Chapter No. 1

MULTIPLE ACCESS COMMUNICATION

1.1 INTRODUCTION

A limited amount of bandwidth is allocated for wireless services. A wireless system is required to accommodate as many users as possible by effectively sharing the limited bandwidth. For multiple users to be able to share a common resource in a managed and effective way requires some form of access protocol that defines how or when the sharing is to take place and the means for identifying individual messages. Process is known as multiplexing in wired networks and multiple access in wireless digital communications. Therefore, in the field of communications, the term multiple access could be defined as a means of allowing multiple users to simultaneously share the finite bandwidth with least possible degradation in the performance of the system.

1.2 HISTORY OF COMMUNICATION

Traditionally radio communication systems have separated users by either frequency channels, time slots, or both. These concepts date from the earliest days of radio. Even spark transmitters used resonant circuits to narrow the spectrum of their radiation. Scheduled net operation was probably the first manifestation of time slotting. Modern cellular systems began with the use of channelized analog FM. More recently several hybrid FDM-TDM digital systems have been developed, ostensibly to enhance service quality and capacity. In all these systems, each user is assigned a particular time-frequency slot.

In large systems the assignments to the time-frequency slots cannot be unique. Slots must be reused in multiple cells in order to cover large service areas. Satisfactory performance in these systems depends critically on control of the mutual interference arising from the reuse. The reuse concept is familiar even in television broadcasting, where channels are not reused in adjacent cities. [1]

An idealized system geometry is shown in the figure. The same frequency obviously cannot be reused in any adjacent pair of cells because a user on the boundary between those cells would receive both signals with equal amplitude, leading to an unacceptably high interference level. A plane can be tiled with hexagonal cells, labeled in accordance with the seven-way pattern shown in the figure. Thus, if a unique set of channels is assigned to each of the seven cells, then the pattern can be repeated without violating the adjacency requirement. Although this idealized pattern is not strictly applicable in all real systems, the seven-way reuse pattern is approximately correct. The capacity of systems built in this way is determined by the bandwidth per channel and the seven-way reuse pattern. In an AMPS system, therefore, the maximum capacity per cell is approximately 416/7 = 59. For three-way sectored cells, the same K=7 reuse applies over all three sectors, that is, only about 19 channels are available in each sector. In an ideal geometry the reuse pattern looks like this, representing channel sets by distinct colors

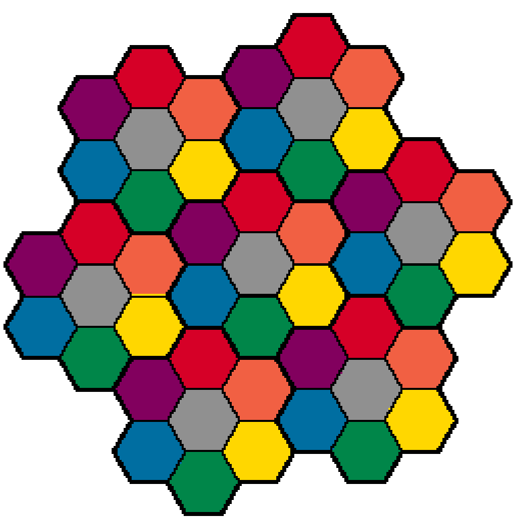

Figure 1: Cell Configuration [1]

In this connection, it should be noted that achievement of the K=7 reuse, rather than an even larger number, depends on the fact that the effective propagation decay law is faster than free space. That is, in a vacuum electromagnetic radiation decays in intensity like R^{-2}. However measurements have consistently shown that the effective propagation law exponent is typically between -3.5 and -5 in the ground mobile environment. Interestingly, it is easy to show that if the propagation law were that of free space, a large cellular system would not be viable at all. The larger-than-free-space propagation exponent means that only the first tier of neighbor cells is significant in the idealized model.

1.3 MULTIPLE-ACCESS TECHNIQUES

The goal in the design of cellular systems is to be able to handle as many calls as possible in a given bandwidth with some reliability. There are several different ways to allow access to the channel.[3] These include the following.

- frequency division multiple-access (FDMA)
- time division multiple-access (TDMA)
- time/frequency multiple-access
- code division multiple-access (CDMA)
    - frequency-hop CDMA
    - direct-sequence CDMA
    - multi-carrier CDMA (FH or DS)

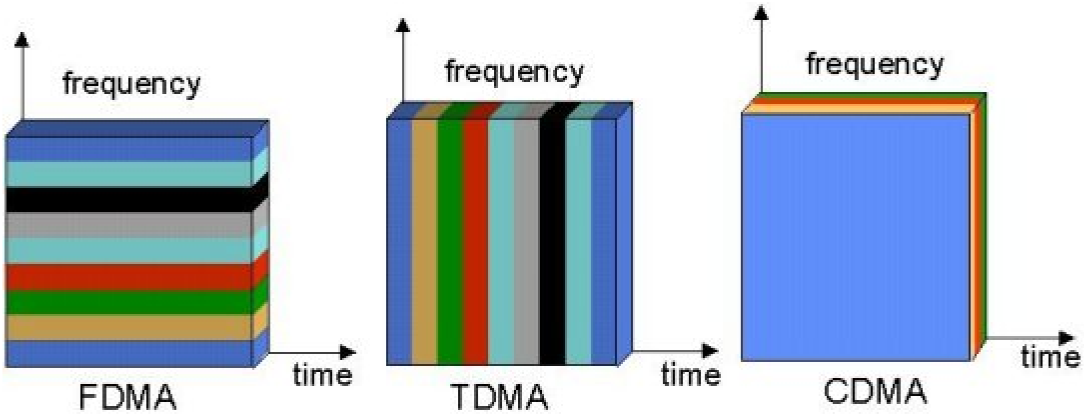

Figure 2: Multiple Access Schemes [3]

1.3.1 Frequency Division Multiple Access (FDMA)

FDMA is one of the three main multiplexing techniques that enable users to share the radio spectrum. In the case of FDMA the receiver will discriminate between the signals by tuning to the particular frequency channel that carries the desired signal. Alternative multiple access techniques are TDMA and CDMA, in which the receivers discriminate between signals by using, respectively, different time slots or different codes. However, in practice TDMA and CDMA are always found in combination with FDMA, i.e. TDMA or CDMA are used to increase capacity on a channel within an FDMA system.

Frequency Division Multiple Access (FDMA) is an analogue transmission technique used for mobile phone communications, in which the frequency band allocated to a network is divided into sub-bands or channels. Each frequency channel can carry either a voice conversation or digital data, and one channel will be assigned to each subscriber for the duration of a call. Using FDMA in this way, multiple users can share the available band without the risk of interference between the simultaneous calls.[5]

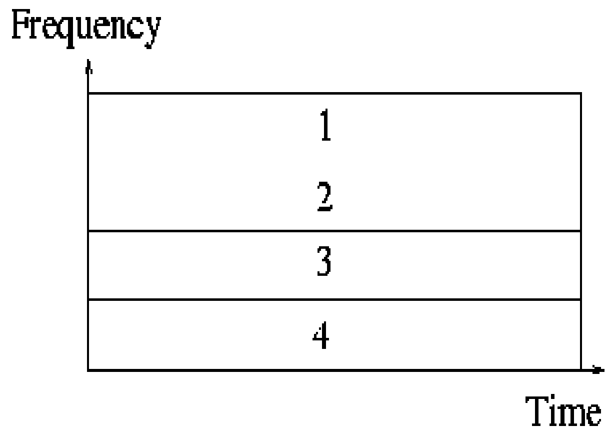

Figure 3: FDMA [5]

<u>Drawbacks</u>

FDMA analog transmissions are the least efficient cellular networks since each analog channel can only be used one user at a time. Analog channels don't take full advantage of band-width. Not only are these FDMA channels larger than necessary given modern digital voice compression, but they are also wasted whenever there is silence during a cell phone conversation. Analog signals are especially susceptible to noise and the extra noise cannot get filtered out [6]. Given the nature of the signal, analog cell phones must use higher power (between 1 and 3 watts) to get acceptable call quality. Given these analog features, it is easy to see why FDMA is being replaced by newer digital networks such as TDMA and CDMA.

1.3.2 Time Division Multiple Access

TDMA is a technology used in digital cellular telephone communication that divides each cellular channel into three time slots in order to increase the amount of data that can be carried. TDMA builds on FDMA (Frequency Division Multiple Access) by dividing conversations by frequency and time. Since digital compression allows voice to be sent under 10 kilobits per second (equivalent to 10 kHz), TDMA can fit three digital conversations into a FDMA/Analog channel (which is 30 kHz) [7]. By sampling a person's voice for, say 30 milliseconds, then transmitting it in 10 milliseconds; the system is able to offer 3 timeslots per channel in a round-robin fashion. This technique allows compatibility with FDMA while enabling digital services and easily boosting the system by 3 times capacity. It increases the efficiency of the network, leads to increase calls, more users on the network and better cellular quality [8].

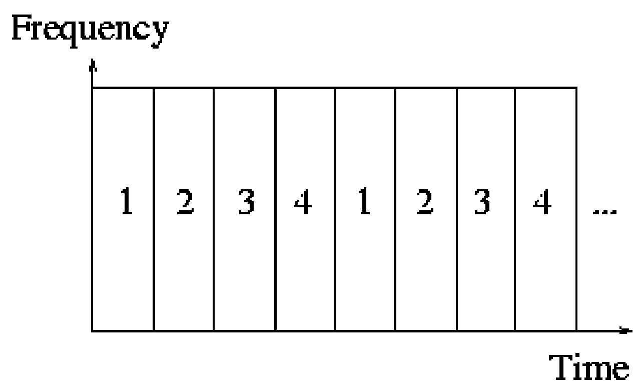

Figure 4: TDMA [5]

Drawbacks

While TDMA is a good digital system, it is still somewhat inefficient since it has no flexibility for varying digital data rates (high quality voice, low quality voice, pager traffic) and has no accommodations for silence in a telephone conversation. In other words, once a call is initiated, the channel/timeslot pair belongs to the phone for the duration of the call. TDMA also requires strict signaling and timeslot synchronization. A digital control channel provides synchronization functionality as well as adding voice mail and message notification. Due to the digital signal, TDMA phones need only broadcast at 600 milliwatts [11].

1.3.3 Code Division Multiple Access (CDMA)

CDMA, for Code Division Multiple Access, is different than those traditional ways in that it does not allocate frequency or time in user slots but gives the right to use both to all users simultaneously. To do this, it uses a technique known as *Spread Spectrum*. In effect, each user is assigned a code which spreads its signal bandwidth in such a way that only the same code can recover it at the receiver end. This method has the property that the unwanted signals with different codes get spread even more by the process, making them like noise to the receiver [5].

<u>Spread Spectrum</u>

Spread Spectrum is a mean of transmission where the data occupies a larger bandwidth than necessary. Bandwidth spreading is accomplished before the transmission through the use of a code which is independent of the transmitted data. The same code is used to demodulate the data at the receiving end. The following figure illustrate the spreading done on the data signal x(t) by the spreading signal c(t) resulting in the message signal to be transmitted, m(t).

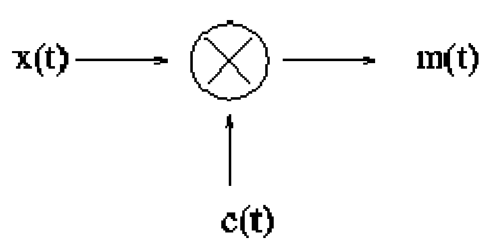

Figure 5: Formation of transmitted signal [5]

Originally for military use to avoid jamming (interference created on purpose to make a communication channel unusable), spread spectrum modulation is now used in personal communication systems for its superior performance in an interference dominated environment.

In spread spectrum, the data is modulated by a spreading signal which uses more bandwidth than the data signal. Since multiplication in the time domain corresponds to convolution in the frequency domain, a narrow band signal multiplied by a wide band signal ends up being wide band.

Here the three signals corresponds to x(t), c(t) and m(t) . The first two signals are multiplied together to give the third waveform.

TYPES OF SPREAD SPECTRUM COMMUNICATIONS

In general, the most popular spread spectrum techniques are categorized into three groups; these are direct sequence, DS, frequency hopping, FH, and a hybrid, DS/FH.

FREQUENCY HOPPING

Frequency hopping is one of two basic modulation techniques used in spread spectrum signal transmission. It is the repeated switching of frequencies during radio transmission, often to minimize the effectiveness of "electronic warfare" - that is, the unauthorized interception or jamming of telecommunications [5]. The signal is rapidly switched between different frequencies within the hopping bandwidth pseudo-randomly, and the receiver knows before hand where to find the signal at any given time. In an FH-CDMA system, a transmitter "hops" between available frequencies according to a specified algorithm, which can be either random or preplanned. The transmitter operates in synchronization with a receiver, which remains tuned to the same center frequency as the transmitter. A short burst of data is transmitted on a narrowband. Then, the transmitter tunes to another frequency and transmits again. The receiver thus is capable of hopping its frequency over a given bandwidth several times a second, transmitting on one frequency for a certain period of time, then hopping to another frequency and transmitting again. Frequency hopping requires a much wider bandwidth than is needed to transmit the same information using only one carrier frequency [8].

The process of frequency hopping is illustrated below:

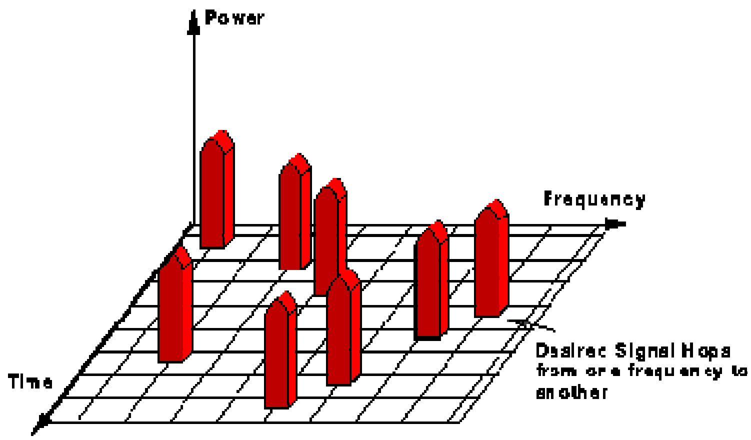

Figure 6: Illustration of the frequency hopping concept [10]

There are two sub-classes of such system, slow frequency hopping, SFH, and fast frequency hopping, FFH. While for the SFH, the hopping rate is of the order of one hop every many symbols, the FFH system completes many hops per symbol. Similar to the DS case, the intended receiver knows the exact hopping pattern and therefore hops in synchronism with the transmitter to extract the data.[5]

DIRECT SEQUENCE

The spread spectrum approach that is an alternative to FH-CDMA is direct sequence code division multiple access which chops the data into small pieces and spreads them across the frequency domain.

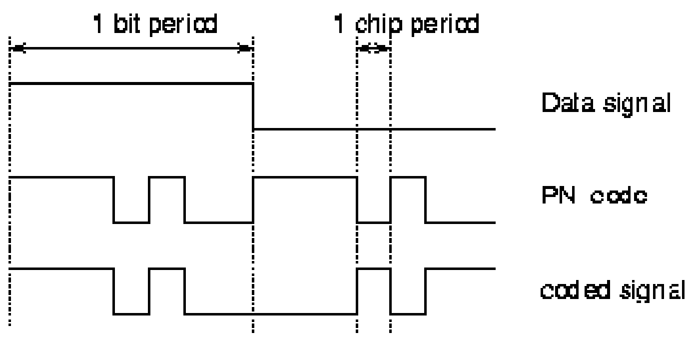

Figure 7: Direct Sequence Spreading [10]

In direct sequence spread spectrum, the stream of information to be transmitted is divided into small pieces, each of which is allocated across to a frequency channel across the spectrum. The digital data is directly coded at a much higher frequency. A data signal at the point of transmission is combined with a higher data-rate bit sequence (also known as a *chipping code*) that divides the data according to a spreading ratio. The redundant chipping code helps the signal resist interference and also enables the original data to be recovered if data bits are damaged during transmission [11].

The code is generated pseudo-randomly, the receiver knows how to generate the same code, and correlates the received signal with that code to extract the data. In the DS type, the data stream first modulates the carrier and then the resultant carrier is combined with a pseudo-noise, PN, sequence. The elements of the PN sequence are normally chosen such that the cross-correlation between different PN sequences is kept very small. Only the intended receiver has a replica of the PN sequence, which is used to extract the transmitted information using a correlator.

The figure shows how a single direct sequence spread spectrum communication channel can have several channels. In this example, there are 3 different code patterns that are used for communication channels. When a receiver uses the reference code, a direct sequence spread spectrum system can build a mask as shown in this figure for each conversation allowing only that information which falls within the mask to be transmitted or received [8].

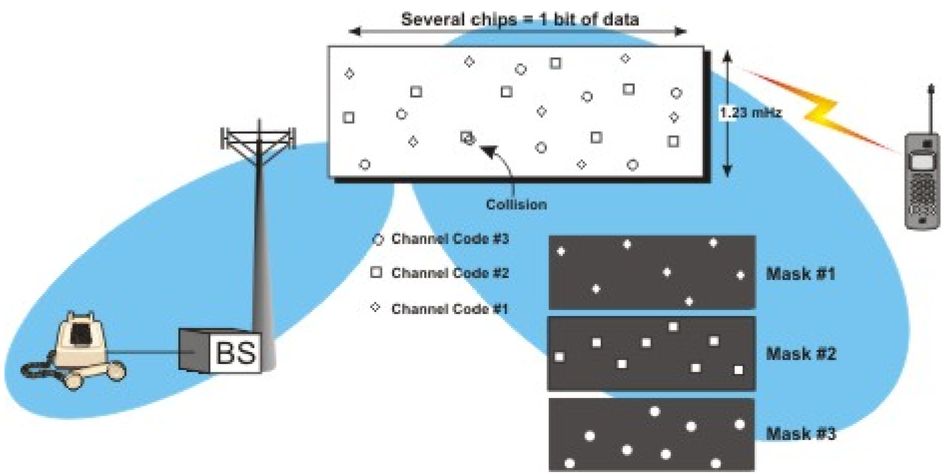

Figure 8: Direct Sequence Spread Spectrum Communication Channels [8]

<u>DS/FH SPREAD SPECTRUM TECHNIQUE</u>

The DS/FH Spread Spectrum technique utilizes both of the former two techniques to achieve both time and frequency diversity, but of course at the expense of extra system complexity. In this system, first DS spreading is employed followed by FH. It combines the advantages of the direct-sequence and frequency-hopping spreading techniques while reducing their shortcomings.

Each data symbol is divided over one or more N_{FH} frequency-hop carriers (carrier frequencies). In each frequency-hop channel a complete pseudo-noise sequence of length N is combined with the data signal (see figure where N_{FH} is taken to be 1).

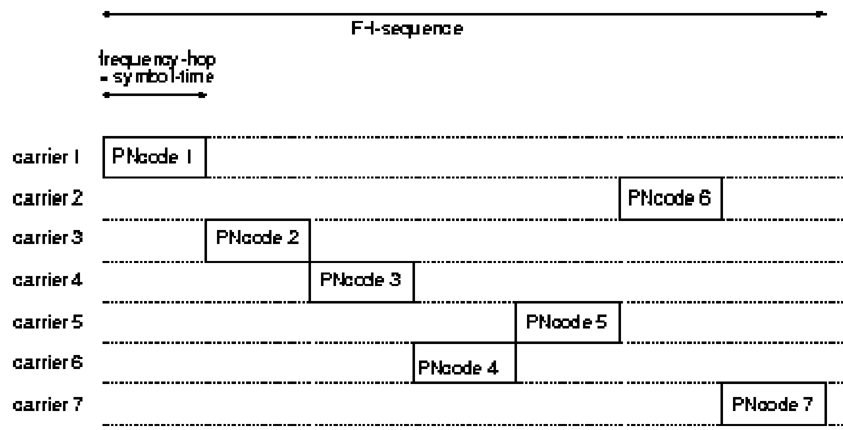

Figure 9: Example of a DS/FFH Spreading scheme [12]

Since the FH-sequence and the PN-codes are coupled, every receiver is identified by a combination of an FH-sequence and a set of PN-codes. To limit the probability that two users share the same frequency channel simultaneously, frequency-hop sequences are chosen in such a way that two transmitters with different FH-sequences share at most two frequencies at the same time during one FH-sequence [13].

Applying several frequency hops within each data bit disqualifies modulation by some kind of Phase Shift Keying. PSK is also quite susceptible to channel distortions. A 16-FSK modulation scheme is therefore chosen.

1.4 SUMMARY

In the present era of telecommunication the prior most requirement is to make a bandwidth efficient multiple access based communication systems. The mostly used multi user technique is TDMA, FDMA and now the most popular is CDMA. CDMA is the inherently secure technique. CDMA is further divided into frequency hopping (FH) and spread spectrum (SS) techniques. There is a unique code assigned to each user generated by the Pseudo-Random Noise (PN) generator.

Chapter No. 2

CODE DIVISION MULTIPLE ACCESS

2.1 INTRODUCTION

The world is demanding more from wireless communication technologies than ever before. More people around the world are subscribing to wireless services and consumers are using their phones more frequently. Add in exciting Third-Generation (3G) wireless data services and applications - such as wireless email, web, digital picture taking/sending and assisted-GPS position location applications - and wireless networks are asked to do much more than just a few years ago. And these networks will be asked to do more tomorrow. [14]

This is where CDMA technology fits in. The great attraction of CDMA technology from the beginning has been the promise of extraordinary capacity increase over narrowband multiple access wireless technologies. Simple models suggest that the capacity improvement may be more than 20 times that of the existing narrowband cellular standards, such as AMPS in North America, NMT in Scandinavia, and TACS in the United Kingdom.

CDMA consistently provides better capacity for voice and data communications than other commercial mobile technologies, allowing more subscribers to connect at any given time, and it is the common platform on which 3G technologies are built. [1]

2.2 CDMA PROCESS

CDMA employs analog-to-digital conversion (ADC) in combination with spread spectrum technology. Audio input is first digitized into binary elements. The frequency of the transmitted signal is then made to vary according to a defined pattern (code), so it can be intercepted only by a receiver whose frequency response is programmed with the same code, so it follows exactly along with the transmitter frequency. There are trillions of possible frequency-sequencing codes; this enhances privacy and makes cloning difficult.

The CDMA channel is nominally 1.23 MHz wide. CDMA networks use a scheme called soft handoff, which minimizes signal breakup as a handset passes from one cell to another. The combination of digital and spread-spectrum modes supports several times as many signals per unit bandwidth as analog modes. CDMA is compatible with other cellular technologies; this allows for nationwide Roaming.

The original CDMA standard, also known as CDMA One and still common in cellular telephones in the U.S., offers a transmission speed of only up to 14.4 Kbps in its single channel form and up to 115 Kbps in an eight-channel form. CDMA2000 and wideband CDMA deliver data many times faster. [15]

2.2.1 Input data

CDMA works on Information data from several possible sources, such as digitized voice or ISDN channels. Data rates can vary, here are some examples:

Voice	Pulse Code Modulation (PCM)
	Differential Pulse Code Modulation (DPCM)
	Adaptive Differential Pulse Code Modulation (ADPCM)
	Low Delay Code Excited Linear Prediction (LD-CELP)
ISDN	Bearer Channel (B-Channel)
	Data Channel (D-Channel)

Table 1: Data Source [16]

The system works with 64 kBits/sec data, but can accept input rates of 8, 16, 32, or 64 kBits/sec. Inputs of less than 64 kBits/sec are padded with extra bits to bring them up to 64 kBits/sec.

2.2.2 Generating Pseudo-Random Codes

For each channel the base station generates a unique code that changes for every connection. The base station adds together all the coded transmissions for every subscriber. The subscriber unit correctly generates its own matching code and uses it to extract the appropriate signals. CDMA codes are not required to provide call security, but create a uniqueness to enable call identification. Codes should not correlate to other codes or time shifted version of itself. Spreading codes are noise like pseudo-random codes, channel codes are designed for maximum separation from each other and cell identification codes are balanced not to correlate to other codes of itself. In order for all this to occur, the pseudo-random code must have the following properties:

1. It must be deterministic. The subscriber station must be able to independently generate the code that matches the base station code.

2. It must appear random to a listener without prior knowledge of the code (i.e. it has the statistical properties of sampled white noise).

3. The cross-correlation between any two codes must be small.

4. The code must have a long period (i.e. a long time before the code repeats itself).

Code Correlation

In general the correlation function has these properties:

- It equals 1 if the two codes are identical
- It equals 0 of the two codes have nothing in common

Intermediate values indicate how much the codes have in common. The more they have in common, the harder it is for the receiver to extract the appropriate signal.

There are two correlation functions:

- Cross-Correlation: The correlation of two different codes. This should be as small as possible.
- Auto-Correlation: The correlation of a code with a time-delayed version of itself. In order to reject multi-path interference, this function should equal 0 for any time delay other than zero.

The receiver uses cross-correlation to separate the appropriate signal from signals meant for other receivers, and auto-correlation to reject multi-path interference. [16]

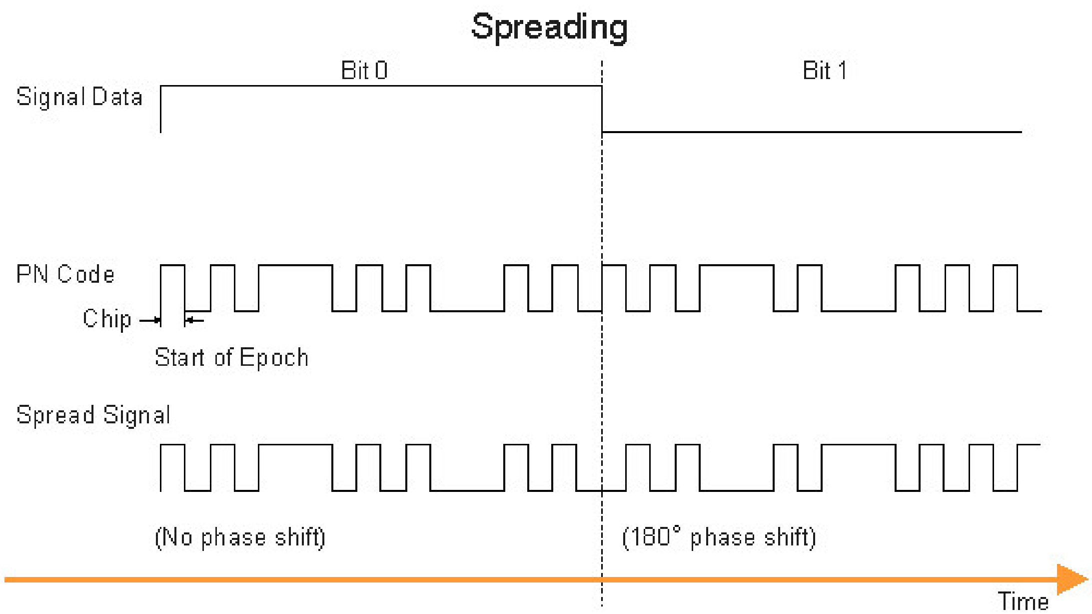

Figure 10: Spreading [16]

2.2.3 Pseudo-Noise Spreading

The spreading process is done by directly combining the baseband information to high chip rate binary code. The Spreading Factor is the ratio of the chips to baseband information rate .CDMA uses unique spreading codes to spread the base band data before transmission. The signal is transmitted in a channel, which is below noise level. The receiver then uses a correlator to despread the wanted signal, which is passed through a narrow bandpass filter. Unwanted signals will not be despread and will not pass through the filter. Codes take the form of a carefully designed one/zero sequence produced at a much higher rate than that of the baseband data. The rate of a spreading code is referred to as chip rate rather than bit rate. [24]

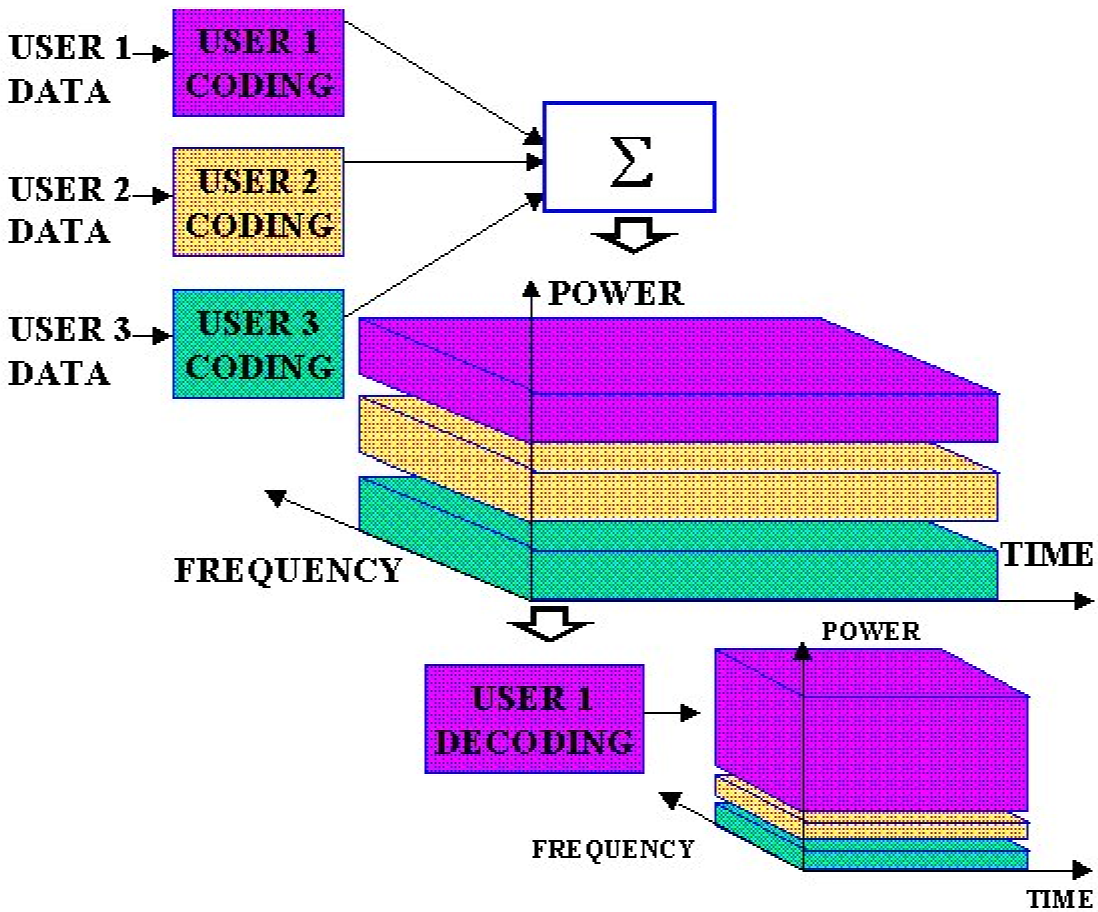

Figure 11: CDMA spreading [3]

Figure below shows the process of frequency spreading. In general, the bandwidth of a digital signal is twice its bit rate. The bandwidths of the information data (f_i) and the PN code are shown together. The bandwidth of the combination of the two, for Chipping Frequency (f_c)> Information rate (f_i), can be approximated by the bandwidth of the PN code. [3]

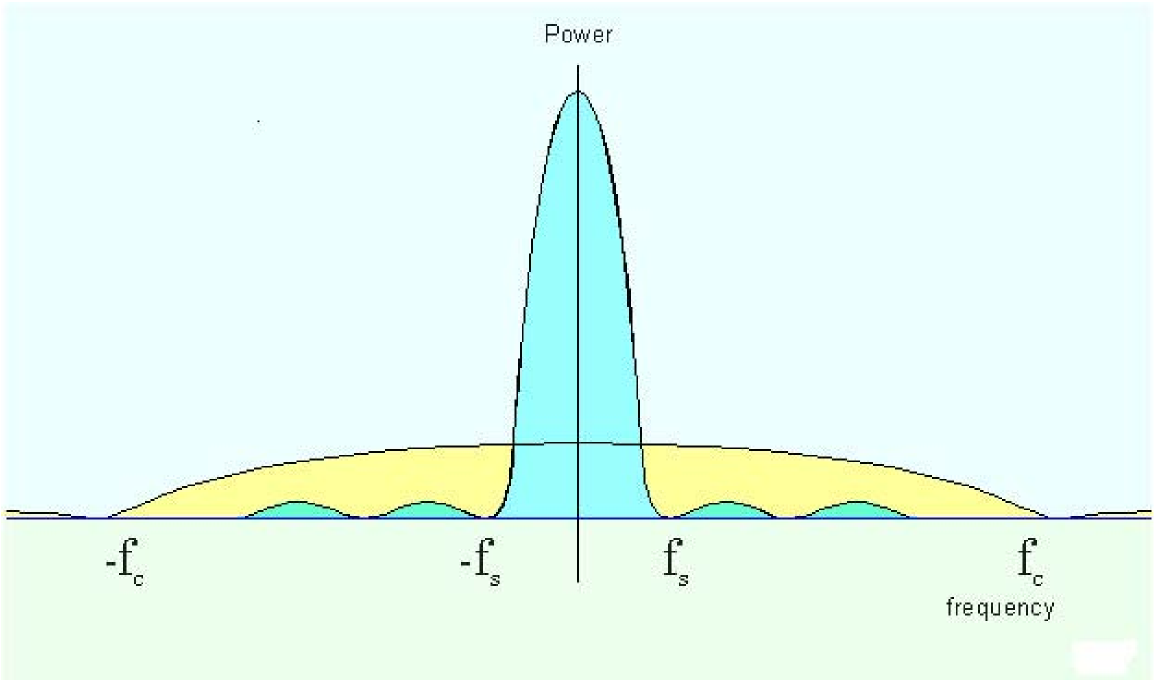

Figure 12: Frequency Spreading [16]

2.2.4 Processing Gain

An important concept relating to the bandwidth is the processing gain (G_p). This is a theoretical system gain that reflects the relative advantage that frequency spreading provides. The processing gain is equal to the ratio of the chipping frequency to the data frequency:

$$G_p \quad \frac{f_c}{f_i}$$

There are two major benefits from high processing gain:

- Interference rejection: the ability of the system to reject interference is directly proportional to G_p.
- System capacity: the capacity of the system is directly proportional to G_p.

So the higher the PN code bit rate (the wider the CDMA bandwidth), the better the system performance.

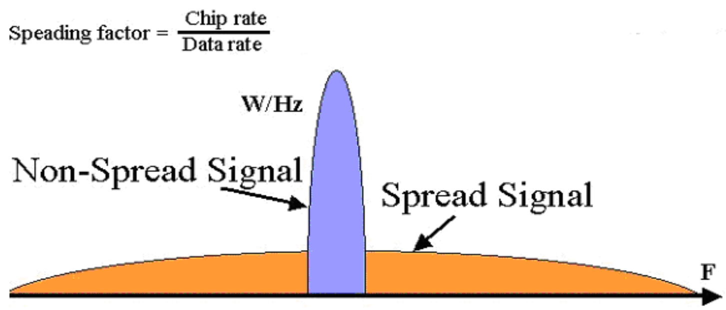

Figure 13: Spread and Non- Spread Signals [16]

2.2.5 Transmitting Data

The resultant coded signal next modulates an RF carrier for transmission using Quadrature Phase Shift Keying (QPSK). QPSK uses four different states to encode each symbol. The four states are phase shifts of the carrier spaced 90_ apart. By convention, the phase shifts are 45, 135, 225, and 315 degrees. Since there are four possible states used to encode binary information, each state represents two bits. This two bit “word” is called a symbol. Figure shows in general how QPSK works.

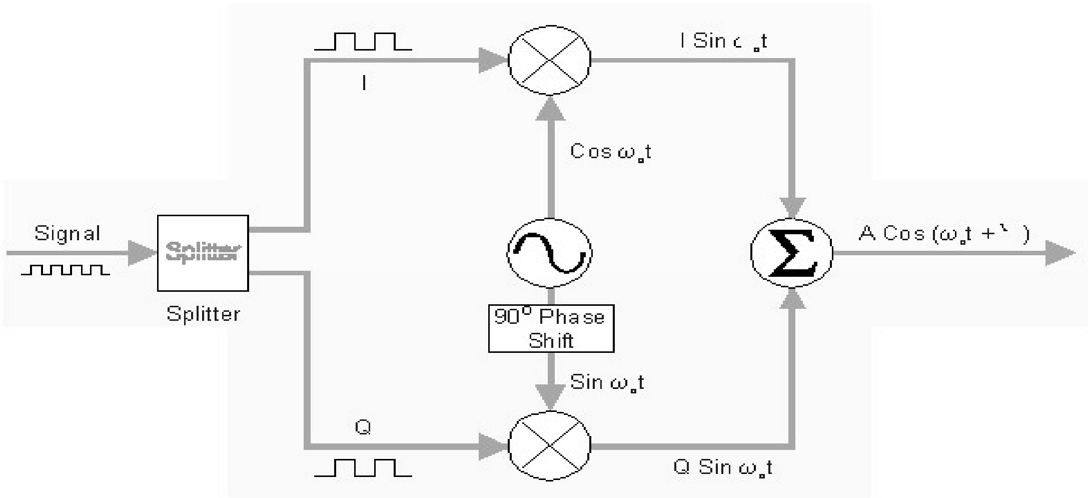

Figure 14: Working of QPSK [3]

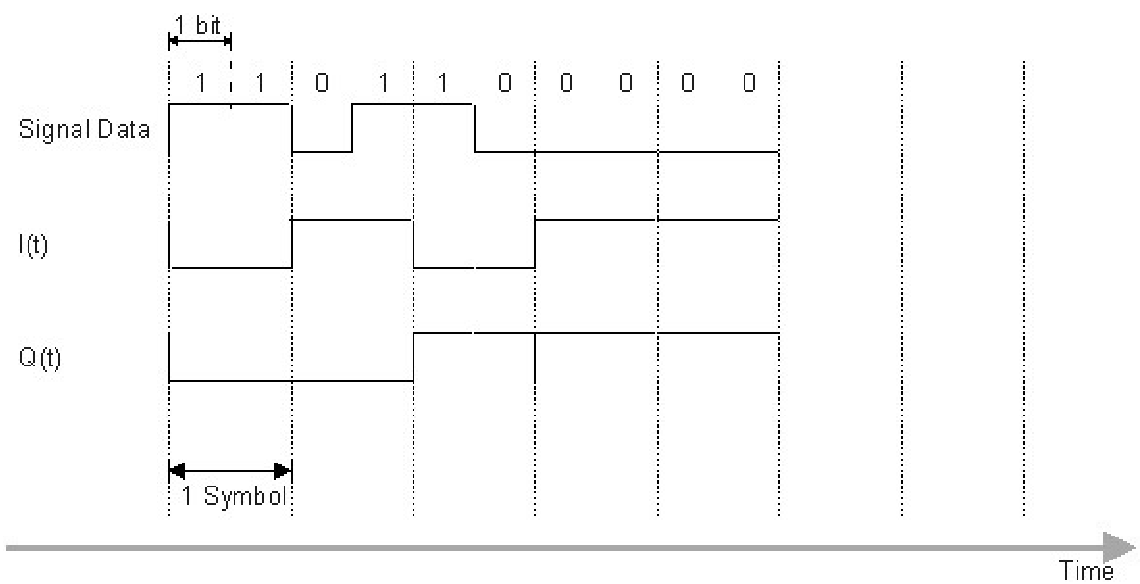

Figure 15: Modulated Signal [3]

2.2.6 Receiving Data

The receiver performs the following steps to extract the Information:

- Demodulation
- Code acquisition and lock
- Correlation of code with signal
- Decoding of Information data

Demodulation

The receiver generates two reference waves, a Cosine wave and a Sine wave. Separately mixing each with the received carrier, the receiver extracts I(t) and Q(t). Analog to Digital converters restore the 8-bit words representing the I and Q chips.

Code Acquisition and Lock

The receiver, as described earlier, generates its own complex PN code that matches the code generated by the transmitter. However, the local code must be phase-locked to the encoded data. The RCS (Reaction Control System) and FSU each have different ways of acquiring and locking onto the other’s transmitted code.

Correlation and Data Despreading

Once the PN code is phase-locked to the pilot, the received signal is sent to a correlator that multiplies it with the complex PN code, extracting the I and Q data meant for that receiver. The receiver reconstructs the Information data from the I and Q data.

2.3 SYSTEM CAPACITY

CDMA offers an answer to the capacity problem. The key to its high capacity is the use of noise-like carrier waves, as was first suggested decades ago by Claude Shannon.

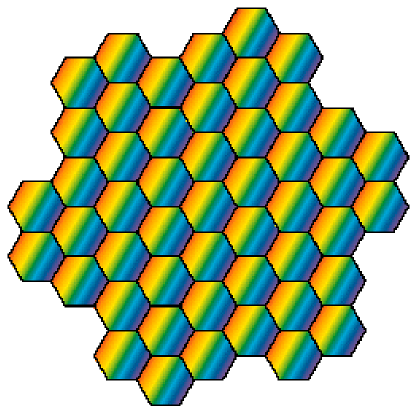

Figure 16: Reuse pattern for CDMA [1]

The rainbow cells indicate that the entire 1.25 MHz passband is used by each user, and that same passband is reused in each cell.
Capacity is determined by the balance between the required SNR for each user, and the spread spectrum *processing gain.* The figure of merit of a well-designed digital receiver is the dimensionless signal-to-noise ratio (SNR)[1].

$$E_b/N_0 = \frac{\text{Energy per bit}}{\text{Power spectral density of noise + interference}}$$

2.4 POWER CONTROL

CDMA is interference limited multiple access system. Because all users transmit on the same frequency, internal interference generated by the system is the most significant factor in determining system capacity and call quality. The transmit power for each user must be reduced to limit interference, however, the power should be enough to maintain the required Eb/No (signal to noise ratio) for a satisfactory call quality. Maximum capacity is achieved when Eb/No of every user is at the minimum level needed for the acceptable channel performance. As the MS moves around, the RF environment continuously changes due to fast and slow fading, external interference, shadowing, and other factors. The aim of the dynamic power control is to limit transmitted power on both the links while maintaining link quality under all conditions. Additional advantages are longer mobile battery life and longer life span of BTS power amplifiers.

2.5 HANDOVER

Handover occurs when a call has to be passed from one cell to another as the user moves between cells. In a traditional "hard" handover, the connection to the current cell is broken, and then the connection to the new cell is made. This is known as a "break-before-make" handover. Since all cells in CDMA use the same frequency, it is possible to make the connection to the new cell before leaving the current cell. This is known as a "make-before-break" or "soft" handover. Soft handovers require less power, which reduces interference and increases capacity. Mobile can be connected to more that two BTS the handover. "Softer" handover is a special case of soft handover where the radio links that are added and removed belong to the same Node B.

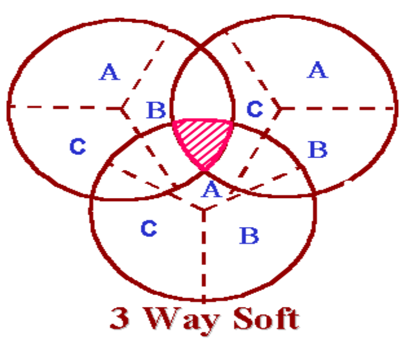

Figure 17: CDMA soft handover [3]

2.6 MULTIPATH AND RAKE RECEIVERS

One of the main advantages of CDMA systems is the capability of using signals that arrive in the receivers with different time delays. This phenomenon is called multipath. FDMA and TDMA, which are narrow band systems, cannot discriminate between the multipath arrivals, and resort to equalization to mitigate the negative effects of multipath. Due to its wide bandwidth and rake receivers, CDMA uses the multipath signals and combines them to make an even stronger signal at the receivers. CDMA subscriber units use rake receivers. This is essentially a set of several receivers. One of the receivers (fingers) constantly searches for different multipaths and feeds the information to the other three fingers. Each finger then demodulates the signal corresponding to a strong multipath. The results are then combined together to make the signal stronger. [3]

2.7 FUTURE PROJECTIONS

With the current rate of development, it is obvious that CDMA will displace TDMA as the primary wireless multiple access technology. Because CDMA is ideally suited for the high traffic demands of metropolitan areas, CDMA systems will dominate the markets. TDMA is a proven technology that has already been widely implemented and supported. For this reason it will remain a viable alternative for markets with lower capacity demands. For the savvy investor, now is a prime time to invest in a growing technology that is sure to provide profitable returns. This technology is, of course, CDMA.

2.8 SUMMARY

CDMA is the multiple access technique in which every user in the network is assigned a unique code. Voice is converted to bits using any of the pulse code modulation techniques. The bits go through the channel coding, which normally is the convolution encoding. The information bits are spread by multiplying the unique code known as the Pseudo Random Noise (PN) with the information bits normally using the Walsh codes. At the receiver end the rake receivers are used for the multipath rejection. Viterbi decoders are used for decoding the convolution codes. The actual voice is then recovered by decoding the DPCM.

Chapter No. 3

SPEECH CODING

3.1 INTRODUCTION

Although humans are well equipped for analog communications, analog transmission is not particularly efficient. When analog signals become weak because of transmission loss, it is hard to separate the complex analog structure from the structure of random transmission noise. If analog signals are amplified, it also amplifies noise, and eventually analog connections become too noisy to use. Digital signals, having only "one-bit" and "zero-bit" states, are more easily separated from noise. They can be amplified without corruption.

A prerequisite for digital transmission systems is that the information to be transmitted can be converted into a sequence of pulse combinations, which are then transmitted practically without any noticeable distortion. Consequently, analog information - such as human speech - must be converted into digital form.

The accuracy of A/D conversion is crucial to the subscriber's perceived quality. The digit combination must be so detailed that the analog speech (or video) can be reproduced without distortion or disturbances in the receiving equipment. At the same time, it is aimed to reduce the amount of digital information in order to better utilize the available network capacity.

3.2 PULSE CODE MODULATION (PCM)

The Pulse Code Modulation (PCM) is the most bit-consuming speech coding technique. Nevertheless, the PCM is the most widely deployed in the telephone network because it is the first telephone speech coding system developed, and, for that reason, the PCM is commonly regarded as the reference system against which other speech coding systems may be calibrated[17]. It is the simplest, low-complexity speech coding technique. It was invented as early as the 1930s, but did not start to predominate until the 1960s when integrated transistor circuits became available. PCM is a type of waveform coding and is standard for voice coding in the telephone network. The bit rate generated per call - 64 kbit/s - has been a decisive factor in switching and transmission design.

3.2.1 Anti-Aliasing Filtering

The first step to convert the signal from analog to digital is to filter out the higher frequency components of the signal. This makes things easier downstream to convert this signal. Most of the energy of spoken language is somewhere between 200 or 300 hertz and about 2700 or 2800 hertz. Roughly 3000-hertz bandwidth for standard speech and standard voice communication is established. Therefore, they do not have to have precise filters (it is very expensive). A bandwidth of 4000 hertz is made from an equipment point if view. This band-limiting filter is used to prevent aliasing. Aliasing happens when the input analog voice signal is undersampled, defined by the Nyquist criterion as $Fs < 2(BW)$ i.e. the sampling frequency is less than the highest frequency of the input analog signal. This creates an overlap between the frequency spectrum of the samples and the input analog signal. The low-pass output filter, used to reconstruct the original input signal, is not smart enough to detect this overlap. Therefore, it creates a new signal that does not originate from the source. This creation of a false signal when sampling is called **aliasing**[18].

3.2.2 Sampling

Reading the amplitude at regular intervals is called sampling. The second step to convert an analog voice signal to a digital voice signal is to sample the filtered input signal at a constant sampling frequency. It is accomplished by using a process called pulse amplitude modulation (PAM). This step uses the original analog signal to modulate the amplitude of a pulse train that has a constant amplitude and frequency. It is important to take the samples on the voice curve at suitable intervals, which means that the quality obtained should allow us to clearly recognize each other's voices. The sampling rate must be sufficiently large so that the analog signal can be reconstructed from the samples with sufficient accuracy. But even taking too many samples is uneconomical. A suitable sampling frequency is 8,000 samples per second. The result will be a pulse amplitude-modulated (PAM) signal where each pulse directly corresponds to the amplitude of the voice curve.

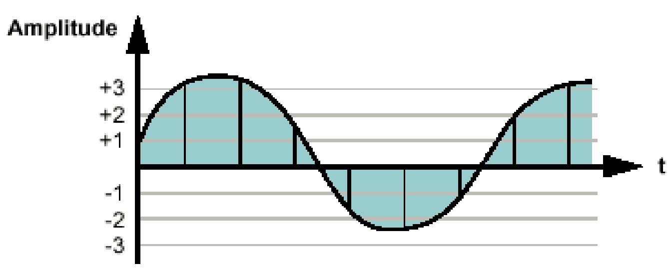

Figure 18: Original Analog Signal [19]

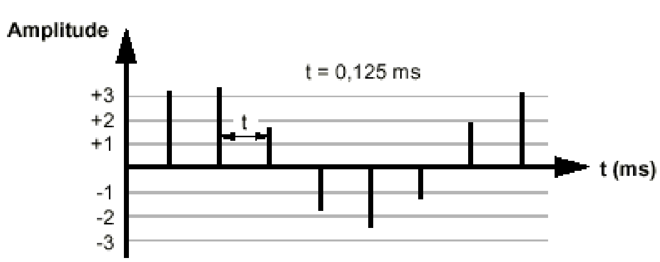

Figure 19: Pulse amplitude-modulated signal [19]

3.2.3 The Sampling Theorem

How is the sampling frequency determined? And how have we reached the conclusion that a frequency of 8,000 samples per second is sufficiently close between readings? A scientist by the name of Harry Nyquist discovered that the original analog signal can be reconstructed if enough samples are taken. He determined that if the sampling frequency is at least twice the highest frequency of the original input analog voice signal, this signal can be reconstructed by a low-pass filter at the destination. The Nyquist criterion is stated like this:

Fs > 2(BW)

Fs = Sampling frequency

BW = Bandwidth of original analog voice signal[18]

Since telephone connections operate in the 300-3,400 Hz band, 8,000 Hz is a sampling frequency that meets the primary requirement for transmission quality: no information should be lost. The sampling frequency is twice the maximum frequency, which is significantly lower than 8 kHz.

3.2.4 Quantization

Quantization is the process of converting each analog sample value into a discrete value that can be assigned a unique digital code word[18].

Quantization means measuring the amplitude of the pulses in the PAM curve and assigning a numerical value to each pulse. As the input signal samples enter the quantization phase, they are assigned to a quantization interval. All quantization intervals are equally spaced (uniform quantization) throughout the dynamic range of the input analog signal.

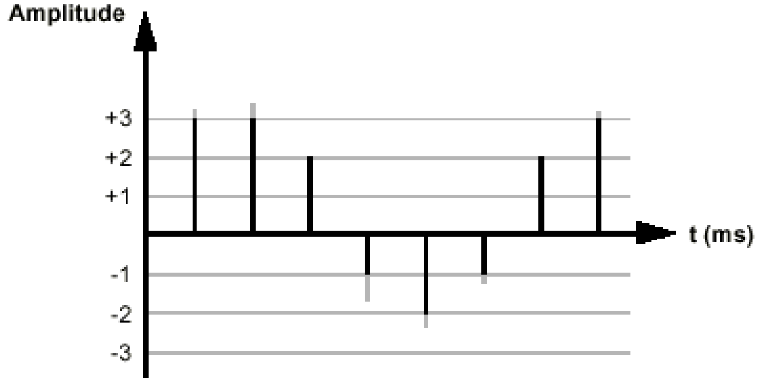

Figure 20: Samples with the corresponding quantized values [19]

Due to quantization, there are a limited number of numerical values to transmit, the equipment can be made less complex, and the risk of transmission errors is reduced. In telephony, 256 quantizing intervals are used. Consequently, there are 256 values to be transmitted.

<u>Quantization Noise</u>

Quantization also means that we forgo accuracy: the series of digits is not really the whole truth about the voice curve. Although each input sample is assigned a quantization interval that is closest to its amplitude height, but if an input sample is not assigned a quantization interval that matches its actual height, then an error is introduced into the PCM process. This deviation is called quantizing distortion or quantization error. This error acts as a noise which is equivalent to the random noise that impacts the signal-to-noise ratio (SNR) of a voice signal. SNR is a measure of signal strength relative to background noise. The ratio is usually measured in decibels (dB). If the incoming signal strength in microvolts is Vs, and the noise level, also in microvolts, is Vn, then the signal-to-noise ratio, S/N, in decibels is given by the formula S/N = 20 log10(Vs/Vn). SNR is measured in decibels (dB). The higher the SNR, the better the voice quality. Quantization noise reduces the SNR of a signal. Therefore, an increase in quantization noise degrades the quality of a voice signal. The figure below shows how quantization noise is generated.

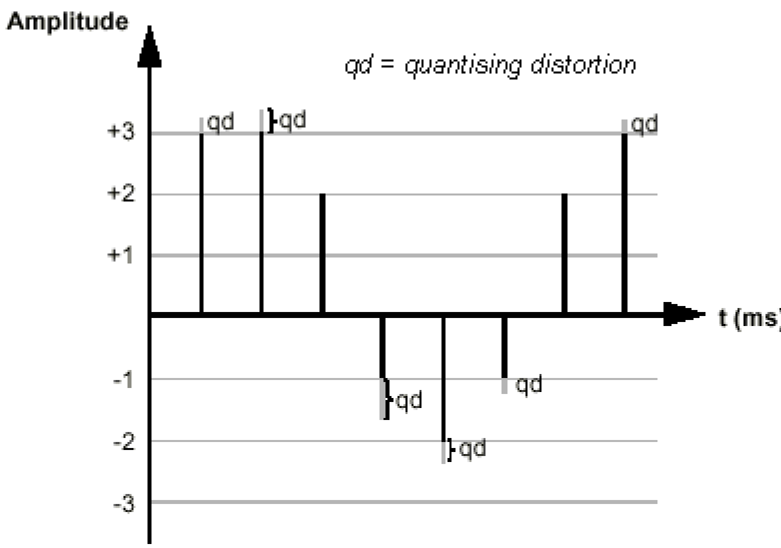

Figure 21: Quantization Error [19]

One way to reduce quantization noise is to increase the amount of quantization intervals. The difference between the input signal amplitude height and the quantization interval decreases as the quantization intervals are increased (increases in the intervals decrease the quantization noise). However, the amount of code words also need to be increased in proportion to the increase in quantization intervals. This process introduces additional problems that deal with the capacity of a PCM system to handle more code words.

SNR (including quantization noise) is the single most important factor that affects voice quality in uniform quantization. Uniform quantization uses equal quantization levels throughout the entire dynamic range of an input analog signal. Therefore, low signals have a small SNR (low-signal-level voice quality) and high signals have a large SNR (high-signal-level voice quality). Since most voice signals generated are of the low kind, having better voice quality at higher signal levels is a very inefficient way of digitizing voice signals. To improve voice quality at lower signal levels, uniform quantization (uniform PCM) is replaced by a nonuniform quantization process called companding.

3.2.5 Coding

Now it remains to give the 256 possible values a suitable layout for transmission. For this purpose, binary pulses, that is, pulses with only two levels are used. Each quantization interval is assigned a discrete value in the form of a binary code word. The standard word size used is eight bits ($2^8 = 256$).

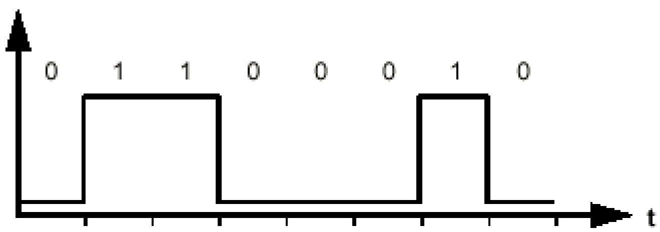

Figure 22: **Binary pulses** [19]

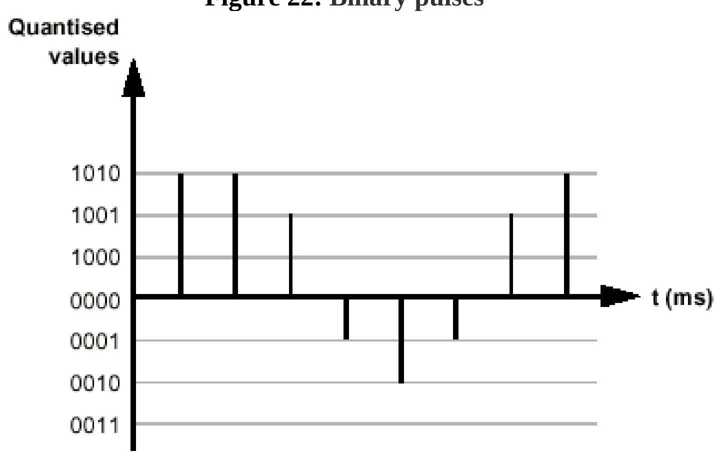

Figure 23: Quantized values with the corresponding binary code[19]

The result of this pulse code modulation process - the eight-bit binary code - is called a PCM word. One PCM word corresponds to one sample. If an input analog signal is sampled 8000 times per second and each sample is given a code word that is eight bits long, then the maximum transmission bit rate for Telephony systems using PCM is 8 x 8000 = 64,000 bits per second. The ITU-T calls this type of voice coding "64 kbit/s PCM".

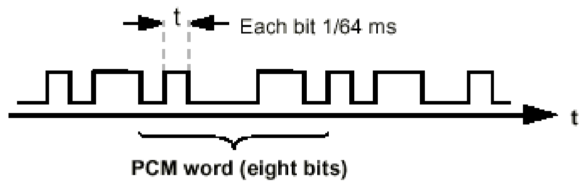

Figure 24: PCM word, eight bits[19]

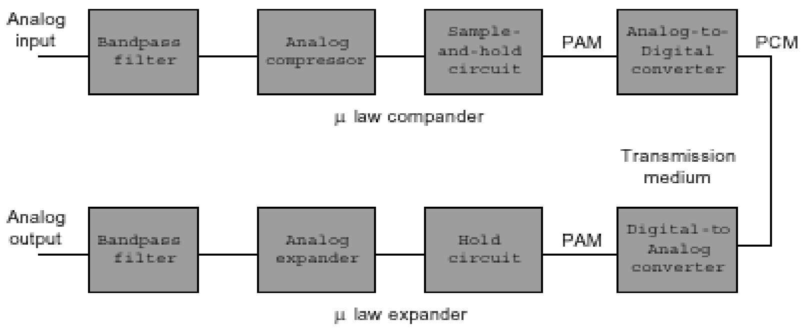

Figure 25: A PCM System[20]

3.3 DPCM

Differential pulse code modulation (DPCM) is a procedure of converting an analog into a digital signal in which an analog signal is sampled and then the difference between the actual sample value and its predicted value (predicted value is based on previous sample or samples) is quantized and then encoded forming a digital value. DPCM code words represent differences between samples unlike PCM where code words represented a sample value[21].

Basic concept of DPCM - coding a difference, is based on the fact that most source signals show significant correlation between successive samples so encoding uses redundancy in sample values. Since the difference between input samples is less than an entire input sample, the number of bits required for transmission is reduced. This allows for a reduction in the throughput required to transmit voice signals. Using DPCM can reduce the bit rate of voice transmission down to 48 kbps.

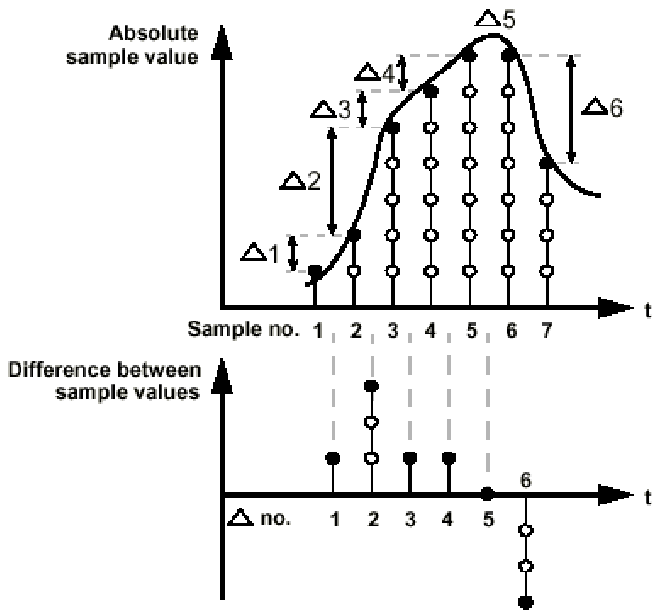

Figure 26: Differential PCM (DPCM)[19]

Realization of basic concept (described above) is based on a technique in which we have to predict current sample value based upon previous samples (or sample) and we have to encode the difference between actual value of sample and predicted value (the difference between samples can be interpreted as prediction error). Because it's necessary to predict sample value DPCM is form of predictive coding.

DPCM compression depends on the prediction technique, well-conducted prediction techniques lead to good compression rates, in other cases DPCM could mean expansion comparing to regular PCM encoding.

3.3.1 DPCM Encoder

The first part of DPCM works exactly like PCM. The input signal is sampled at a constant sampling frequency (twice the input frequency). Then these samples are modulated using the PAM process. At this point, the DPCM process takes over. The sampled input signal is stored in what is called a predictor. The predictor takes the stored sample signal and sends it through a differentiator. The differentiator compares the previous sample signal with the current sample signal and sends this difference to the quantizing and coding phase of PCM (this phase can be uniform quantizing or companding with A-law or µ-law). After quantizing and coding, the difference signal is transmitted to its final destination.

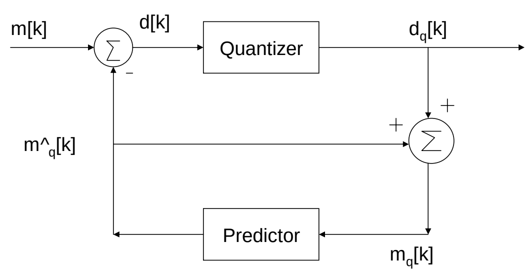

Figure 27: DPCM Encoder[22]

3.3.2 DPCM Decoder

At the receiving end of the network, everything is reversed. First the difference signal is dequantized. Then this difference signal is added to a sample signal stored in a predictor and sent to a low-pass filter that reconstructs the original input signal.

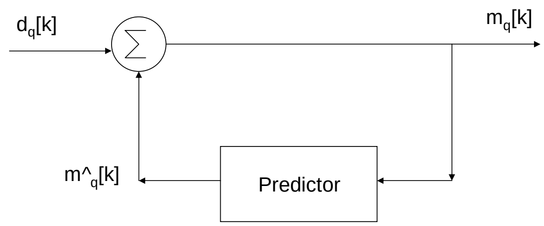

Figure 28: DPCM Decoder[22]

DPCM is a good way to reduce the bit rate for voice transmission. However, it causes some other problems that deal with voice quality. DPCM quantizes and encodes the difference between a previous sample input signal and a current sample input signal. DPCM quantizes the difference signal using uniform quantization. Uniform quantization generates an SNR that is small for small input sample signals and large for large input sample signals. Therefore, the voice quality is better at higher signals. This scenario is very inefficient, since most of the signals generated by the human voice are small. Voice quality needs to focus on small signals.

3.4 DELTA MODULATION

Sample correlation used in DPCM is further exploited in delta modulation (DM) by oversampling the baseband signal. This increases the correlation between adjacent samples, which results in a small prediction error that can be encoded using only one bit. Thus, DM is basically a 1-bit DPCM, that is, a DPCM that uses only two levels for quantization. It is a very simple and inexpensive method of analog to digital conversion.

In DM, analog input is approximated by staircase function which moves up or down by one quantization level at each sampling interval. The bit stream approximates derivative of analog signal (rather than amplitude). It generates 1 if function goes up and 0 otherwise.

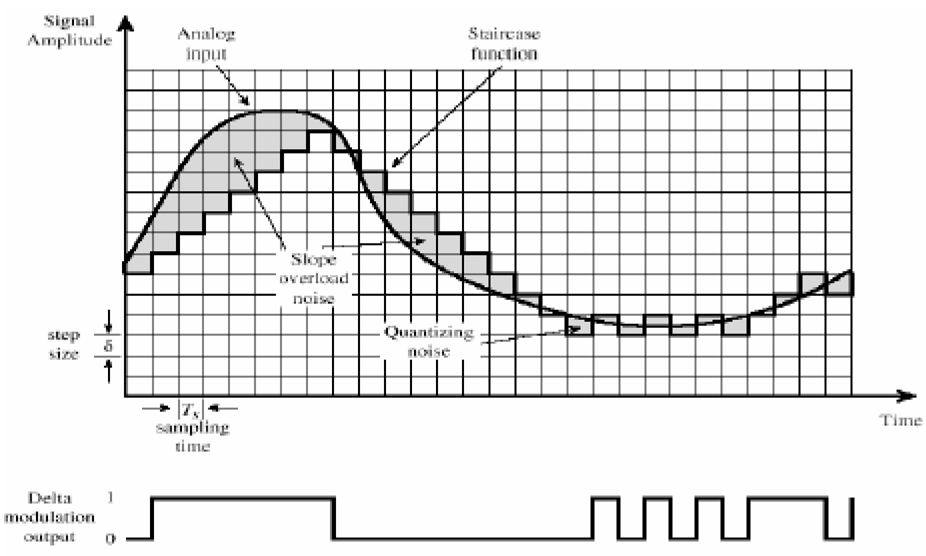

Figure 29: Example of Delta Modulation[20]

The step size in DM is fixed which may cause noise. When the slope of the analog signal is too steep, the step size proves to be too small to accurately follow the signal. This generates slope overload noise. On the other hand when the slope is very gradual the step size might prove to be too large to follow the signal. This generates a noise called granular noise. To overcome these problems, adaptive delta modulation is used.

3.5 ADAPTIVE DELTA MODULATION

To overcome the overload effects in DM, some type of signal compression is necessary. A suitable method appears to be the adaptation of the step value according to the level of the input signal derivative. For example, when the signal is falling rapidly, slope overload occurs. If we increase the step size during this period, the overload is avoided. On the other hand, if the slope is small, a reduction of step size reduces the threshold level as well as the granular noise. This method is called Adaptive Delta Modulation.

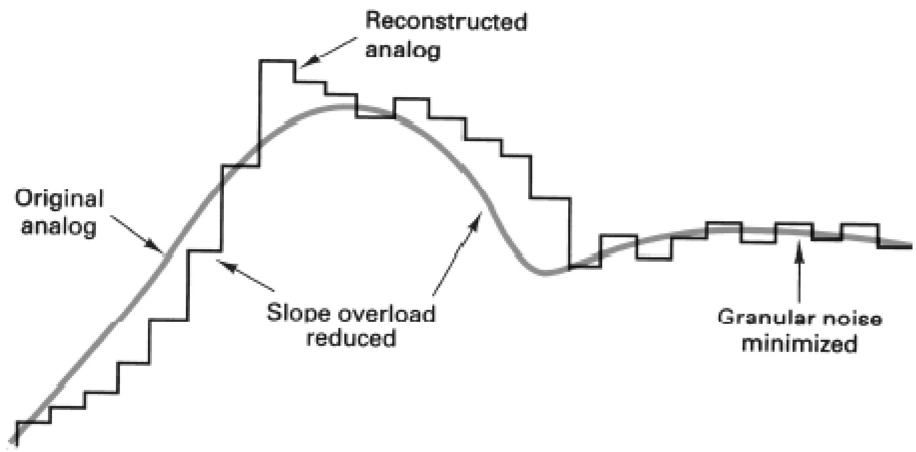

Figure 30: Adaptive Delta Modulation[20]

3.6 SUMMARY

The voice input in the computer is converted to digital pulses using the Differential Pulse Code Modulation, DPCM, because it is bandwidth efficient. In this process first the voice is sampled, difference is calculated between consecutive samples, this difference is quantized to various levels, then encoded to bits and sent to the receiver; unlike PCM where actual sample values go through this process instead of the differences. In delta modulation, this difference is further minimized to one bit. Quantization generates a noise called quantization noise because of not being assigned an appropriate level.

Chapter No. 4

CHANNEL CODING

4.1 INTRODUCTION

Channel coding is defined as a way of encoding data in a communications channel that adds patterns of redundancy into the transmission path in order to lower the error rate[23]. Channel coding is often used in digital communication systems to protect the digital information from noise and interference and reduce the number of bit errors. Channel coding is mostly accomplished by selectively introducing redundant bits into the transmitted information stream. These additional bits allow detection and correction of bit errors in the received data stream and provide more reliable information transmission. The cost of using channel coding to protect the information is a reduction in data rate or an expansion in bandwidth.

4.2 TYPES OF CHANNEL CODES

Convolutional coding and block coding are the two major forms of channel coding. Convolutional codes operate on serial data, one or a few bits at a time. Block codes operate on relatively large (typically, up to a couple of hundred bytes) message blocks[24].

4.2.1 Block Codes

Block codes are based rigorously on finite field arithmetic and abstract algebra. They can be used to either detect or correct errors. Block codes accept a block of k information bits and produce a block of n coded bits. By predetermined rules, n-k redundant bits are added to the k information bits to form the n coded bits. Commonly, these codes are referred to as (n,k) block codes. Some of the commonly used block codes are Hamming codes, Golay codes, BCH codes, and Reed Solomon codes. There are many ways to decode block codes and estimate the k information bits.

4.2.2 Convolutional Codes

In telecommunication, a convolutional code is a type of error-correction code in which

(a) each *m*-bit information symbol (each *m*-bit string) to be encoded is transformed into an *n*-bit symbol, where *m/n* is the code *rate* (*n* >= *m*) and

(b) the transformation is a function of the last *k* information symbols, where *k* is the constraint length of the code[25].

Convolutional codes are one of the most widely used channel codes in practical communication systems. These codes are developed with a separate strong mathematical structure and are primarily used for real time error correction. Convolutional codes convert the entire data stream into one single codeword. The encoded bits depend not only on the current k input bits but also on past input bits. The main decoding strategy for convolutional codes is based on the widely used Viterbi algorithm.

Convolutional encoding with Viterbi decoding is a technique that is particularly suited to a channel in which the transmitted signal is corrupted mainly by additive white gaussian noise (AWGN). AWGN can be thought as noise whose voltage distribution over time has characteristics that can be described using a Gaussian, or normal, statistical distribution, i.e. a bell curve. This voltage distribution has zero mean and a standard deviation that is a function of the signal-to-noise ratio (SNR) of the received signal[24]. If received signal level is fixed then if the SNR is high, the standard deviation of the noise is small, and vice-versa. In digital communications, SNR is usually measured in terms of E_b/N_0, which stands for energy per bit divided by the one-sided noise density.

Convolutional codes are usually described using two parameters: the code rate and the constraint length. The code rate, k/n, is expressed as a ratio of the number of bits into the convolutional encoder (k) to the number of channel symbols output by the convolutional encoder (n) in a given encoder cycle. The constraint length parameter, K, denotes the "length" of the convolutional encoder, i.e. how many k-bit stages are available to feed the combinatorial logic that produces the output symbols. Closely related to K is the parameter m, which indicates how many encoder cycles an input bit is retained and used for encoding after it first appears at the input to the convolutional encoder. The m parameter can be thought of as the memory length of the encoder.

4.2.3 Turbo Codes

Recently a new parallel-concatenated convolutional coding technique known as turbo coding has emerged. A turbo code is a type of channel coding that uses a convolutional code and a type of Viterbi decoder that outputs a continuous value rather than a 0 or 1[26]. It is a near channel capacity error correcting code.

Initial hardware encoder and decoder implementations of turbo coding have already appeared on the market. This technique achieves substantial improvements in performance over concatenated Viterbi and Reed-Solomon coding. A variant in which the codes are product codes has also been developed, along with hardware implementations.

This error correcting code is able to transmit information across the channel with arbitrary low (approaching zero) bit error rate. This code is a parallel concatenation of two component convolutional codes separated by a random interleaver. It has been shown that a turbo code can achieve performance within 1 dB of channel capacity. Random coding of long block lengths may also perform close to channel capacity, but this code is very hard to decode due to the lack of code structure. Without a doubt, the performance of a turbo code is partly due to the random interleaver used to give the turbo code a "random" appearance. However, one big advantage of a turbo code is that there is enough code structure (from the convolutional codes) to decode it efficiently.

There are two primary decoding strategies for turbo codes. They are based on a maximum a posteriori (MAP) algorithm and a soft output Viterbi algorithm (SOVA). Regardless of which algorithm is implemented, the turbo code decoder requires the use of two (same algorithm) component decoders that operate in an iterative manner.

4.3 CONVOLUTIONAL CODE

4.3.1 Encoder Structure

A convolutional code introduces redundant bits into the data stream through the use of linear shift registers as shown in this figure.

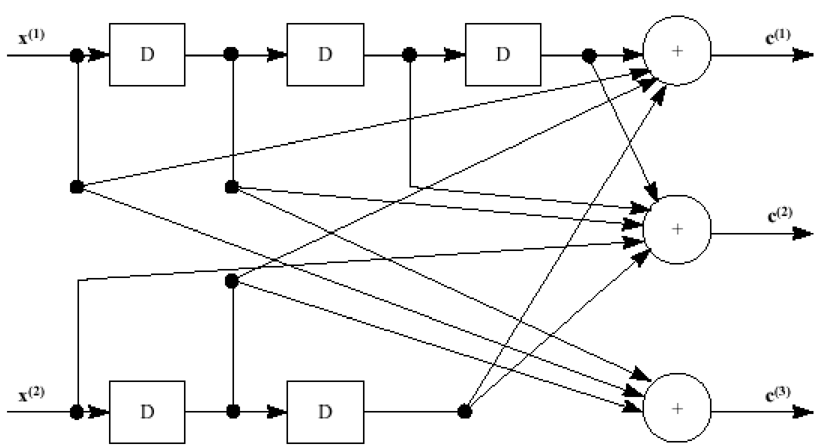

Figure 31: Example convolutional encoder where x (i) is an input information bit stream and c (i) is an output encoded bit stream[27]

The information bits are input into shift registers and the output encoded bits are obtained by modulo-2 addition of the input information bits and the contents of the shift registers. The connections to the modulo-2 adders were developed heuristically with no algebraic or combinatorial foundation. The code rate r for a convolutional code is defined as

$$r = k/n$$

where k is the number of parallel input information bits and n is the number of parallel output encoded bits at one time interval. The constraint length K for a convolutional code is defined as

$$K=m+1$$

where m is the maximum number of stages (memory size) in any shift register. The shift registers store the state information of the convolutional encoder and the constraint length relates the number of bits upon which the output depends. For the convolutional encoder shown above, the code rate r=2/3, the maximum memory size m=3, and the constraint length K=4. A convolutional code can become very complicated with various code rates and constraint lengths.

4.3.2 Encoder Representations

The encoder can be represented in several different but equivalent ways. They are

1. Generator Representation
2. Tree Diagram Representation
3. State Diagram Representation
4. Trellis Diagram Representation

<u>a. Generator representation</u> shows the hardware connection of the shift register taps to the modulo-2 adders. A generator vector represents the position of the taps for an output. A "1" represents a connection and a "0" represents no connection. For example, the two generator vectors for the encoder in the following figure are $\mathbf{g_1}$ = [111] and $\mathbf{g_2}$ = [101] where the subscripts 1 and 2 denote the corresponding output terminals.

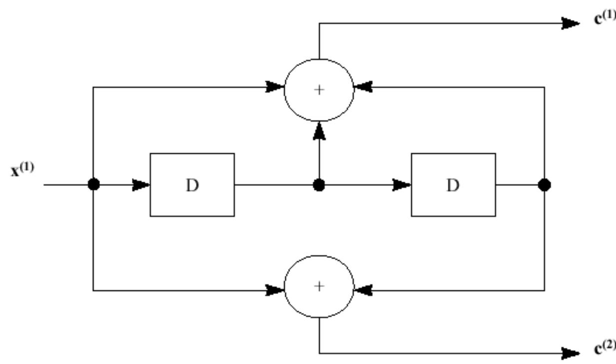

Figure 32: Generator Representation [27]

<u>b. Tree Diagram Representation</u> shows all possible information and encoded sequences for the convolutional encoder. The following figure shows the tree diagram for the encoder in last figure for four input bit intervals.

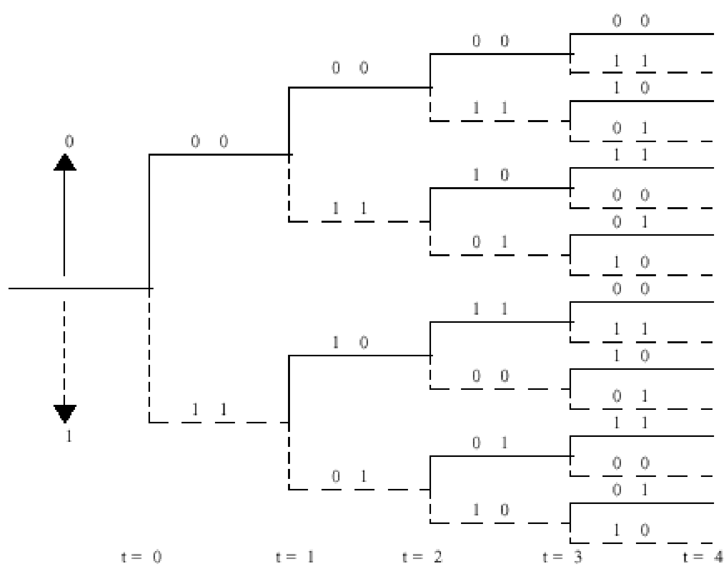

Figure 33: Tree Diagram [27]

In the tree diagram, a solid line represents input information bit 0 and a dashed line represents input information bit 1. The corresponding output encoded bits are shown on the branches of the tree. An input information sequence defines a specific path through the tree diagram from left to right.

c. State Diagram Representation shows the state information of a convolutional encoder. The state information of a convolutional encoder is stored in the shift registers. Figure shows the state diagram of the encoder in previous two figures.

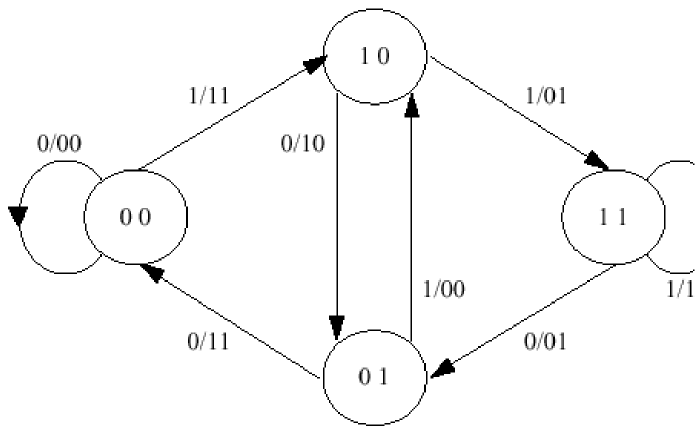

Figure 34: State Diagram [27]

In the state diagram, the state information of the encoder is shown in the circles. Each new input information bit causes a transition from one state to another. The path information between the states, denoted as x/c, represents input information bit x and output encoded bits c. It is customary to begin convolutional encoding from the all zero state.

d. Trellis Diagram Representation is basically a redrawing of the state diagram. It shows all possible state transitions at each time step. Frequently, a legend accompanies the trellis diagram to show the state transitions and the corresponding input and output bit mappings (x/c). This compact representation is very helpful for decoding convolutional codes as discussed later. The figure shows the trellis diagram for the same encoder as before.

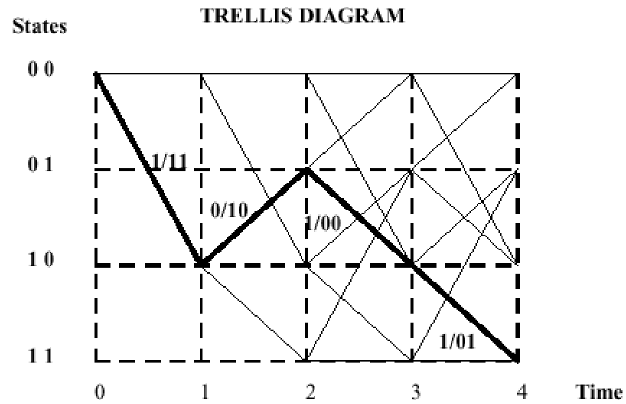

Figure 35: Trellis Diagram [27]

4.3.3 Hard-Decision and Soft-Decision Decoding

Hard-decision and soft-decision decoding refer to the type of quantization used on the received bits. Hard-decision decoding uses 1-bit quantization on the received channel values. Soft-decision decoding uses multi-bit quantization on the received channel values. For the ideal soft-decision decoding (infinite-bit quantization), the received channel values are directly used in the channel decoder. The following figure shows hard- and soft-decision decoding.

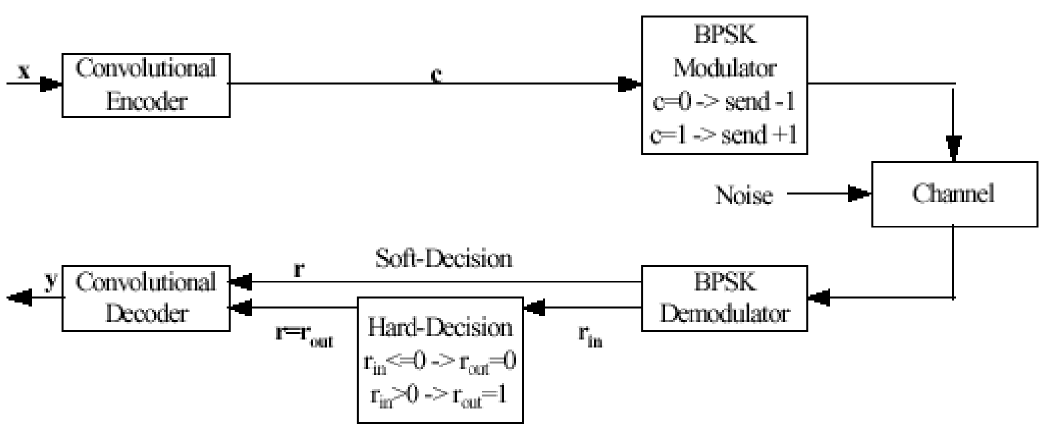

Figure 36: Hard and Soft Decision [27]

a. Hard-Decision Viterbi Algorithm

For a convolutional code, the input sequence **x** is “convoluted” to the encoded sequence **c.** Sequence **c** is transmitted across a noisy channel and the received sequence **r** is obtained. The Viterbi algorithm computes a maximum likelihood (ML) estimate on the estimated code sequence **y** from the received sequence **r** such that it maximizes the probability p(r|y) that sequence **r** is received conditioned on the estimated code sequence **y.** Sequence **y** must be one of the allowable code sequences and cannot be any arbitrary sequence. Following figure shows the described system structure.

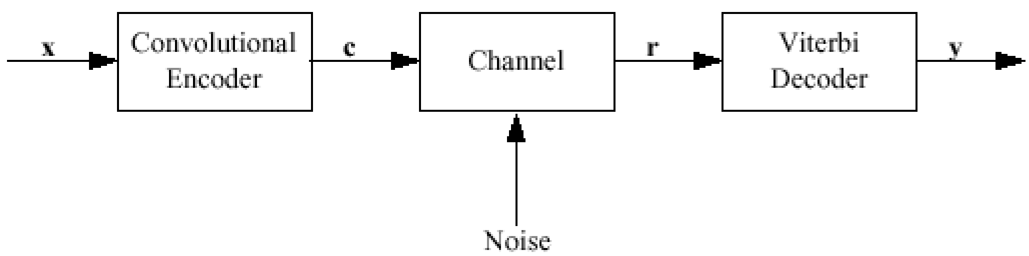

Figure 37: System Structure of Hard Decision Viterbi Decoder [27]

b. Soft-Decision Viterbi Algorithm

In soft-decision decoding, the receiver does not assign a zero or a one (single-bit quantization) to each received bit but uses multi-bit or infinite-bit quantized values. Ideally, the received sequence **r** is infinite-bit quantized and is used directly in the soft-decision Viterbi decoder. The soft-decision Viterbi algorithm is similar to its hard-decision algorithm except that squared Euclidean distance is used in the metric instead of Hamming distance.

4.3.4 Decoding Depth

The decoding depth is a window in time that makes a decision on the bits at the beginning of the window and accepts bits at the end of the window for metric computations. This scheme gives up the optimum maximum likelihood (ML) decoding at the expense of using less memory and smaller decoding delay. It has been experimentally found that if the decoding depth is 5 times greater than the constraint length K then the error introduced by the decoding depth is negligible.

4.4 TURBO CODE

4.4.1 Encoder

The fundamental turbo code encoder is built using two identical recursive systematic convolutional (RSC) codes with parallel concatenation. An RSC encoder is typically *r* = 1/2 and is termed a component encoder. The two component encoders are separated by an interleaver. Only one of the systematic outputs from the two component encoders is used, because the systematic output from the other component encoder is just a permuted version of the chosen systematic output. The following figure shows the fundamental turbo code encoder.

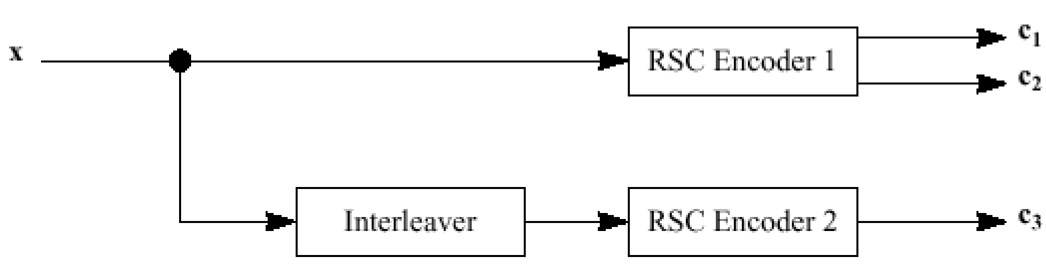

Figure 38: Fundamental Turbo Code Encoder[28]

The first RSC encoder outputs the systematic $\mathbf{c_1}$ and recursive convolutional $\mathbf{c_2}$ sequences while the second RSC encoder discards its systematic sequence and only outputs the recursive convolutional $\mathbf{c_3}$ sequence.

4.4.2 Recursive Systematic Convolutional (RSC) Encoder

The recursive systematic convolutional (RSC) encoder is obtained from the nonrecursive nonsystematic (conventional) convolutional encoder by feeding back one of its encoded outputs to its input. The following figure shows a conventional convolutional encoder.

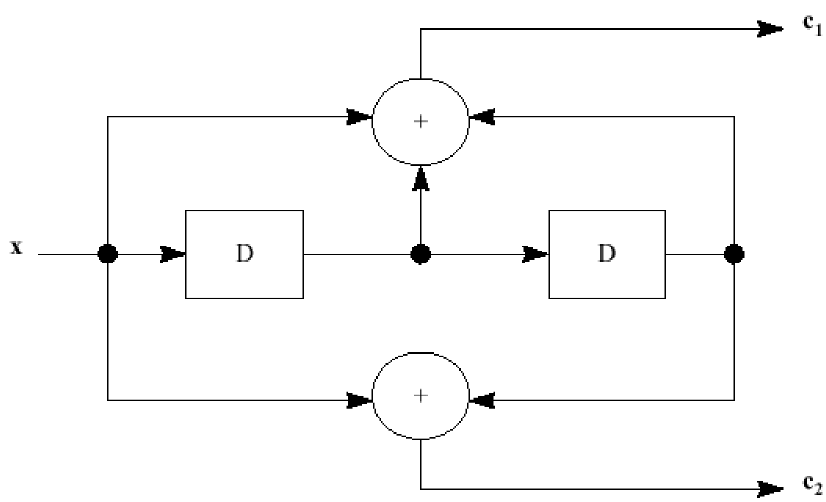

Figure 39: Conventional Convolutional Encoder with r=1/2 and K=3[28]

The conventional convolutional encoder is represented by the generator sequences $\mathbf{g_1}$ = [111] and $\mathbf{g_2}$ = [101] and can be equivalently represented in a more compact form as $\mathbf{G}$ = [$\mathbf{g_1}$, $\mathbf{g_2}$]. The RSC encoder of this conventional convolutional encoder is represented as G = [1, $\mathbf{g_2}$ / $\mathbf{g_1}$] where the first output (represented by $\mathbf{g_1}$) is fed back to the input. 1 denotes the systematic output, $\mathbf{g_2}$ denotes the feedforward output, and $\mathbf{g_1}$ is the feedback to the input of the RSC encoder. The resulting RSC encoder is shown below.

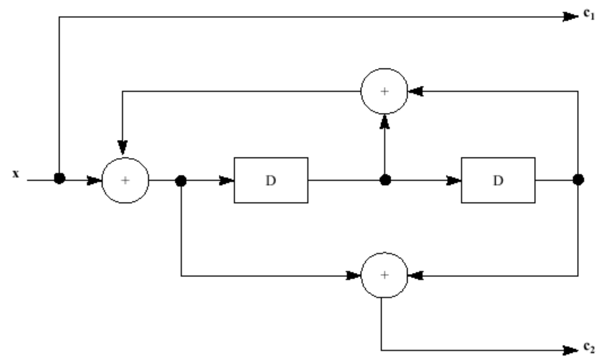

Figure 40: The RSC Encoder obtained from Conventional Convolutional Encoder with r=1/2 and K=3 [28]

4.4.3 Trellis Termination

For the conventional convolutional encoder, the trellis is terminated by inserting m = K-1 additional zero bits after the input sequence. These additional bits drive the conventional convolutional encoder to the all-zero state (trellis termination). However, this strategy is not possible for the RSC encoder due to the feedback. The additional termination bits for the RSC encoder depend on the state of the encoder and are very difficult to predict. Furthermore, even if the termination bits for one of the component encoders are found, the other component encoder may not be driven to the all-zero state with the same m termination bits due to the presence of the interleaver between the component encoders. Following figure shows a simple strategy that has been developed which overcomes this problem.

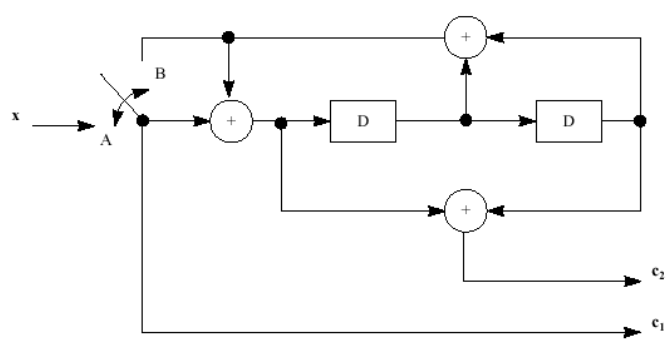

Figure 41: Trellis Termination Strategy for RSC Encoder [28]

For encoding the input sequence, the switch is turned on to position A and for terminating the trellis, the switch is turned on to position B.

4.4.4 Concatenation of Codes

A concatenated code is composed of two separate codes that are combined to form a larger code. There are two types of concatenation, namely serial and parallel concatenations.

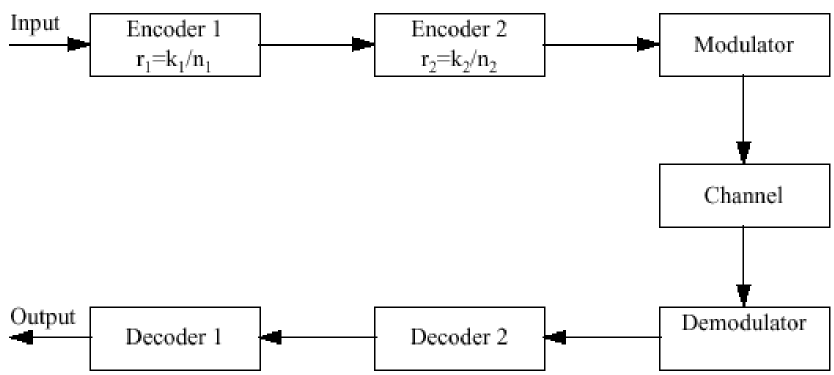

Figure 42: Serial Concatenated Code [28]

The total code rate for serial concatenation is

$$r_{tot} = k_1 k_2 / n_1 n_2$$

which is equal to the product of the two code rate.

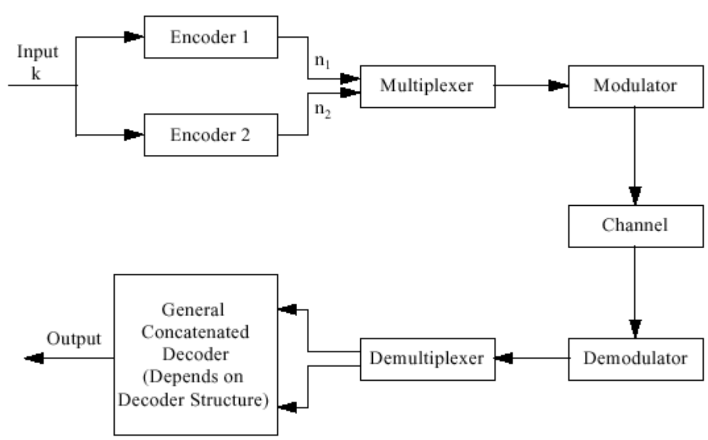

Figure 43: Parallel Concatenated code [28]

The total code rate for parallel concatenation is

$$r_{tot} = k / n_1 n_2$$

For both serial and parallel concatenation schemes, an interleaver is often used between the encoders to improve burst error correction capacity or to increase the randomness of the code.

Turbo codes use the parallel concatenated encoding scheme. However, the turbo code decoder is based on the serial concatenated decoding scheme.

4.4.5 Interleaver

For turbo codes, an interleaver is used between the two component encoders. The interleaver is used to provide randomness to the input sequences. Also, it is used to increase the weights of the codewords as shown in the figure below.

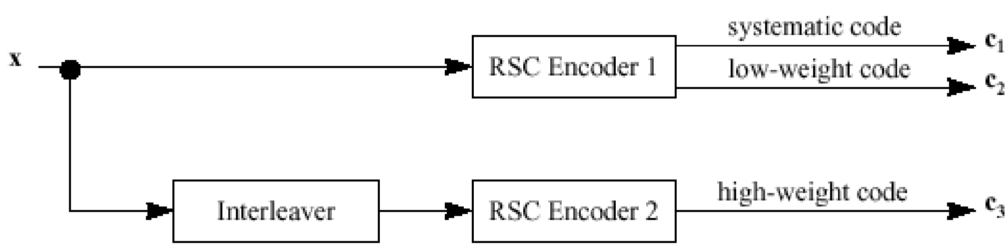

Figure 44: Interleaver increases the code rate for RSC encoder 2 as compared to RSC encoder 1 [28]

The input sequence **x** produces a low-weight recursive convolutional code sequence $\mathbf{c_2}$ for RSC Encoder 1. To avoid having RSC Encoder 2 produce another low-weight recursive output sequence, the interleaver permutes the input sequence **x** to obtain a different sequence that hopefully produces a high-weight recursive convolutional code sequence $\mathbf{c_3}$. Thus, the turbo code's code weight is moderate, combined from Encoder 1's low-weight code and Encoder 2's high-weight code.

4.4.6 Decoder

The turbo code decoder is based on a modified Viterbi algorithm that incorporates reliability values to improve decoding performance.

For concatenated (multistage) convolutional codes, there are two main drawbacks to conventional Viterbi decoders. First, the inner Viterbi decoder produces bursts of bit errors which degrades the performance of the outer Viterbi decoders. Second, the inner Viterbi decoder produces hard decision outputs which prohibits the outer Viterbi decoders from deriving the benefits of soft decisions. Both of these drawbacks can be reduced and the performance of the overall concatenated decoder can be significantly improved if the Viterbi decoders are able to produce reliability (soft-output) values. The reliability values are passed on to subsequent Viterbi decoders as a-priori information to improve decoding performance. This modified Viterbi decoder is referred to as the soft-output Viterbi algorithm (SOVA) decoder. Following figure shows a concatenated SOVA decoder.

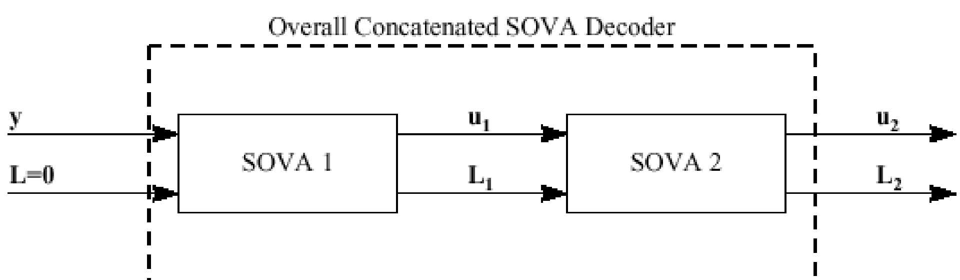

Figure 45: A concatenated SOVA decoder where y represents the received channel values, u represents the hard decision output values, and L represents the associated reliability values. [29]

4.4.7 SOVA Implementation

The SOVA decoder can be implemented in various ways. The straightforward implementation of the SOVA decoder may become computationally intensive for large constraint length K codes and long frame sizes because of the need to update all of the survivor paths. Because the update procedure is meaningful only for the maximum likelihood (ML) path, an implementation of the SOVA decoder that only performs the update procedure for the ML path is shown in the following figure.

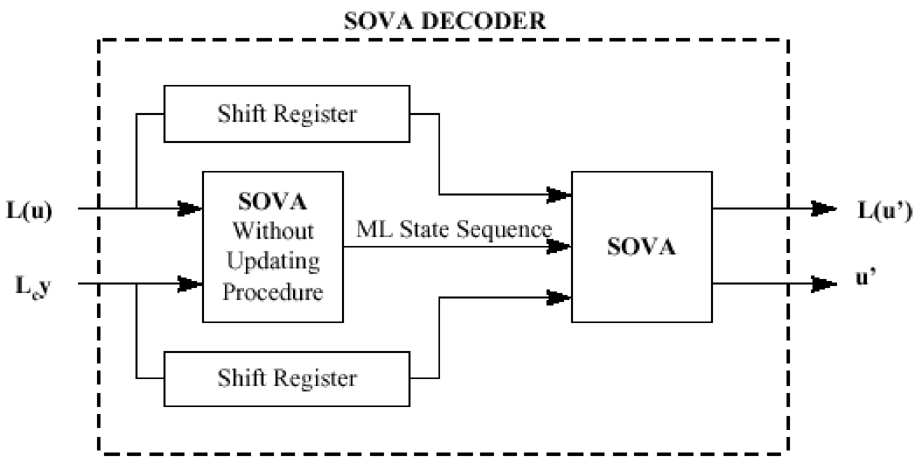

Figure 46: SOVA Decoder Implementation [29]

The SOVA decoder inputs are L(u) and L_cy, the a-priori values and the weighted received values respectively and outputs are u' and L(u'), the estimated bit decisions and its associated "soft" or L-values respectively. This implementation of the SOVA decoder is composed of two separate SOVA decoders. The first SOVA decoder computes the metrics for the ML path only and does not compute (suppresses) the reliability values. The shift registers are used to buffer the inputs while the first SOVA decoder is processing the ML path. The second SOVA decoder (with the knowledge of the ML path) recomputes the ML path and also calculates and updates the reliability values. As it can be seen, this implementation method reduces the complexity in the updating process. Instead of keeping track and updating 2^m survivor paths, only the ML path needs to be processed.

4.5 SUMMARY

The information bits have a high Bit Error Rate (BER) when they go wireless. These bits are thus error coded to cater for the channel noise. This type of channel coding is called as Forward Error Correction (FEC). Turbo codes have been used because of their high error tolerance. At the receiver end two Soft Output Viterbi decoders have been used to recover the encoded bits.

Chapter No. 5

SPREADING

5.1 INTRODUCTION

Bandwidth spreading or spreading of the data signal in a CDMA system is done by applying a code, independent of the data-signal. Each user is assigned a code which spreads its signal bandwidth in such a way that only the same code can recover it at the receiver end. This method has the property that the unwanted signals with different codes get spread even more by the process, making them like noise to the receiver. Originally for military use to avoid jamming, spread spectrum modulation is now used in personal communication systems for its superior performance in an interference dominated environment.

5.2 IMPORTANT DEFINITIONS

5.2.1 Correlation

Correlation is a measure of how related two entities are.[30]

• 1 means the second sequence matches the first sequence

• 0 means there is no relation at all between the two sequences

A high correlation means that there is a lot of resemblance between the two compared entities.

5.2.2 Auto-Correlation

It defines how much a function correlates with a time shifted version of itself, with respect to that time shift.[30]

5.2.3 Cross-Correlation

It defines how much a function correlates with other functions. It is the comparison between two sequences from different sources.[31] The codes used for spreading have low cross-correlation values and are unique to every user. This is the reason that a receiver which has knowledge about the code of the intended transmitter, is capable of selecting the desired signal.

5.2.4 Processing Gain (Spreading Factor)

The *Processing Gain(PG)* or sometimes called the *Spreading Factor(SF)* is defined as the ratio of the information bit duration over the chip duration:

$$PG = SF = Tb / Tc$$ [32]

where *Tb* represents the period of one data bit and *Tc* represents the period of one chip. The chip rate, 1/Tc, is often used to characterize a spread spectrum transmission system.

Hence, processing gain represents the number of chips contained in one data bit. Higher Processing Gain (PG) means more spreading. High PG also means that more codes can be allocated on the same frequency channel.

5.3 TYPES OF SPREADING SEQUENCES

For DS-CDMA, there are two types of spreading sequences:

i. PN sequences

ii. Orthogonal codes

5.3.1 PN Sequences

Pseudo-noise sequences are binary sequences which exhibit noise like randomness properties.[33] These sequences appear to be random; but are in fact perfectly deterministic. The sequence appears to be random in the sense that the binary values and groups of the same binary value occur in the sequence in the same proportion they would if the sequence were being generated based on a fair "coin tossing" experiment. In the experiment, each head could result in one binary value and a tail the other value. The PN sequence appears to have been generated from such an experiment. A software or hardware device designed to produce a PN sequence is called a PN generator.

PN spreading is the use of a PN sequence to distribute or spread the power of a signal over a bandwidth which is much greater than the bandwidth of the signal itself. PN despreading is the process of tasking a signal in its wide PN spread bandwidth and reconstituting it in its own much narrower bandwidth.

PN sequences can be used in at least two ways to spread the signal power over a wide bandwidth. In Frequency Hopping (FH) the center frequency of a narrowband signal is shifted pseudo randomly using the PN code. In Direct Sequence (DS) the signal power is spread over a wide bandwidth by in effect multiplying the narrow-band signal by a wideband PN sequence.

To be usable for direct-sequence spreading, a pn-code must meet the following constraints:

- The pn-codes are 2-leveled (bit-sequences).
- The codes must have a sharp (1-chip wide) autocorrelation peak to enable code-synchronization and to achieve equal spreading over the whole frequency-band.

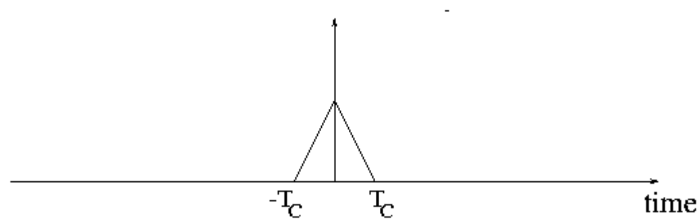

Figure 47: Auto-Correlation Function of a Random Binary Wave [32]

- The codes must have low cross-correlation values. The lower this cross-correlation, the more users can be allowed in the system. This holds for both full-code and partial-code overlap. The latter because in most situations there will not be a full-period correlation of two codes and it is more likely that codes will only correlate partially (due to random-access nature).
- To avoid a DC-component in the spread signal, the codes should be ``balanced'': the difference between ones and zeros in the code must be 1.

5.3.2 Orthogonal Codes

Two periodic signals of period *T* are orthogonal when their cross-correlation is null for a zero time shift.[32] Therefore, two orthogonal signals can be transmitted at the same time and will not interfere with each other. A set of orthogonal codes is used in CDMA where perfect synchronism can be achieved, as in the forward link. In orthogonal codes all pairwise cross correlations are zero. Or in other words, their inner product is zero. The inner product, in the case of codes with elements values +1 and -1, is the sum of all the terms we get by multiplying two codes element by element.

For example, (1, 1, 1, 1) and (1, 1, -1, -1) are orthogonal:

(1 * 1) + (1 * 1) + (1 * -1) + (1 * -1) = 0

5.4 WALSH CODES

An important set of orthogonal code is the *Walsh* set. Set of Walsh codes of length *n* consists of the *n* rows of an *n* x *n* Walsh-Hadamard matrix. Every row is orthogonal to every other row and to the logical not of every other row. Each row in the matrix is actually a separate code for each user.

5.4.1 Generation of Walsh Sequence

Walsh functions are generated using an iterative process of constructing a Hadamard matrix starting with $H_1 = [0]$. The Hadamard matrix is built by:

$$H_{2n} = \begin{pmatrix} H_n & H_n \\ H_n & \overline{H_n} \end{pmatrix}$$

For example, here are the

Walsh-Hadamard codes of length 2 and 4 respectively:

$$H_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} H_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$ [32]

From the corresponding matrix, the Walsh-Hadamard codewords are given by the rows. Usually the binary data is mapped to polar form so that real numbers arithmetic can be used when computing the correlations. So 0's are mapped to 1's and 1's are mapped to -1's.

5.4.2 Spreading of Information Bits

In spread spectrum, the data is modulated by the spreading signal (e.g. Walsh code) which uses more bandwidth than the data signal. Since multiplication in the time domain corresponds to convolution in the frequency domain, a narrow band signal multiplied by a wide band signal ends up being wide band.

Bits of the spreading signal are called *chips*. Spreading sequences of length n are generated for n users. Each incoming info bit is multiplied by a spreading sequence. Spreading factor determines how many chips are to be multiplied with one information bit.

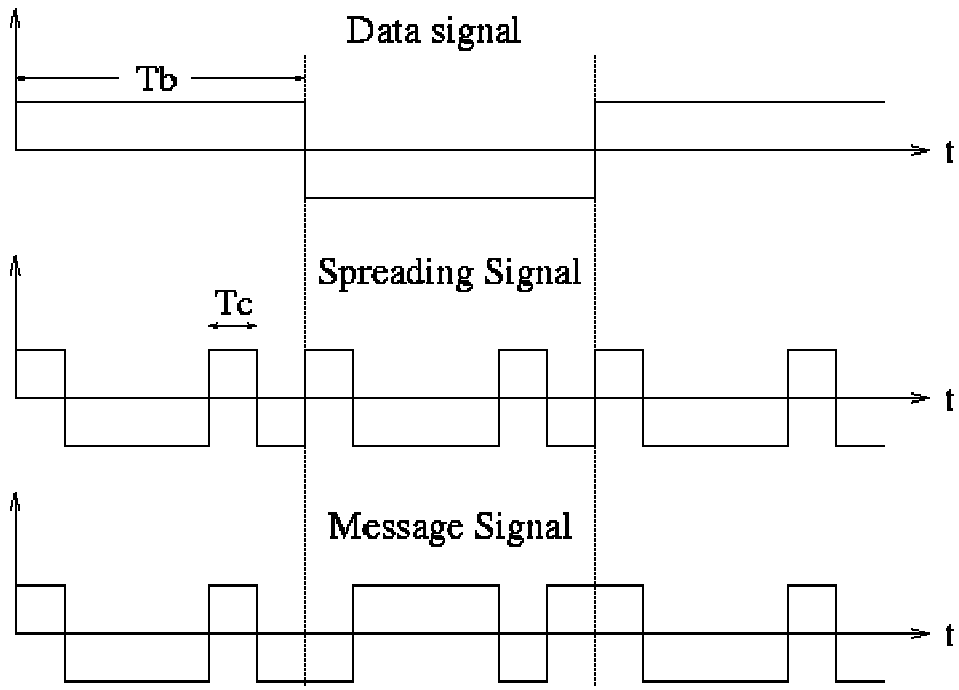

Figure 48: Spreading of a Data Signal [32]

5.4.3 Advantages of Walsh Codes

Since all pairwise cross correlations are zero, therefore, spreading data rate by an orthogonal code provides mutual orthogonality among all users in the same cell.

Walsh-Hadamard codes are also important because they form the basis for orthogonal codes with different spreading factors. This property becomes useful when signals with different Spreading Factors are wanted to share the same frequency channel. The codes that posses this property are called Orthogonal Variable Spreading Factor (OVSF) codes[32].

5.4.4 Drawbacks of Walsh Codes

Walsh-sequences have the advantage to be orthogonal, in this way one could get rid of any multi-access interference. There are however a number of drawbacks:

- The codes do not have a single, narrow autocorrelation peak. As a consequence code-synchronization becomes difficult and it is not possible for the receiver to detect the beginning of the codeword without an external synchronization scheme.
- Walsh-Hadamard codes do not have the best spreading behavior. They do not spread data as well as PN sequences do because their power spectral density is concentrated in a small number of discrete frequencies.
- Although the full-sequence cross-correlation is identically zero, this does not hold for partial-sequence cross-correlation function. The consequence is that the advantage of using orthogonal codes is lost when all users are not synchronized to a single time base.
- Orthogonality is also affected by channel properties like multi-path. In practical systems equalization is applied to recover the original signal.

5.5 SUMMARY

The error encoded bits are spread by using the modulo 2 addition with the unique code assigned to every particular user. A set of Orthogonal codes namely Walsh code has been used for the spreading purpose. Spreading factor, which determines the extent of spreading, has been kept dynamic.

CONCLUSION

A complete PC to PC CDMA link has been successfully established over voice. To make this link cost effective the P4 Intel Processors have been used, which were readily available. MATLAB has been used as the processing platform and the complete code has been implemented in it. Audio input is given to the PC using a microphone. The incoming voice is stored by the MATLAB as a WAV file which is then sampled at the rate of 8000 samples per second and binary coded using DPCM. 8 bits have been used to code each sample. Turbo code has been used for the Forward Error Correction, having constraint length 3. For the spreading of the information bits, Walsh code of the length 8 has been used, which is currently good for 8 users but can be increased if used with more powerful processors. Spreading factor has been kept dynamic which can be varied according to the requirement. The resulting spreaded information is stored as a binary file and transmitted through the serial cable to the receiving PC. If any transceiver module compatible with P4 PCs is connected to the serial port in place of the serial cable, a wireless link can be successfully established between the two PCs. An AWGN channel simulation has already been introduced at the receiver computer to give a wireless effect to the transmission and introduce the same type of errors as are added in a wireless channel. At the receiver end, after detecting the user, the incoming information is first despreaded, using the same spreading code as was used at the transmitting end. Then error decoding is done using Soft Output Viterbi Decoder (SOVA) having decoding depth 15. Then the original waveform is rebuilt from the binary information using the DPCM decoder. This waveform is output from the PC to the listener through the speakers.

FUTURE RECOMMENDATIONS

- Conversion of MATLAB codes to Visual C. It is going to drastically improve the processing speed.
- Real time implementation of the CDMA link in place of the record and store procedure.
- Use of wireless transceivers at the serial port instead of the serial cable, as the codes are highly capable of dealing with wireless transmission.
- Use of more powerful processors to accommodate maximum number of users.
- Introduction of rake receivers at the receiving end to cater for multi-path effects.

BIBLIOGRAPHY

1. http://www.cdg.org/technology/cdma_technology/a_ross/cdmarevolution.asp
2. http://www.wtec.org/loyola/wireless/02_04.htm
3. http://www.umtsworld.com/technology/cdmabasics.htm
4. http://www.mobile-phone-directory.org/Glossary/F/FDMA.html
5. http://www.tsp.ece.mcgill.ca/Telecom/Docs/cdma.html
6. http://www.cellphonecarriers.com/analog-fdma.html
7. http://www.cellphonecarriers.com/tdma-cell-phone.html
8. http://searchnetworking.techtarget.com/sDefinition/0,,sid7_gci214175,00.html
9. http://66.102.7.104/search?q=cache:hhuKGZnkM00J:etd.lib.fsu.edu/theses/available/etd-04092004143712/unrestricted/Ch_2MultipleAccess.pdf
10. http://cas.et.tudelft.nl/~glas/ssc/techn/techniques.html
11. http://www.see.ed.ac.uk/~eaa/cdma.html?http://oldeee.see.ed.ac.uk/~eaa/cdma.html
12. http://store.voipbooks.com/intocd.html
13. http://cas.et.tudelft.nl/~glas/stanford/node3.html
14. http://www.cdg.org/technology/index.asp
15. http://searchmobilecomputing.techtarget.com/gDefinition/0,294236,sid40_gci213842,00.html
16. http://www.bee.net/mhendry/vrml/library/cdma/cdma.htm
17. http://murray.newcastle.edu.au/users/staff/eemf/ELEC351/SProjects/ChanChng/Wave.html
18. http://www.cisco.com/warp/public/788/signalling/waveform_coding.html

19. http://www.ericsson.com/support/telecom/part-a/a-2-7.shtml

20. http://www.sis.pitt.edu/~dtipper/PM.pdf

21. http://einstein.informatik.uni-oldenburg.de/rechnernetze/dpcm.htm

22. B.P.Lathi, "Sampling and Pulse Code Modulation", *Modern Digital And Analog Communication Systems,* p. 281

23. http://www.answers.com/topic/channel-coding

24. http://home.netcom.com/~chip.f/viterbi/tutorial.html

25. http://www.answers.com/main/ntquery;jsessionid=2gbb3gqhdd0p1?method=4&dsid=1512&dekey=convolutional+code&gwp=8&curtab=1512_1&sbid=lc04ahttp://home.netcom.com/~chip.f/viterbi/tutorial.html

26. http://www.answers.com/main/ntquery;jsessionid=2gbb3gqhdd0p1?method=4&dsid=1512&dekey=turbo+code&gwp=8&curtab=1512_1&sbid=lc04a

27. http://scholar.lib.vt.edu/theses/available/etd-71897-15815/unrestricted/chap2.pdf

28. http://scholar.lib.vt.edu/theses/available/etd-71897-15815/unrestricted/chap3.pdf

29. http://scholar.lib.vt.edu/theses/available/etd-71897-15815/unrestricted/chap4.pdf

30. http://www.tsp.ece.mcgill.ca/Telecom/Docs/index.html

31. http://www.telecom.otago.ac.nz/tele303/notes/lec10_ss2.pdf

32. http://www.tsp.ece.mcgill.ca/Telecom/Docs/cdma.html

33. http://www-mobile.ecs.soton.ac.uk/bjc97r/pnseq-1.1/node1.html