4 Access and Radio Theory

Objectives

After this chapter the student will:
- be able to describe the general problems of radio transmission.
- be able to describe the solutions for these problems as specified in GSM.

4.1 Introduction

In the first part of this chapter the most common problems with radio transmission will be discussed. Some of them are general radio problems while others are only present in a digital or a time multiplexed system. In the latter part of the chapter the solutions to these problems, as applied in the GSM system, will be presented.

4.2 Problems with digital radio

Limited Spectrum

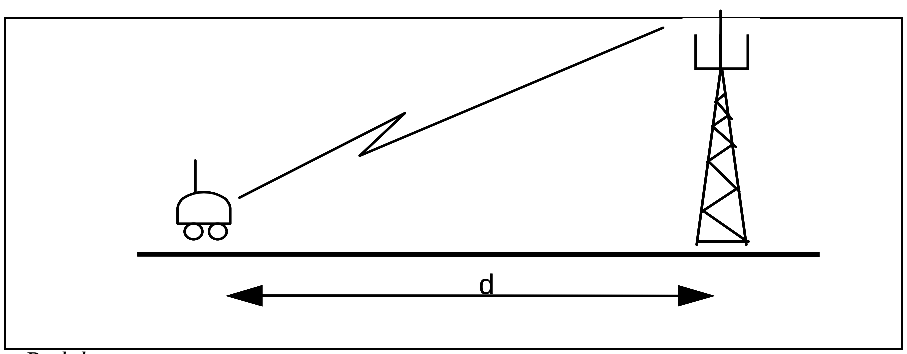
In the public switched telephone network, one time slot on a PCM-link is needed to carry the speech generated by one subscriber. One PCM time slot has a bit rate of 64 kbit/s.

One problem encountered with radio is that the available spectrum is limited. In GSM there is a separation of 200 kHz between the carrier frequencies and eight time slots per carrier. To keep the carrier within these 200 k Hz the bit rate per traffic channel has to be less than 13 kbit/s.

Thus, what normally takes 64 kbit/s to transfer in a traditional telecommunication network, must in GSM be handled with only 13kbit/s.

Path loss

Path loss or attenuation of the signal causes the received signal at the MS to get weaker the further away from the transmitting BTS antenna the MS is. Path loss can be a problem, making it difficult to get sufficient signal strength levels in a cell, but it is also the principle on which cellular systems are built. Due to the attenuation of the signal, frequencies can be reused over and over again all over the network.



Path loss

For a given frequency, path loss depends on the distance between the BTS and the MS antennae. One way to estimate this is to use the free space formula, where Lp stands for path loss, d for distance and λ for wavelength:

$$L_p = 20 \log (4 \pi d/\lambda)$$

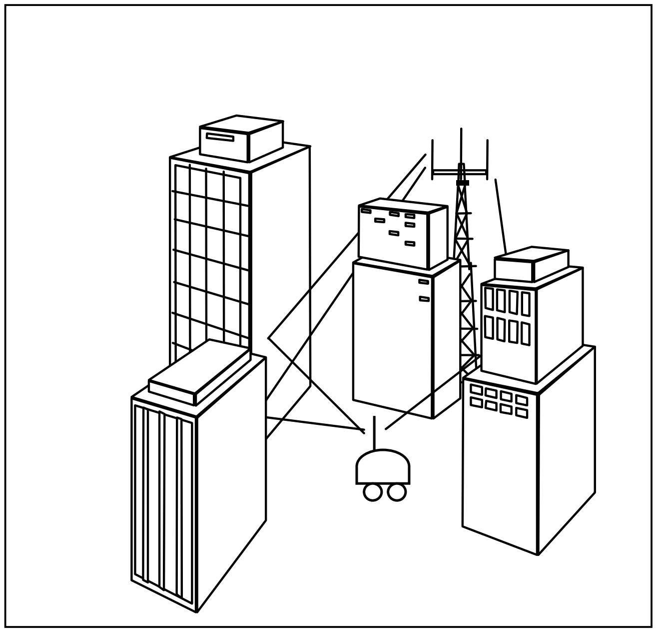
This formula assumes a line of sight condition between the transmitting and receiving antennae. It also assumes that there are no reflections interacting with the direct radio wave. Also, as indicated by the formula, the higher the frequency used, the higher the path loss.

Shadowing

If the radio path does not have free line of sight between transmitter and receiver, the obstacles will cause shadowing. Shadowing is also called ” log normal fading” or ”long term fading”. Since the MS is normally located in a low position, transmission will most likely be affected by shadowing objects, e.g. buildings, hills, the user or virtually anything in the radio path.

Multi-path fading

Another effect that might occur from not having line of sight between transmitting and receiving antennae is multi-path fading. In the absence of a direct radio wave, the reflected waves will be used instead. Normally we would receive not one, but several reflected radio waves and the resulting wave could be stronger, or weaker, than the individual waves. The reflected radio waves may be identical but are probably slightly delayed in time. If there is no phase difference between the waves, the resulting wave may have considerably better signal strength, but if the phase difference between the two signals is close to 180 degrees they may cancel each other out. This may cause very deep fading dips. The phenomenon is called multi-path or Rayleigh fading.



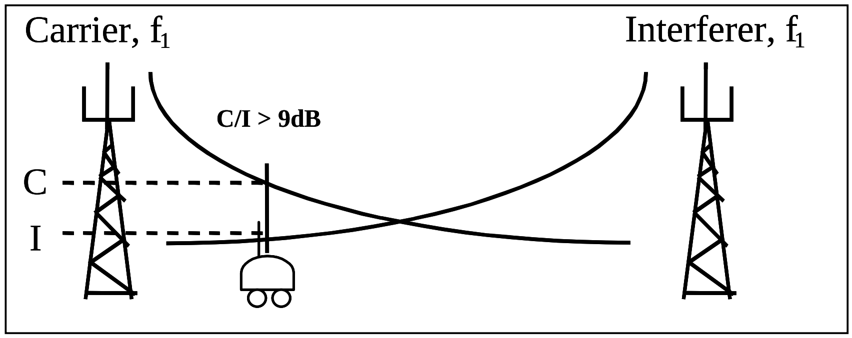
Multi-path fading

These fading dips may appear at a distance of half a wave length, which is approximately 16 centimetres for GSM 900 and 8 centimetres for GSM 1800. When the MS is moving fast, the fading dips will swiftly be passed through, not causing too much of a problem. The problems occur when the MS is moving slowly or standing still, e.g. stopping at a red light whilst in a fading dip.

This problem is present only when there is no direct radio wave and there are several reflectors nearby. This problem is particularly evident in cities and densely populated areas.

Carrier to Interference, C/I

Reusing an identical carrier frequency in different cells is limited by co-channel interference or C/I. Co-channel interference is the relation between the desired signal C and the undesired re-used signal I, both using the same carrier frequency.



Co-channel interference

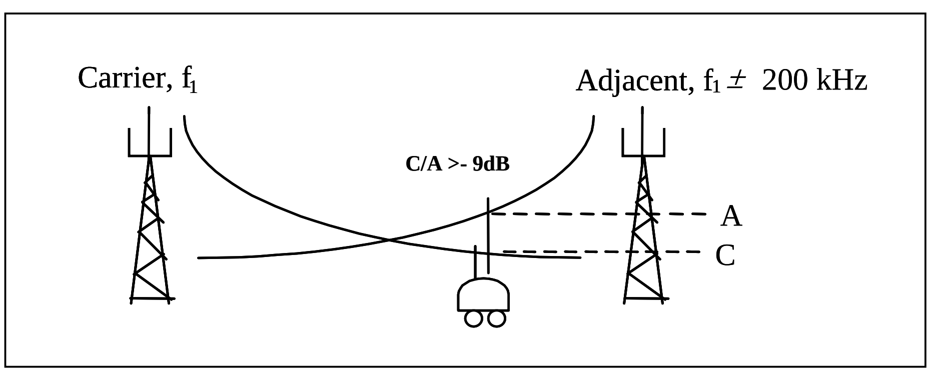
Suitable values for this ratio are settled by evaluation by a large group of listeners as to what is acceptable speech quality. The values given in the GSM recommendation values are:

C/I ≥ 9 dB

The margin for Rayleigh fading is already included in this value. But when the value 9 dB was decided, frequency hopping was assumed. In phase 1A in the GSM specification, frequency hopping is not used and in that case the value 12 dB should be used instead of 9 dB.

Carrier to Adjacent, C/A

As the filters, limiting each carrier to its domain of 200kHz, are not ideal, the carriers will somewhat affect each other. This means that some of the energy of the adjacent frequency will leak into the serving cell and cause interference. The relation between the desired signal C from the correct carrier and the undesired signal A from the carrier 200 kHz away is called adjacent channel interference or C/A.



Carrier to Adjacent, C/A

The limit for this ratio in GSM is:

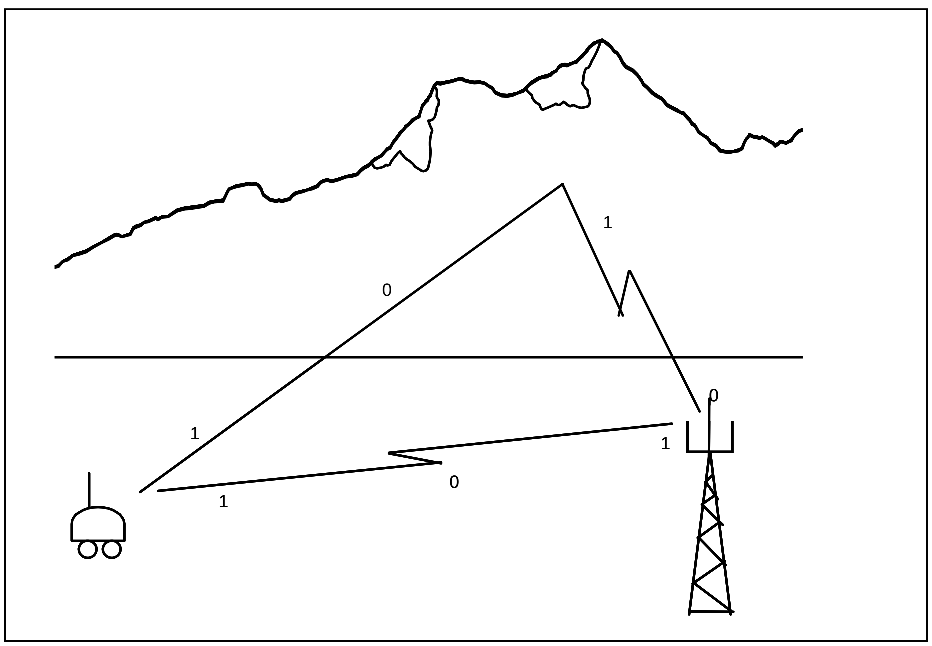
$C/A_1 \geq -9$ dB
$C/A_2 \geq -41$ dB

This means that we can allow the adjacent carrier frequency A_1 to be up to 9 dB stronger than our desired signal C. The index A_1 and A_2 denotes that we refer to the first and second adjacent channels, 200 kHz and 400 kHz away.

Time dispersion and Inter Symbol Interference

Time dispersion may be present when the MS is in an open area with a very large reflecting object, such as a mountain present. In this scene the MS may receive a direct radio wave from the base station, but also a fairly strong reflection from the reflector. This will cause interference with the direct wave.

In the case of digital modulation, this may cause the individual bits to overlap each other. The bit stream from the direct wave may arrive several bit times earlier than the identical reflected bits. This effect, caused by time dispersion, is called Inter Symbol Interference (ISI). Whereas multi-path fading is mainly a concern in large urban areas, ISI is a concern in rural areas.



If the delayed bits from the reflected radio wave interfere with the bits from the direct signal, ISI may occur.

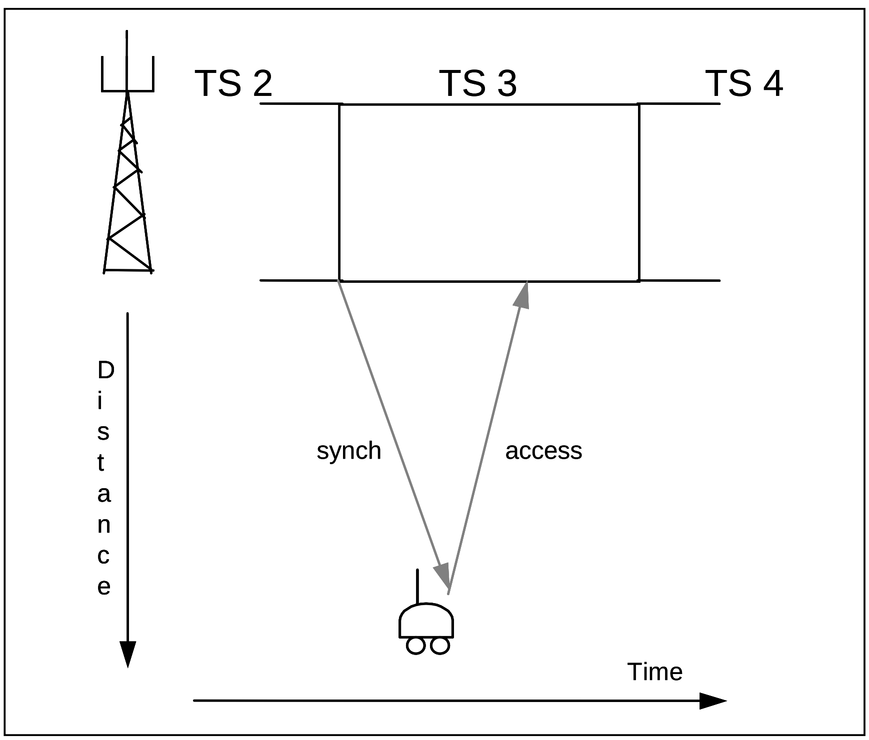
In cellular systems this phenomena is called carrier to reflection C/R. Problems like these might occur in mountain areas, hilly regions or places where a significant reflector is present.

The problem with ISI is on a bit level, i.e. a number of bits in a burst may interfere with a number of bits in the same burst. In contrast to ISI, there is also a problem on burst level; time alignment.

Time alignment

Time alignment is a pure TDMA concern. Each user is allowed to transmit in a specific time slot, and in that time slot only. If the signal, the burst, is sent too late it may interfere with the signals in the next time slot.

If the MS is close to the base station the burst will have no problems arriving within the allocated time slot. When the MS moves away from the BTS the burst will arrive later due to the greater distance that the burst has to travel before reaching the base station. What was meant for time slot three might arrive in time slot four, which is reserved for another user.



Time alignment

4.3 Solutions in GSM

Speech coding

The basic problem in digital radio transmission is to transform the analogue voice signals into digital ones. The difficulty is to maintain a good speech quality, while using as little as possible of the limited spectrum available, remember the problem of limited spectrum. There are a few different ways of coding the speech, of which waveform coding, vocoding and hybrid coding are mentioned below.

<u>Waveform coding:</u>
In a waveform coder three main steps are always involved; sampling, quantification and coding.

- *Sampling*

The analogue signal is sampled periodically. The more frequently the analogue signal is sampled, the better quality the wave form coder will have.

According to the sampling theorem:

$$f_s \geq 2 * f_{max}$$

To be able to reproduce an analogue signal without distortion, it must be sampled with a frequency at least twice as high as the highest frequency component in the analogue signal.

It is sufficient for good speech quality to be able to reproduce frequencies up to 3.4 kHz, so the sampling frequency should be at least 6.8 kHz (higher frequency components has quite low energy and can be omitted). In the PCM-system a sampling frequency of 8 kHz is used.

- *Quantification*

To describe the amplitude of the sampled signal we define a number of fixed levels. For each sample it must then be decided which of the fixed levels is closest to the sampled amplitude. The more levels the more accurate the speech coder becomes. The PCM-system uses 256 levels.

The PCM-system uses a method called A-law quantification, which gives the same resolution for low as well as for high amplitudes. This means that there are more levels defined for lower amplitudes than for high (to get the same of percentage of accuracy). GSM on the other hand uses uniform quantification, meaning that the different fixed amplitude levels are at equal distances from each other.

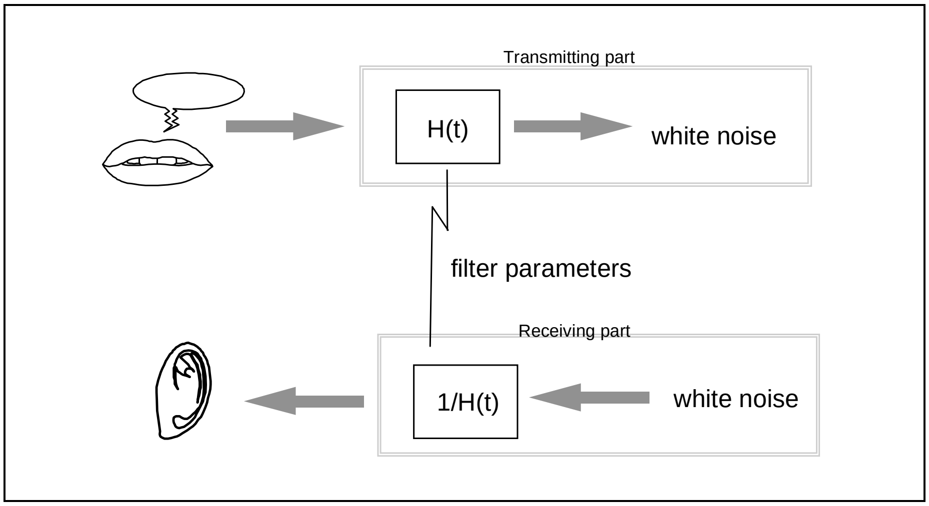
- *Coding*

The decided level of the sampled value must be coded into binary bits. In the PCM-system 8 bits are needed to code the 256 (2^8) levels. In GSM 13 bits are needed to get an equally good resolution with a uniform quantification.

The waveform coder can produce very good speech quality and is independent of the appearance of the signal. The only drawback is that it needs a high bit rate, normally between 16 and 64 kbit/s, to produce good speech quality.

Vocoding:

The vocoder has another approach to speech coding . It uses filters to reconstruct the speech in the decoder. Let us illustrate this by an example:



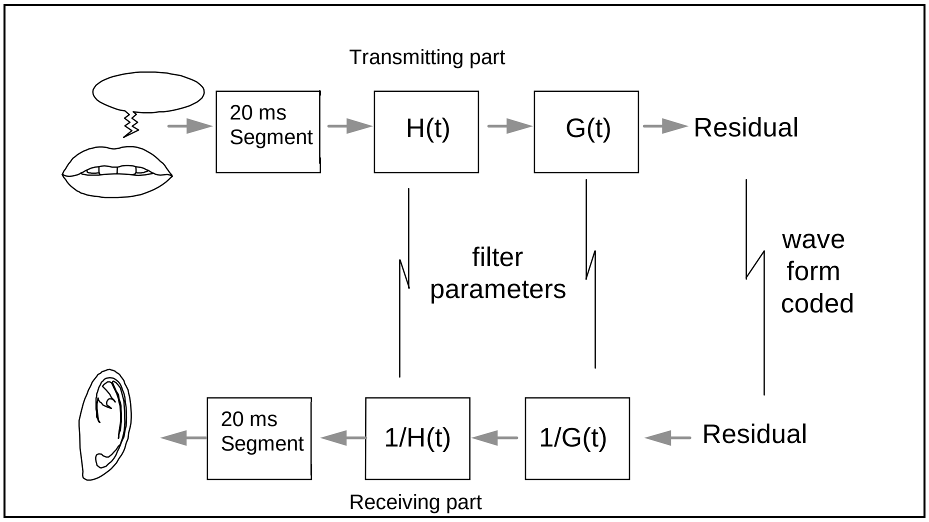
The principle of the vocoder

The vocoder tries to find the filter H(t) that at a certain moment neutralises the tone, or segment of speech, from the user. If this filter could be figured out and simultaneously built, the output of the filter would be white noise (in white noise all frequencies are present and there is no information). In that case, only the filter parameters have to be transferred. The receiving side could then construct an inversion filter 1/H(t) using the parameters. If the inversion filter is then fed with white noise the output would be the original coded signal.

This is the principle of the vocoder. Instead of sending the actual signals, it sends information about how to reconstruct the signal. As there are some problems involved in momentarily analysing and building filters, not to mention creating inversion filters, and as the speech segment never is constant, it is hard to achieve good speech quality using vocoders only. The advantage however is that the bit rate needed is low.

Hybrid coding:

A hybrid coder is a mix of different types of coders. In GSM a mix of vocoding and waveform coding is used. First the speech is sliced into 20 ms segments. Each segment is then run through a filter trying to remove as much information from the signal as possible. This first filter is constant during the whole 20 ms segment. Then the output from the first filter is run through another filter, which takes away the variation of the signal every 5 ms. Next the output, which unfortunately is not quite white noise, is run through a wave form coder, coding the residual product.



Principle of the GSM hybrid coder

The hybrid coder is combining the best of both coding methods; the vocoder taking care of most of the information at a low bit rate, while the residual product is taken care of by a wave form coder to ensure speech quality. The drawback of the hybrid coder is the added complexity, and also the higher sensitivity to variety in the signal appearance, i.e. male-female, language and other sounds than voice.

Channel coding

In an analogue network the loss of some information will only decrease the quality somewhat. The ear is able to correct the analogue signals that are slightly incorrect. In a digital network, however, the importance of each bit of information is crucial. The symbol "1" interpreted as a "0" gives a totally different piece of information.

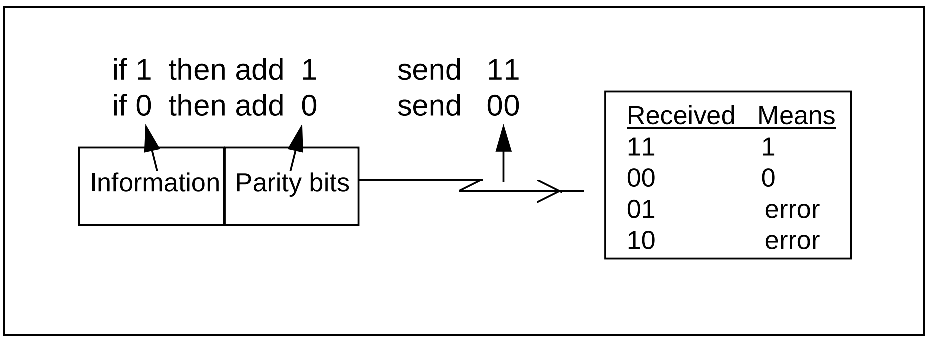
Remember the problem of multipath fading where you while passing one of these dips could loose several bits. With means of channel coding single bit errors can be detected and corrected.

The different methods of channel coding used in GSM are block coding and convolutional coding. The philosophy of both of them is basically the same; the bits are protected by adding a number of redundant bits to help detect or correct the errors.

Block coding:
When block coding is used, one or several check bits are added to the information block. The check bits only depend on the bits in that very block.

A simple form of block coding is using a parity bit. The parity bit could be set to zero if the 1´s in the block equal an even number. Otherwise the parity bit is set to one, so that the number of 1´s in the total block are always even.

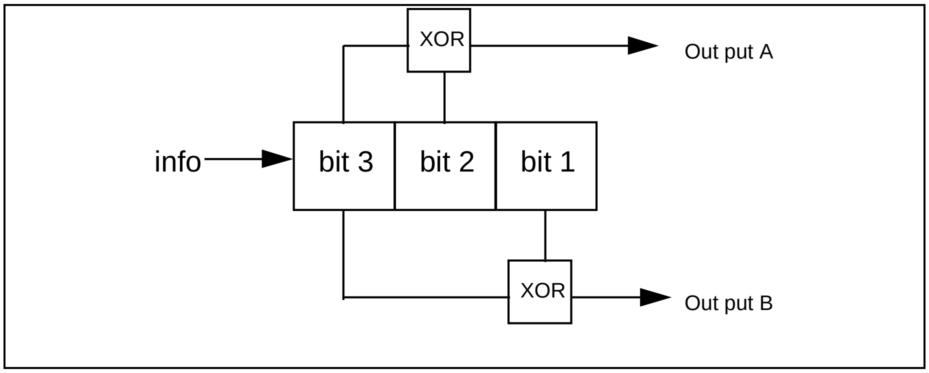
Block coding is mainly used for detecting errors. In the computer world block coding is often used together with a retransmission command, demanding the transmitting part to resend. This is not so useful when dealing with a real time application such as speech.



The principle of Block coding

Convolutional coding:
The convolutional coder consists of a shift register into which the information bits are shifted one by one. Doing logical operations on the positions of the bits in the register produces the coded information bits. This will make several coded bits dependent on one of the information symbols shifted into the coder. When all the information bits are shifted through the register we have produced the coded bits that will be sent.

Convolutional coding is not only good for detecting errors, but also for correcting them. The condition for being able to correct errors is that only few errors appear at a time, with a certain number of correct bits in between the incorrect ones.



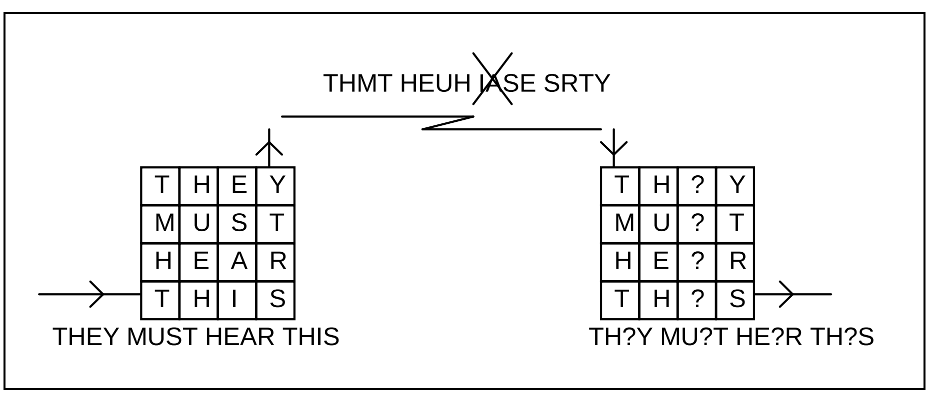
The principle of Convolutional coding

Interleaving

The error detection and correction methods mentioned, work best when the bits lost are spread out at a certain distance. In reality complete bursts with a large number of bits are lost.

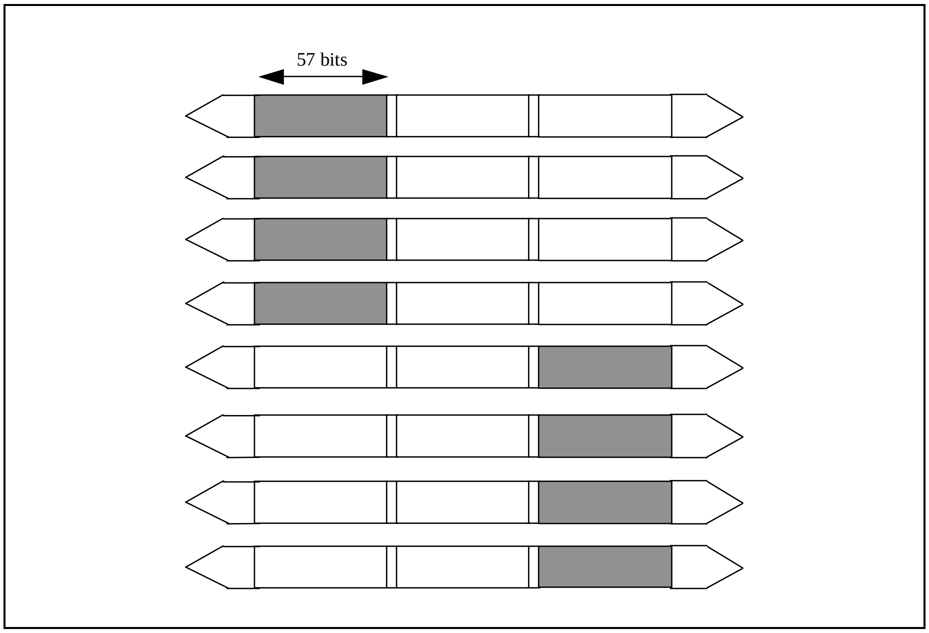
Interleaving is a method of spreading the potential losses, so that they can be taken care of by ”Channel Coding” thus minimising the harm of a lost burst. An analogy of this is, if the last 20 pages are torn out of an Agatha Christie novel, it will be more difficult to reconstruct the plot than if every 10th page, totalling 20 pages are lost.

As an example, let us assume that each message block contains four symbols, like in the picture below. Assume also that it is likely that we loose not only a part of the block, but the whole block. If we would use the principle of interleaving we would re-arrange the symbols so that, for example, all number one symbols are put together in one block, all the number two symbols in another, etc. If we now loose one whole block, we are actually only loosing every forth symbol, which is much easier to handle. If only parts of a block are lost, the chances of reconstructing the information improve dramatically.



If the information is regrouped or interleaved, the loss of symbols from a lost burst will be "shared" by several blocks.

In GSM the channel coder produces a total of 456 bits for every 20 ms segment of speech. These are interleaved in blocks of 57 bits as shown below.



In GSM the 456 bits from a 20 ms speech segment are interleaved over eight bursts

Diversity

Having gone through the negative effects, of multi-path fading earlier in this chapter, we now know that the fading dips will occur approximately every half wavelength. If the receiving antenna is located in a fading dip, there are only two ways of getting away from the dip, the receiving antenna must move away from the dip, or the fading dip must be moved away from the antenna. The latter is done by changing the frequency.

<u>Space diversity:</u>
By using two receiving antennae at the base station, chances are that at least one of them are free from multi-path fading problems. To achieve this a certain distance between the antennae is necessary, and 4-6 meters is recommended for GSM900. For GSM 1800 the recommended minimum distance is shorter.

There are different methods of combining the two received signals. Either the system can alternate between the two antennae, always using the antenna with the highest signal strength, or both the receiving antennae can be used all the time.

The use of space diversity reception will improve the C/I properties of the system as the problem with fading dips is reduced.

Frequency Hopping

Another effective way to fight multi-path fading is to change the frequency, thus changing the positions of the dips. When frequency hopping is applied in GSM, each consecutive burst will be transmitted (and of course received) at a different frequency. The frequencies used are changed either according to a cyclic pattern or a pseudo-random pattern (a "random" pattern decided in advance so that both parts in the transmission know what to expect).

Equaliser

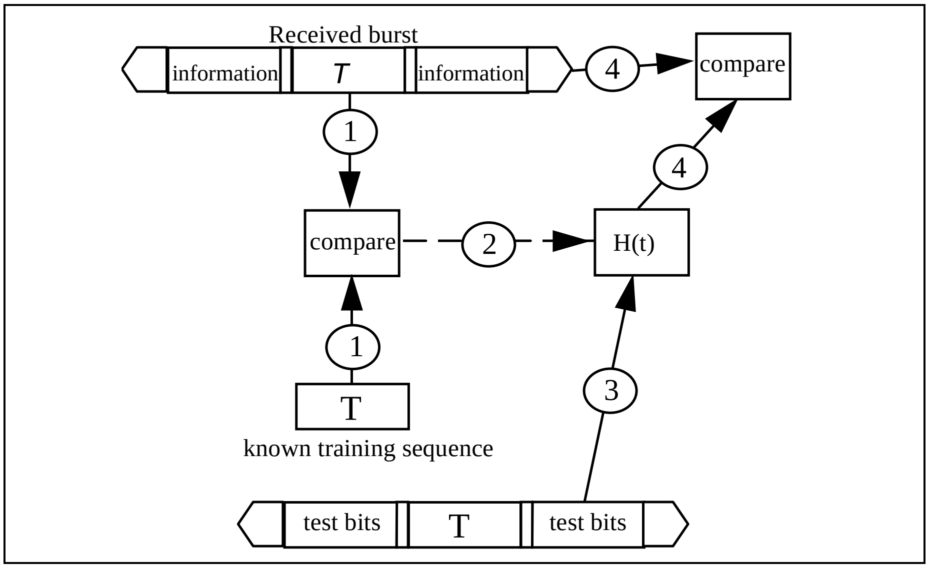
The equaliser will mainly address the problems of Inter Symbol Interference, described earlier. The problem occurs when the air interface affects the signal in some way that causes bit errors on the receiving side. If it was possible to get information about in what way the signals were affected, the system could take measures to correct the errors.

In a normal burst, used for traffic, there is a 26 bit training sequence in the middle of the burst. The bit pattern in this training sequence is known to the system. By analysing the received training sequence and compare it with what it should look like, the system will know how the air interface have affected the training sequence and also being able to guess how the other parts of a burst has been changed.

TB 3	Encrypted bits 57	1	Training sequence 26	1	Encrypted bits 57	TB 3

Training sequence in the middle of normal burst

By analysing the received training sequence a channel model can be built. This channel model is like a filter affecting the transmitted bits in the same manner that the air interface is affecting them. By running different bit patterns through the channel model and comparing the resulting signal with what was actually received, the system can reach a conclusion as to what was actually sent



Equaliser

According to the GSM-specification, the system must be able to cope with signals with a delay of 15 µs. This corresponds to a path distance of 4.5 km ($15\ \mu s * 3 * 10^8$ m/s = 4.5 km), meaning that time dispersion caused by a reflector 2.25 km or less away has to be handled by the system.

Timing advance and access burst

We earlier mentioned this timing problem, i.e. that each user must keep his bursts within the designated time slot. This could be a problem if the user is far away, and his signalling will be delayed as it travels trough the air.

To cope with this, the system will repeatedly send ”timing advance” orders to the MS. The system will simply tell the MS how many bit times earlier, or later, to send its bursts. These decisions are based on an analysis of how the bursts are received in the base station.

One problem remains however, the first access. When the MS is in idle mode, there is no real connection between the MS and the base station. In other words there is no way to judge how far the MS is from the BTS. When the MS makes the first access to the system, maybe to initiate a call, the access burst may come in too late.

Therefore the burst used for access is as short as possible to allow the burst to come in late without interfering with the next burst. Every burst is slightly shorter than the time slot in which it must fit. The normal burst, used for traffic, has 8.25 bit times as a buffer, while the access burst has 68.25 bit times as a buffer.

TB 8	Information	GuardPeriod 68.25

The access burst with an extended guard period

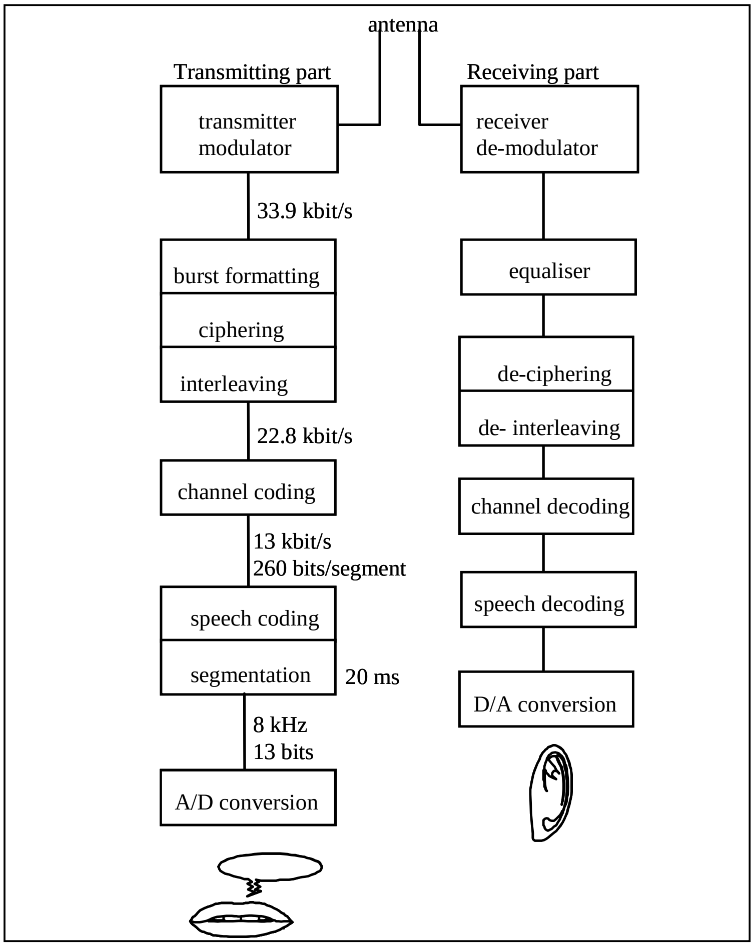
The MS tries to synchronise to the base station but will actually hear the synchronisation delayed because of the distance. Then the MS will send an access burst, based on that delayed timing, this burst will be delayed even more on the way to the base station. To compensate for this delay the MS will be informed by the system to transmit its bursts 0 to 63 bit times in advance.

63 bit times corresponds to 63 bits $*$ 3.7 μs / bit $*$ $3 * 10^8$ m / s = 70 km (3.7 μs is the time it takes to send one bit and $3 * 10^8$ m / s is the speed of light and also of the radio wave). In other words, the maximum distance an MS can be from a base station is 70 / 2 = 35 km, (division by two since the MS has to compensate for the delay on both the downlink as well as for the uplink).This is also the maximum radius of any cell.

4.4 Signal path through the network

Let us summarise the way the signal travels through the system. In the transmitting part of the MS, the speech is digitised and sliced into segments of 20 ms. Each segment is then run through the speech coder resulting in a bit rate of 13 kbit/s. The next step is channel coding to add redundant bits for error detection and correction. Now the bit rate is 22.8 kbit/s per traffic channel. Interleaving and ciphering is done later, and finally the 20 ms segment is formatted into 8 half burst. Here 33.9 kbit/s is needed. To be able to send information for 8 users (8 TS on a frequency) each with a rate of 33.9 k bit/s, the total bit rate per frequency will be (8x33.9) 270 kbit/s.

In the receiving part the bursts are received and a channel model created in the equaliser. After all 8 half bursts have been received and deciphered, they are reassembled into the 456 bit block again. The sequence is then decoded to detect and correct errors. Finally the bit stream is speech decoded and transformed into analogue speech signals.



Signal path through the MS