta Modulation

For the basic Delta Modulator, the local demodulator is a simple RC low pass filter that can be considered as a simple and slightly leaky integrator. For double integration delta modulation, the local demodulator consists of two single RC low pass filters in cascade. This results in more filtering of the out of band quantisation noise in the feedback loop and gives more “noise shaping”. This term “noise shaping” is often used for ADC and DAC applications and shows how by filtering the feedback signal, the noise power is reduced in the audio frequency band. The extra phase shift from the second “integrator” can cause instability. The design of a Double Integration feedback network is thus a compromise between filtering and stability. Third order feedback networks are possible, but higher order networks are increasingly difficult to design.

It is possible to apply noise shaping to Delta Sigma Modulation as well. Some commercial ADC's using ΔΣM use a fifth order local demodulator. This requires a highly specialised feedback network to still result in a stable modulator.

System Comparison

Figure 9 shows a comparison of PCM, ΔPCM and Delta Modulation (ΔM). From this graph it can be seen that for data rates at 40 kSPS or below, there is little difference in performance between PCM and ΔM.

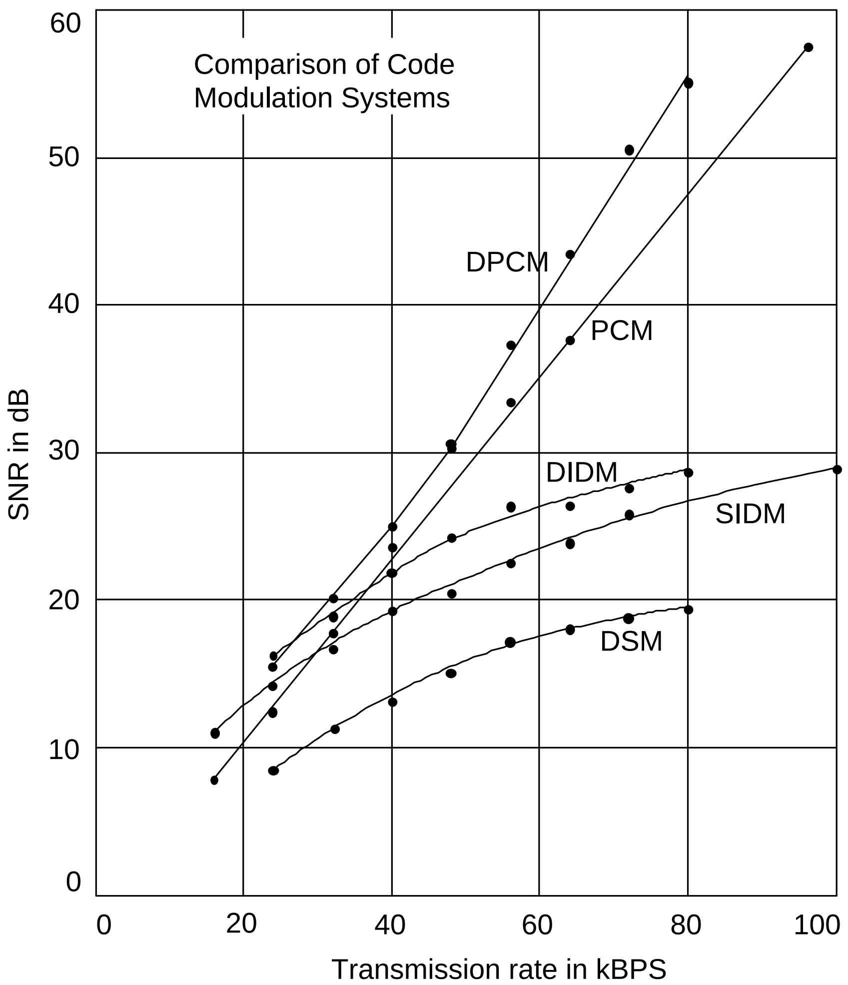

Figure 9. Comparison of PCM, ΔPCM and ΔM performance versus data rate

It can also be seen that ΔΣM has the worst performance of all the systems, and yet is is being used for the highest performances Analogue to Digital and Digital to Analogue Conversion systems.

Instantaneous Companding

In Instantaneous companding the spacing between the quantum levels in PCM is changed, such that small increments are used at low signal levels and larger increments are used at higher levels. The quantisation step size is thus matched to the signal amplitude, giving a more constant SNR with changing input level.

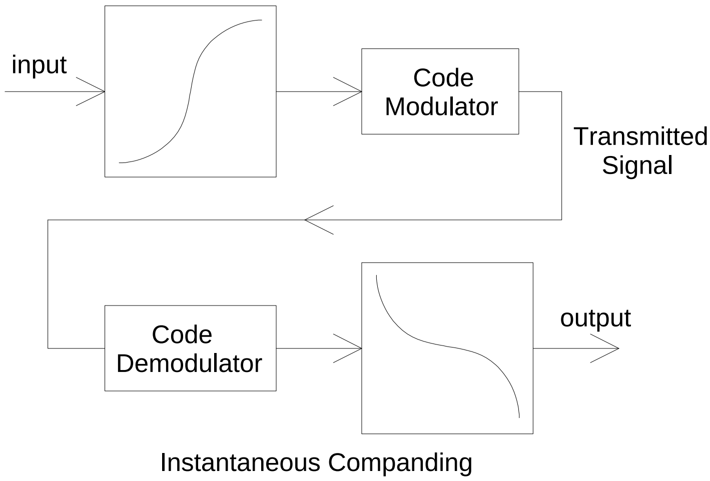

Figure 10. Non-Uniform step sizes used in instantaneous companding.

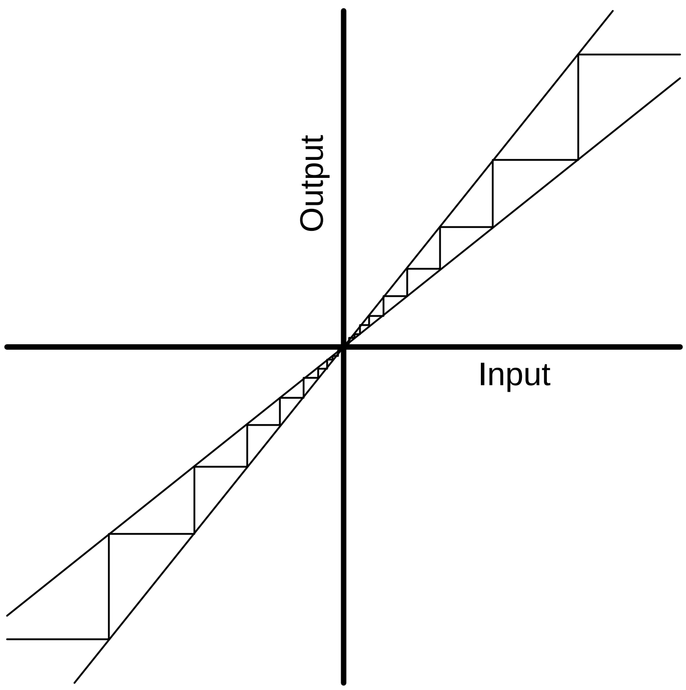

Figure 11. Overall transfer function for instantaneous companding.

There are three commonly used types of companding:

μ-law Companding

$$Y = \frac{\log(1+\mu|x|)}{\log(1+\mu)}$$ where X is the input and Y is the output of the compressor.

A-law Companding

$$Y = \frac{A|X|}{1+\log(A)} \qquad \text{if} \qquad 0 \le X \le \frac{1}{A}$$

$$Y = \frac{1+\log(A|X|)}{1+\log(A)} \qquad \text{if} \qquad \frac{1}{A} \le X \le 1$$

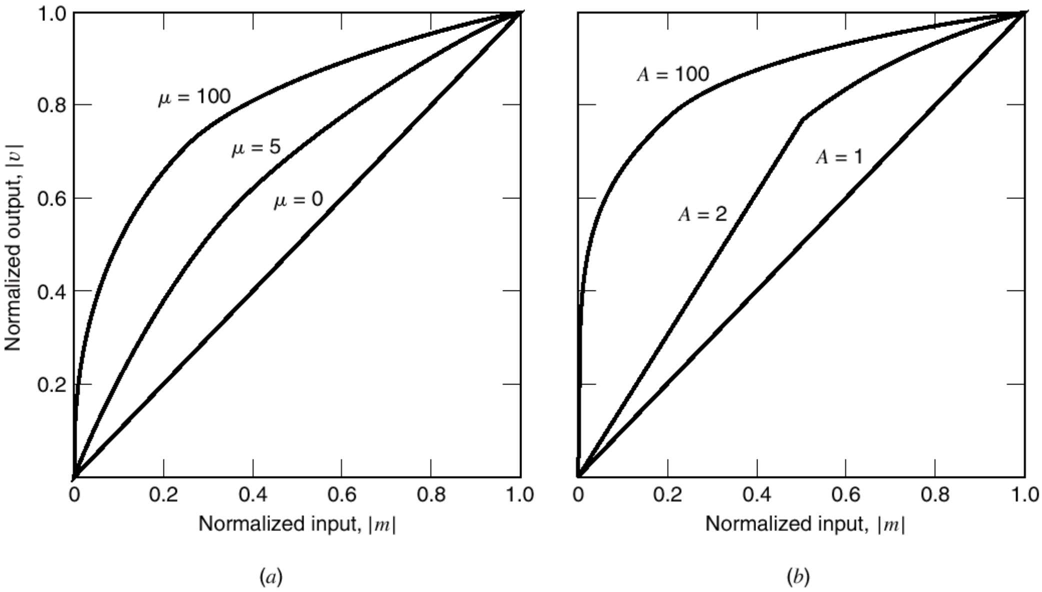

Figure 12. (Haykin Figure 3.14) Compression laws. (*a*) μ-law. (*b*) A-law.

The μ and A law characteristics were developed for analogue compression techniques. It is however difficult to keep an exact match between the compression curve and the expansion curve shown in figure 10. With the advent of low cost large scale IC's, it is much easier to implement this using a higher accuracy ADC and a lookup table to perform the compression. Some ADC's for communication applications include these compression techniques as part of the IC.

The compression and expansion curve can thus best be generated using a lookup table. The ITU (CCITT) 13 segment lookup table is one such standard compression curve and the 15 segment (TI System) used in the USA is another.

ITU (CCITT) 13 segment compression

12 bit input (±2048) and 8 bit output ±128.

256 output levels, straight lines joining break points at:

Y	X	Steps (X scaled by 2048)	First
0 to 32	0 to 1/64	Y at 0.5, 1.5, 2.5 for X at 0.5, 1.5, 2.5 etc	0 + 0.5
33 to 48	1/64 to 1/32	Y at 32.5, 33.5, 34.5 for X at 33, 35, 37 etc	32 + 1
49 to 64	1/32 to 1/16	Y at 49.5, 50.5, 51.5 for X at 66, 70, 74 etc	64 + 2
65 to 80	1/16 to 1/8	Y at 65.5, 66.5, 67.7 for X at 132, 140, 144 etc	128 + 4
81 to 96	1/8 to ¼	Y at 81.5, 82.5, 83.5 for X at 264, 280, 296 etc	256 + 8
97 to 112	¼ to ½	Y at 97.5, 98.5, 99.5 for X at 528, 560, 592 etc	512 + 16
113 to 128	½ to 1	Y at 113.5, 114.5, 115.5 for X at 1056, 1120, 1184 etc	1024 +32

Table 1. CCITT 13 segment compression curve.

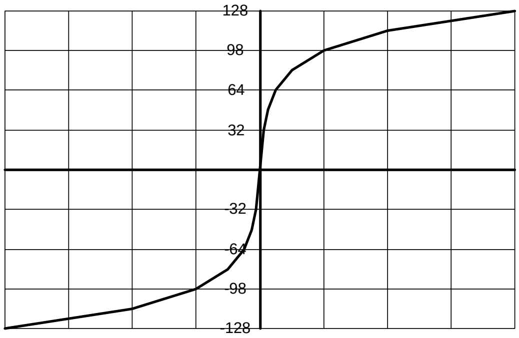

Figure 13. Plot of 13 segment CCITT compression curve.

In the transmitter, the amplitude signal is compressed using the A, μ or ITU curve. At the receiver the signal is expanded with the inverse characteristic, resulting in an overall linear characteristic with the non-uniform quantum levels shown in figures 10 and 11.

The signal is thus COMpressed at the transmitter and exPANDed at the receiver, the combined action of this is called Companding. The above companding is called instantaneous companding since it operates on the amplitude of the input waveform.

Syllabic Companding

In syllabic companding the gain of the modulator and demodulator is changed, to ensure that the input corresponds to the peak SNR at all times. The gain is varied slowly, but fast enough to keep track of the variations of power during each syllable, hence the name syllabic companding. Syllabic comprerssion is used in many audio system, the AGC on microphones in video recorders is one example. In many of those systems the expansion is not done and on a video camera, the output sound power is virtually constant. For code modulation systems, the expansion is normally done. The Companding Ratio is the ratio of the smallest to the largest difference in quantum level, or step size in delta modulation. A minimum of 30 dB companding ratio is desirable.

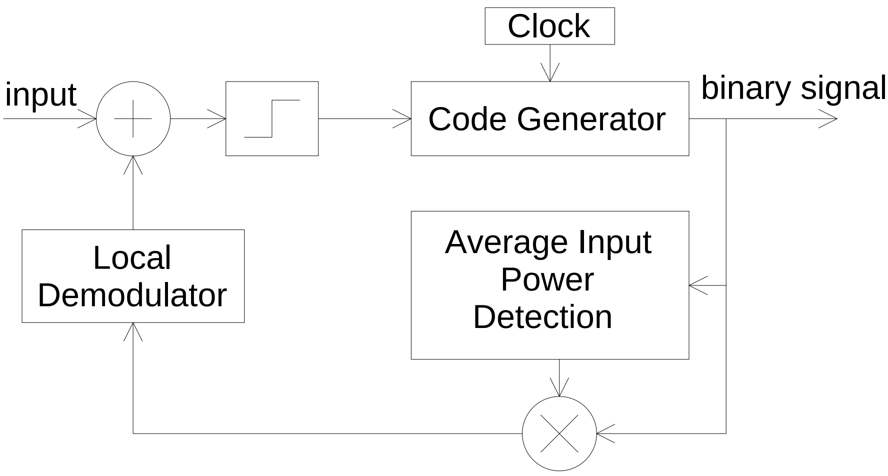

Figure 14. Syllabic Companding Code Modulator

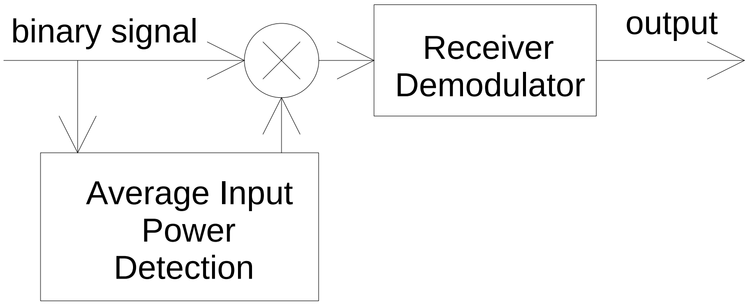

Figure 15. Syllabic Companding Cede Demodulator

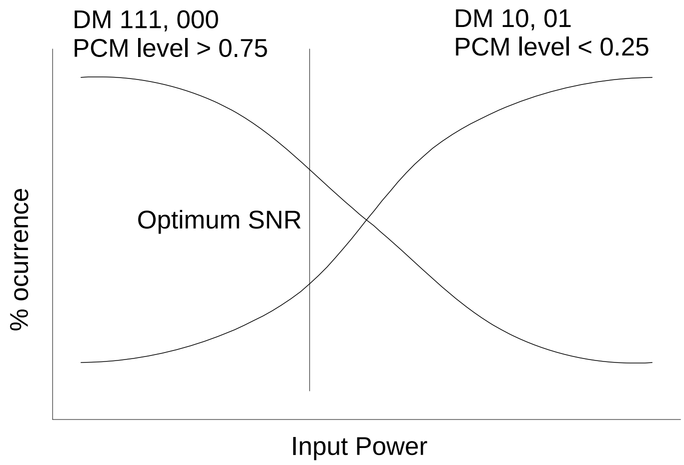

Figure 16. Gain Control Indicators

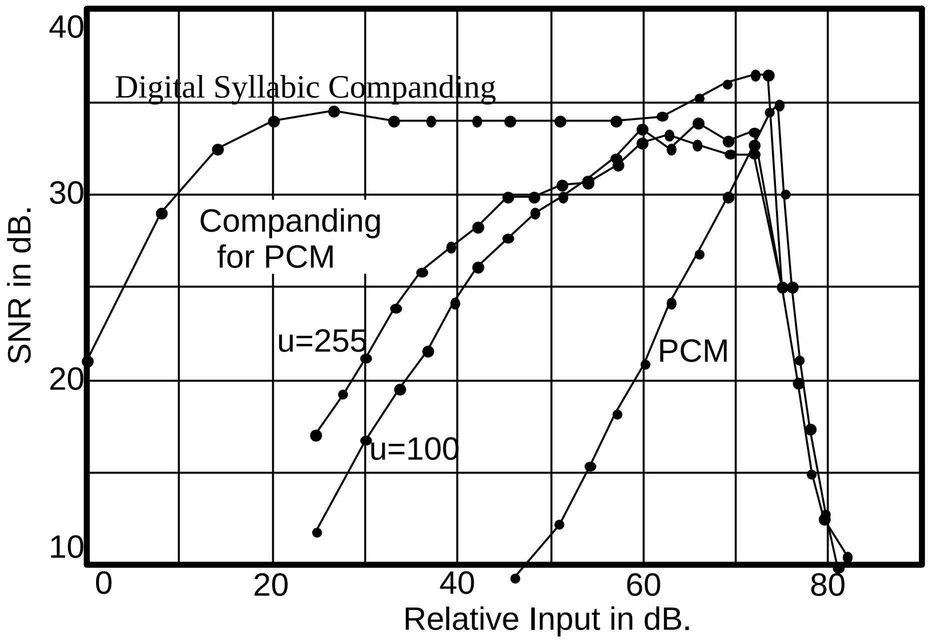

Figure 17. Plot of Companding Performance

Error performance

Using Instantaneous or Syllabic Companding increases the performance of the code modulator but makes the system more susceptible to error. One must ensure that the error is not kept forever. For example an transmission error in a syllabic companding code modulation system will give the wrong gain at the receiver. If the gain limits occasionally by having the maximum or minimum gain setting, then the error is reset.

Vocoders

For voice communication good quality speech at low data rates is obtained using Vocoders. Vocoders attempt to describe speech by modelling the operation of the vocal tract and lips and describing the speech production mechanism in terms of a few independent parameters. Voiced and unvoiced sounds are produced by exciting this model with repetitive pulses or white noise respectively.

The Rectangular Pulse Excitation - Long Term Prediction (RPE-LTP) coder processes speech in 20 ms blocks and then calculates the parameters to represent that block of speech. The RPE-LTP codec operates at a bit rate of 13.2 kbps with a delay of 40 ms and produces good quality speech using a single DSP device. RPE-LPT is used in the European GSM cellular radio system. Vector Sum Excited Linear Predictive (VSELP) coders are used in the North American Digital Cellular radio (NADC) and the Japanese Digital Cellular radio (JDC). This coder produces good quality speech at 8 kbps.

Codebook Excitation Linear Predictive (CELP) coding is proposed for the Australian MobileSat service, a satellite based mobile radio system, and provides good quality speech at 4.8 kbps. In CELP coders the Vocal tract model is excited with a series of stored waveforms rather than rectangular pulses or white noise. The waveforms are stored in a codebook and the one selected matches the currently produced speech as closely as possible. There are many varieties of CELP coders. The USA Department of Defence uses a CELP type coder to produce speech with a quality comparable that produced by the 32 kbps Continuous Variable Slope Delta Modulation (CVSD or CVSDM) system currently used. CVSDM digitises the speech waveform and does not consider the properties of the vocal tract.

The Improved Multi-Band Excitation (IMBE) coder is used in the APCO-25 digital trunk radio system. This coder produces speech at a data rate of 4.4 kbps. Trunk radio is a modern fleet mobile radio system, where a mobile user always communicates to a base station. Examples are police and taxi mobile radio systems. The network is trunked in that the base station and the mobiles can use any of the available frequencies and the frequencies are allocated dynamically, depending on occupancy. With the digital coding and supervisory techniques, secure communications are obtained.

In practice error correcting coding is used to provide a better performance during poor Carrier to Noise Ratio conditions. The error correcting coding will increase the gross data rate to up to 2 times the net data rate. The data rate of the GSM transmissions, which uses a 13.2 kbps coder, is 22.8 kbps after the error correcting codes are applied to part of the data. In addition the data stream will have some overheads, such as channel identification, synchronisation bits, a training sequence for equalisation and other control bits. The amount of this overhead depends on the communication system. For the GSM system 34 overhead bits are used together with 114 data bits to make the 148 bit packet used in each data channel. A 8.25 bit guard period is also used, making the effective data rate 33.8541667 kbps.

Some Speech Coders like the one used in Qualcom’s CDMA cellular radio system use a variable data rate. During active speech the data rate is 9600 bps and during pauses or when the user listens, the rate drops to 1200 bps. Since normal telephony has an activity factor of 40%, a significant reduction in the overall data rate is made.

Not all mobile telephone systems use Predictive Speech Coders. The Cordless telephone systems such as the CT2 system used in the UK, the DECT system used in Europe and the JCT system used in Japan all use Adaptive Delta PCM (ADPCM) coding techniques to convert the speech to a 32 kbps data stream. The advantages of the ADPCM coding are a shorter processing delay, which allow the cordless phones to satisfy the 5 millisecond maximum round trip delay of phones connected to a fixed telephone network. Cordless phones have a very small range, allowing typically 5000 users/km2.