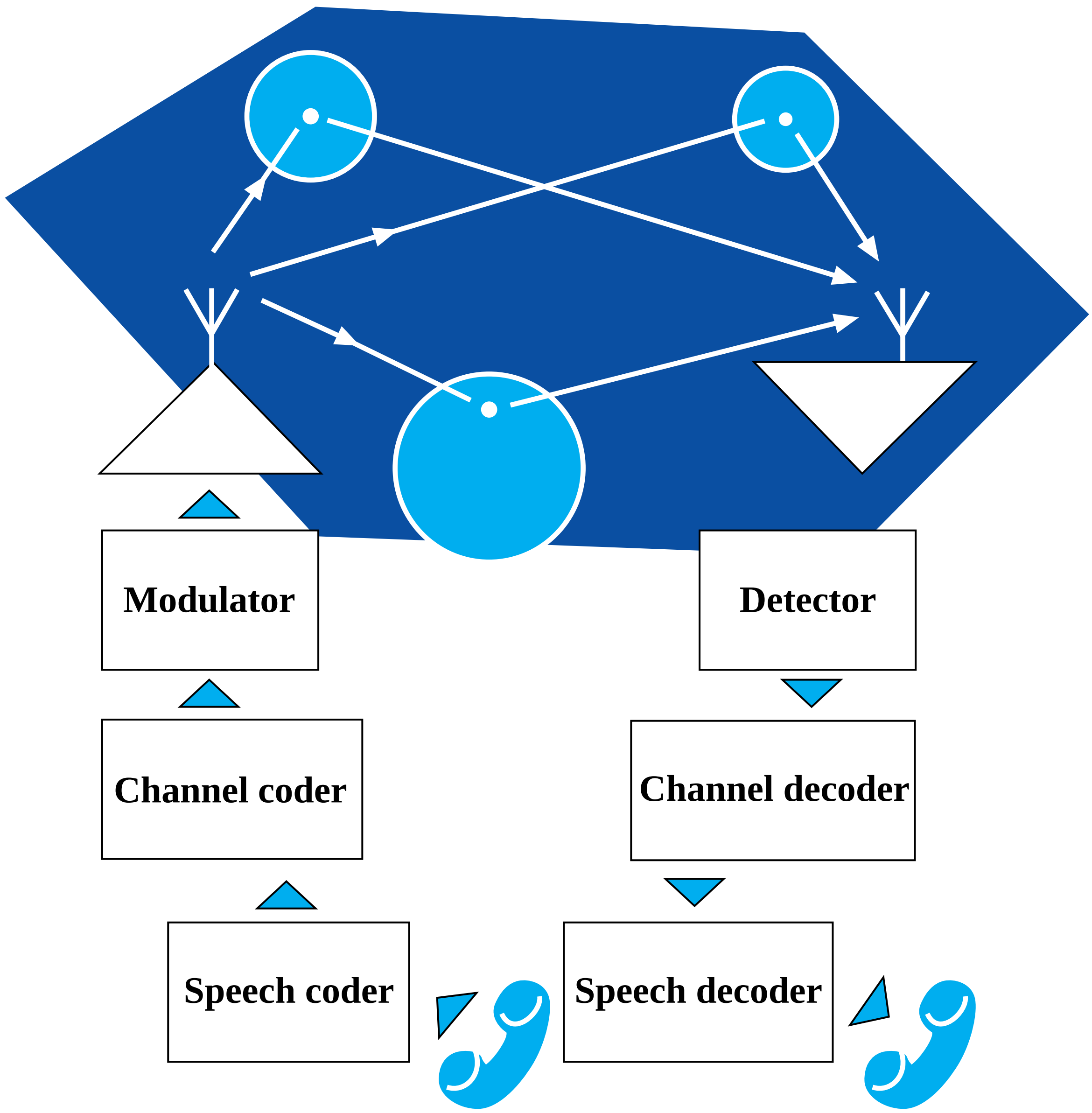
Radio School

DM1 Digital Mobile Telephony



R C U R

Core Unit Radio Systems and Technology

ERICSSON

DM1 GSM
Index

Digital Mobile Telephony
DM1 GSM (Global System for Mobile Communications)

Contents **Page**

1. The background to GSM

1.1 System specification

The initiative for a digital mobile telephone system came from the Scandinavian Telecommunication Administrations, who 1981 submitted a joint proposal to the CEPT for the specification of a pan-European mobile telephone system, conceivably to be based on digital transmission. The reason for proposing serious consideration of a digital transmission system was based on the findings of studies conducted by a Scandinavian working group. In 1982, the CEPT appointed a GSM group (Groupe Special Mobile), whose members consisted of representatives from a number of countries in western Europe, to investigate the idea.

Following system studies coordinated by the GSM group, a decision was taken in 1985 to draw up a specification for a digital system. The general criteria stipulated were that the new system should provide at least the same speech quality and spectrum efficiency as the existing analogue mobile telephone systems. Another requirement was that the estimated cost of the fully developed system when in mass production should be lower than that of the existing analogue ones. In addition, the system must be able to interface with the ISDN on the public telephone network, even if some services requiring a wide bandwidth might not be available owing to the frequency shortage.

At that time, simulations and experiments of digital-speech transmission systems based on FDMA had progressed far enough to predict with considerable certainty that a new system based on digital transmission would be able to offer a higher performance than existing analogue systems. However, it seemed likely that further development work could result in alternative forms of multiple access to FDMA with improved system performance. Due to technical uncertainties it was not yet possible to recommend any other multiple-access arrangement. The main unknown factor was if it would be possible to suppress strong intersymbol interference caused by time dispersion in wideband radio transmission. The GSM group therefore decided that an evaluation should be made of systems based on other types of multiple access.

Nine R & D groups in western Europe set up test systems, to be evaluated in Paris in autumn 1986 by means of laboratory evaluations, employing fading simulators, and field tests. It was very much on the basis of these comparative tests that the GSM group recommended in spring 1987 that a joint pan-European mobile telephone system should be developed based on digital transmission and ”narrowband TDMA”. The system would be called the GSM. This was followed by a Memorandum of Understanding signed by 13 countries, under which they agreed to introduce the GSM by 1 July 1991.

Key features of the outlines specification of 1987 were TDMA with 8 time slots in a time frame of 4.6 ms; RPE-LPC speech coders with a data rate of 13 kb/s; convolutional coding for error correction; and GMSK modulation with 200-kHz channel spacing.

A comprehensive specification drawn up by a consolidated GSM group (the permanent nucleus) was ready by the end of 1988. The extensive documentation covered not only the different radio subsystems but also the network services to be offered and interfaces to the fixed network. However, a great deal of work still remained on the fine details of the design, and this was made the responsibility of the ETSI. The work on developing the GSM as a commercial product proved to require considerably more resources than had been foreseen. In consequence, the project overran the original time plan by about a year. The first large scale introduction of GSM was in Germany in 1992-93, where the capacity of the existing analogue mobile telephone system had inadequate traffic capacity and high costs for subscribers.

The growth in the number of subscribers during the same period was slower in the Nordic countries. The main explanation for this was that the NMT network had still not reached its capacity limit, and many mobile telephone subscribers held the opinion that the service offered by NMT, with wide coverage in Scandinavia, was adequate and relatively low priced. However, a sharp upturn in the number of GSM subscribers came in the beginning of 1994. In middle 95 the Swedish frequency administration authority decided that part of the frequency band for NMT 900 should be used for GSM.

When the system was introduced, its full name was changed to Global System for Mobile Communications, which meant that the abreviation GSM could still be used.

1.2 System technology

FDMA

Development of the fundamental technology for digital mobile telephone systems started in Sweden at the beginning of the 1980s, and some years later in other western European countries. Initially, the Swedish studies focused on the simplest system configuration based on FDMA (see Fig. 1.1). The principal subsystems were speech coding, channel coding and modems.

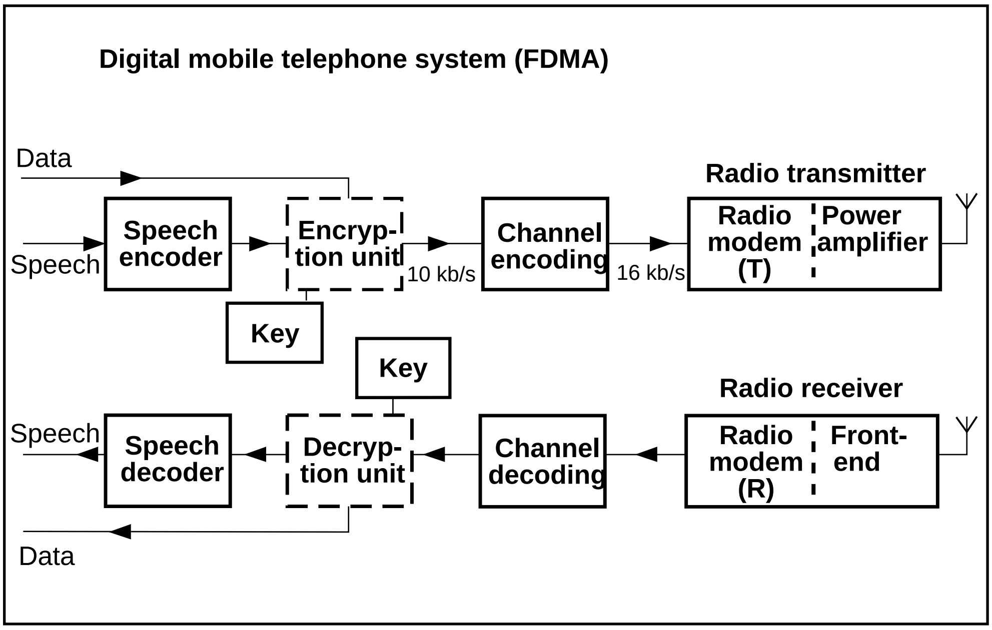


Fig. 1.1

The main finding of the work was that digital radio transmission eventually could provide better speech quality and higher spectrum efficiency. The conclusion was based on the results of a combination of computer simulations, laboratory evaluations using a Rayleigh fading simulator, and field tests.

The improvement in spectrum efficiency compared with analogue mobile telephone systems was largely due to major advances in speech coding and channel coding. The Swedish FDMA system incorporated a RELP-type 16-kb/s speech coder with a permissible bit error rate of 1%, and channel coding the performance of which was optimized to suppress the effect of fading dips caused by the rapid fading due to multipath propagation. Additional facilities to deal with fading were soft channel decoding and interleaving. These measures to counter fading yielded a significant reduction in the required protection ratio. The GMSK modem also contributed to the good spectrum efficiency through its combination of moderate protection ratio and fairly narrow modulation spectrum.

The difference in spectrum efficiency between analogue and digital transmission is shown in Fig. 1.2, which compares three cellular system options:

a) Analogue speech transmission and FM modulation

b) Digital speech transmission without channel coding

c) Digital speech transmission with channel coding

In all three cases it is assumed that a duplex band of 2 x 10 MHz is available. In option b), the input data rate to the modulator is 16 kb/s, which corresponds to a necessary channel separation of 15 kHz. In option c), the data rate going through the channel coder increases from 16 to 27 kb/s, which requires channel spacing of 25 kHz – in other words, the network in this case has available 400 two-way speech channels. The overall spectrum efficiency is also determined by the required reuse distance between co-channel cells (cluster size), which, in turn, depends on the local mean of the protection ratio, K_I over the rapid fading. A typical requirement for analogue mobile telephone systems (without diversity) is K_I = 18 dB. Results from lab tests indicate that K_I = 20 dB is required in option b) and K_I = 13 dB in option c). The improvement going from b) to c) can be explained by the considerable diversity gain from channel coding. The protection ratio is the required power ratio between the wanted signal C and the cochannel interference I that gives adequate transmission quality. e.g. $K_I = (C/I)min$

Comparison of spectrum efficiency

	Analogue system (companded FM)	Digital system without channel coding	Digital system with channel coding
Data rate, speech encoder	-	16 kb/s	16 kb/s
System data rate	-	16kb/s	27kb/s
Channel spacing	25 kHz	15 kH/z	25 kH/z
Protection ratio (local mean)	18 dB	20 dB	13 dB
Cluster size	3x7	3x9	3x3
Spectrum efficiency:			
Channels per MHz per cell	1.9	2.4	4.4
Traffic per cell for 10 MHz system	12.4 e	17.2 e	35.1 e

C and I subject to Rayleigh fading
120° sector antennas. Each base station serves 3 cells

Fig. 1.2

Besides by K_I, the required reuse distance, D, is also determined by the distance-dependence of the global propagation attenation (propagation exponent), the structure of the shadow fading (variance of the log-normal distributions and the correlation between the fading in C and I), and the required area availability (the proportion of the area of a cell in which the local mean of C/I exceeds the protection ratio). The required cluster size is derived from the normalized reuse distance, D/d (d: cell radius). The cluster size is the number of cells with different channel allocations that is required to enable co-channel cells to be adequately separated. The cluster size of 3 x 3 shown in Fig. 1.3 is often used for the GSM. (The local mean of the protection ratio for GSM is 9-10 dB). Each base station site serves three cells.

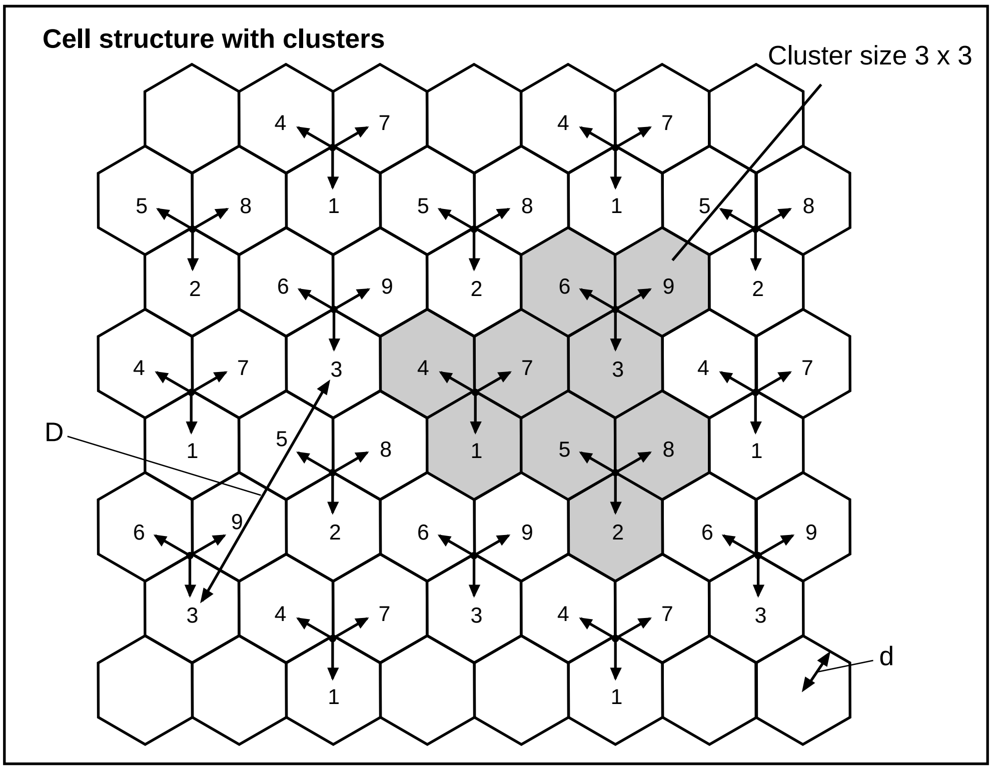


Fig. 1.3

Simulations based on typical propagation characteristics gave the relationship shown in Fig. 1.4. The figure shows, for instance, that for 90% area reliability (C/I > K_I over 90% of the cell) and K_I = 18 dB (local mean over the Rayleigh fading), a cluster size of 21 is required. This implies that seven base-station sites, each serving three cells, must have different channel assortments. Accordingly, the total number of radio channels available to the system must be distributed over 21 cells. (See also module S4).

The number of radio channels available per cell is derived from the cluster size, the channel spacing and the total frequency band available. If a loss system (Erlang B) with 2% permissible blocking is assumed, the carried traffic per cell will be that shown in Fig. 1.2. (See also module S3).

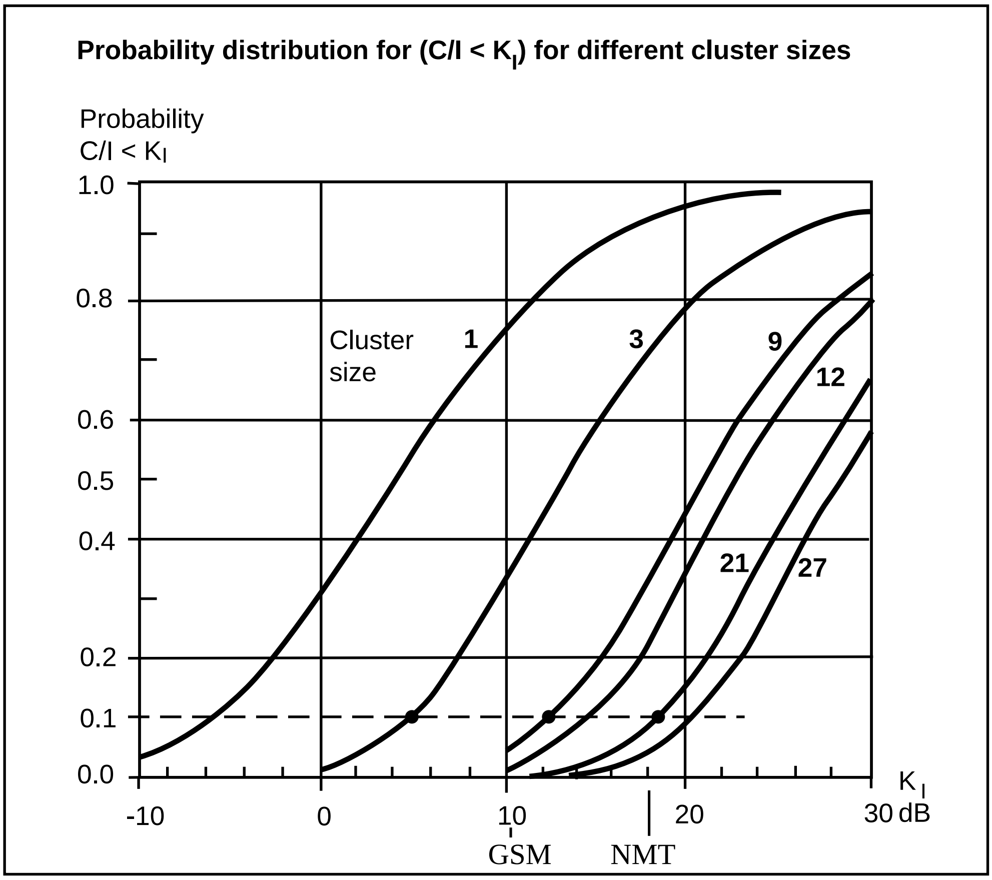


Fig. 1.4

As is evident from Fig. 1.2, digital transmission with channel coding achieves a spectrum efficiency three times higher than existing analogue mobile telephone systems. Although channel coding implies an increased input data rate to the modulator – in other words, wider channel spacing than in a corresponding system without channel coding – this is more than compensated by the considerable reduction in the required protection radio. A significant improvement in the overall spectrum efficiency is obtained.

TDMA

The use of digital transmission means that other forms of multiple access can be used besides FDMA. The most readily available option is TDMA, possibly combined with time duplex. This offers further advantages in terms of system performance and cost savings. A summary of the advantages of TDMA is presented in Fig. 1.5.

Advantages of TDMA (+ time duplex)

Wide radio channels reduce requirements on frequency stability and selectivity

TDM instead of FDM replaces analogue high-Q filters with digital VLSI

Fewer radio units and simpler antenna filters at the base

No duplex filter needed at terminals

Idle periods between the time slots for transmission and reception of user traffic can be used for determination of the local radio environment at the terminal and system signalling.

Fig. 1.5

The possibility of listening or transmitting in other frequency or time slots in idle periods during each frame affords important system benefits. Such periods can be used for system signalling, preparing for handoff and, with an antenna diversity arrangement, for selecting and connecting a suitable antenna to the receiver input before the reception time slot occurs. TDMA also allows a terminal to transmit and receive in different time slots (time duplex). This eliminates relatively expensive and bulky duplex filters.

These advantages often outweigh the disadvantages of TDMA. The drawbacks are listed in Fig. 1.6.

Drawbacks of TDMA

Higher peak power level from transmitter at a given mean power level (determines range)

Wide modulation bandwidth can result in intersymbol interference due to multipath propagation (need for adaptive channel equalization)

Greater equipment complexity (requires advanced VLSI with low power consumption)

Increased channel spacing which reduces flexibility of frequency planning

Fig. 1.6

1.3 System options for the pan-European GSM

The choice of multiple-access arrangement for GSM was largely based on the results of the evaluations made in Paris in late 1986 and early 1987.

The majority of the test systems were based on TDMA. The main contenders for the GSM were narrowband TDMA and wideband TDMA. Several different versions of narrowband TDMA were evaluated. One French proposal (SHF-900) combined TDMA with an advanced frequency-hopping procedure; The proposal from the Swedish Telecom Administration was for 8-PSK and as little as four time slots in the TDMA frame. This gave such a narrow modulation bandwidth that channel equalization would only have been necessary in very difficult propagation conditions.

However, the TDMA option specified by the GSM group corresponded most closely to the experimental system developed by Ericsson (DMS 900). The main competitor was wideband TDMA, which, in terms of performance, for the most part was on a par with the best versions of narrowband TDMA. An interesting finding of the Paris tests was that several systems achieved roughly the same spectrum efficiency (see Fig. 1.7).

The differences among the systems in respect of required channel spacing per speech channel were offset by different protection-ratio requirements. As described earlier, different protection ratios result in different cluster sizes. Roughly speaking, if a system can cope with half the channel width, an increase in the cluster size by a factor of two can be allowed without any impact on the total spectrum efficiency.

The CD 900 system (SEL, Germany), based on wideband TDMA, incorporated very powerful channel coding, which increased the required bandwidth but resulted in a much lower protection ratio than in narrowband TDMA. The pilot system, MAX 2, proposed by Swedish Telecom Administration, was designed for the narrowest possible channel width per speech channel, which resulted in a fairly high protection ratio.

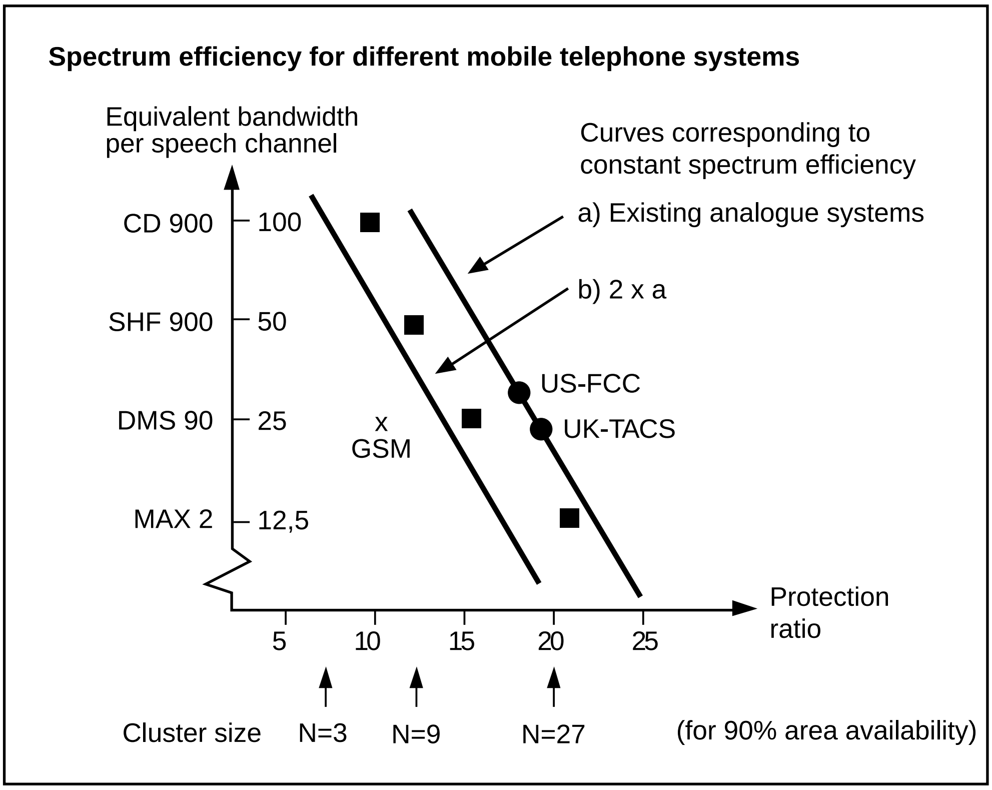


Fig. 1.7

All the systems achieved better spectrum efficiency than in analogue mobile telephone systems. This was one of the criteria stipulated at the outset that the digital systems would have to meet. The evaluation also seemed to show that digital systems can provide better speech quality also during rapid fading.

Main features of the DMS-90 system (narrowband TDMA; ERA proposal)

A block diagram of the ERA's test system is shown in Fig. 1.8, and the system's multiple-access structure in Fig. 1.9. The channel coding was supported by interleaving and frequency hopping. This gave a considerable diversity gain – i.e. a low protection ratio – even in communications with portable terminals (quasistationary radio channel). The interleaving divided a 384-bit block from the channel coder into four subblocks of 96 bits, which were distributed among four time slots. (Each time slot could accommodate 2 x 96 user bits.) The influence of the time dispersion of the propagation channel was compensated for by the adaptive equalizer, which also might give multi-path diversity. The impulse response of the radio channel was determined with the help of a training sequence at each time slot of the TDMA frame.

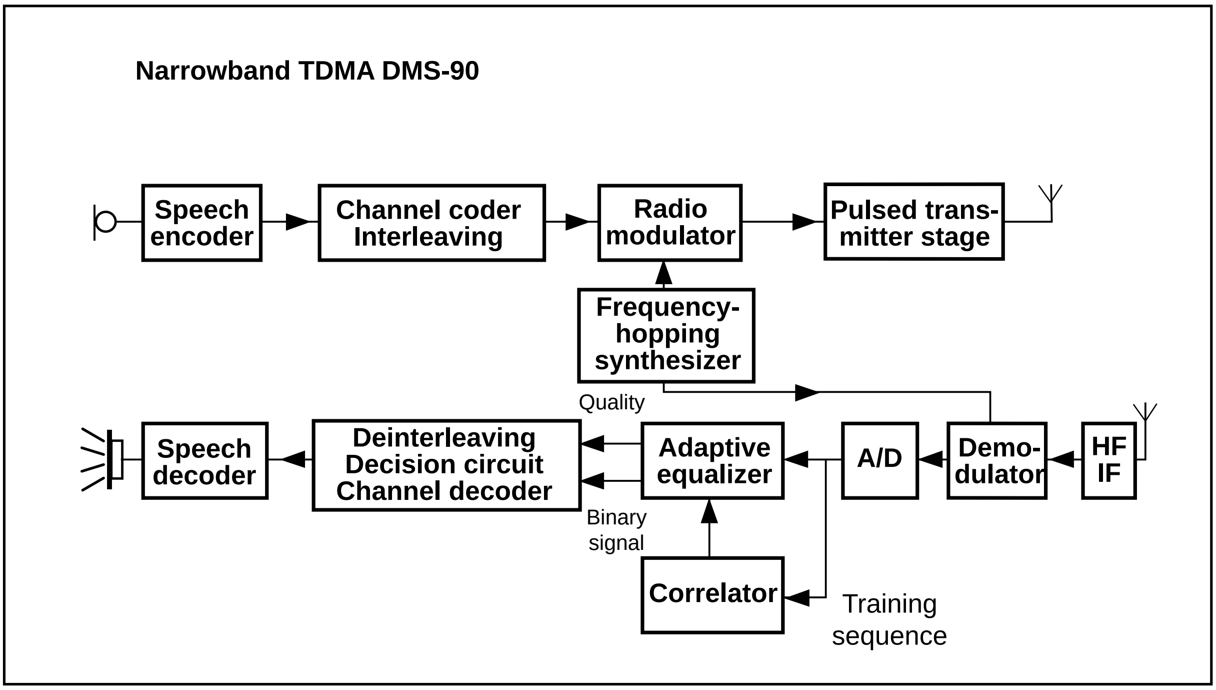


Fig. 1.8

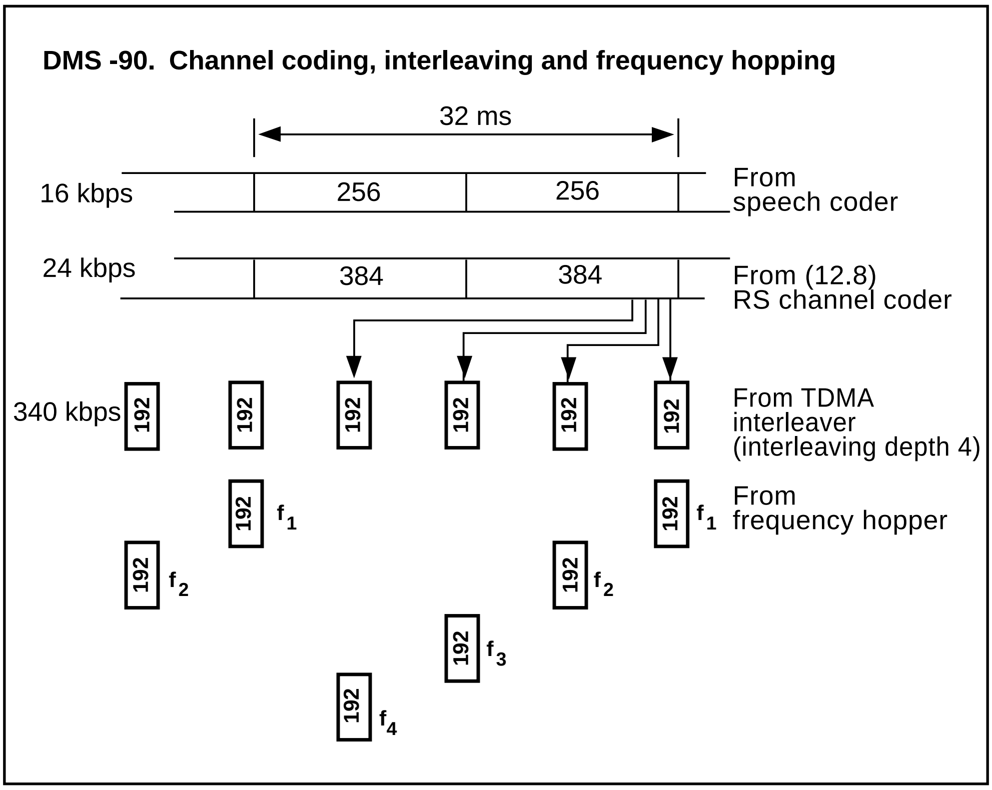
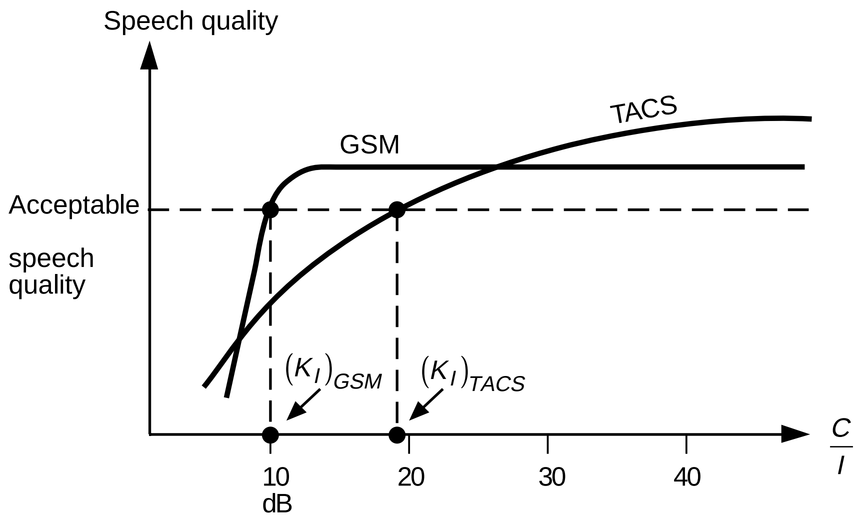


Fig. 1.9

Since the initial development phase and the Paris evaluations, further development of the radio transmission system has resulted in substantially better performance. The final GSM specification gives a much better protection ratio: $K_I < 10$ dB (largely achieved through further refinement of the channel coding). Further rapid advances in speech coding have also made possible a gradual reduction in the data rate from the speech coder without any degradation of speech quality. In 1994 a half-rate speech coder has been standardized, further improving the spectrum efficiency by a factor of two. (Some people have expressed the opinion that, instead of reducing the data rate, the advances made in speech coding should be used to improve speech quality.)

The relationship between C/I or C/N on the one hand and subjective speech quality on the other differs between analogue and digital transmission (see Fig. 1.10). (TACS is the British analogue mobile telephone system.)

Speech quality in the GSM and TACS



K_I: Local mean of the protection ratio

Fig. 1.10

With digital transmission using FEC channel coding, speech quality is almost constant down to a threshold that corresponds to the error-correction limit of the channel decoder. If the input signal to the receiver falls below this level, the error-correction fails and speech quality rapidly degrades. If the quality of the input signal to the receiver is high, analogue mobile telephone systems are superior, since the speech coder causes some quality degradation even if the C/I and C/N are high enough for no transmission errors to occur.

Main features of CD-900 (wideband TDMA, German-French proposal)

One of the demonstration systems evaluated in Paris was the CD-900 or wideband TDMA system, which was developed by a consortium led by German SEL. The system solution was a further development of an earlier military project – Autotel. The technical performance and spectrum efficiency of the CD-900 system were on a par with those in the system options based on narrowband TDMA. Because the published information available is limited, to gain a general idea of the system we need to examine the combined available information on Autotel and the CD-900.

The main system characteristics of wideband TDMA are a wide modulation bandwidth through a combination of many time slots per TDMA frame and a substantial bandwidth expansion through low-rate channel coding. The channel coding is based on near-orthogonal codes, i.e. optimum soft decoding can be based on matched filters implemented by a correlation procedure. This is shown in diagrammatic form in Fig. 1.11, which applies to Autotel. A group of four information bits is coded into 16 chips. In the CD-900 system, five information bits are coded into 32 chips, and an additional sixth bit is transmitted via the polarity of the chip sequence. A group of 32 chips can be considered to form a symbol in an alphabet of size $2^6 = 64$. If the QAM arrangement is included the size of the symbol alphabet is 128.

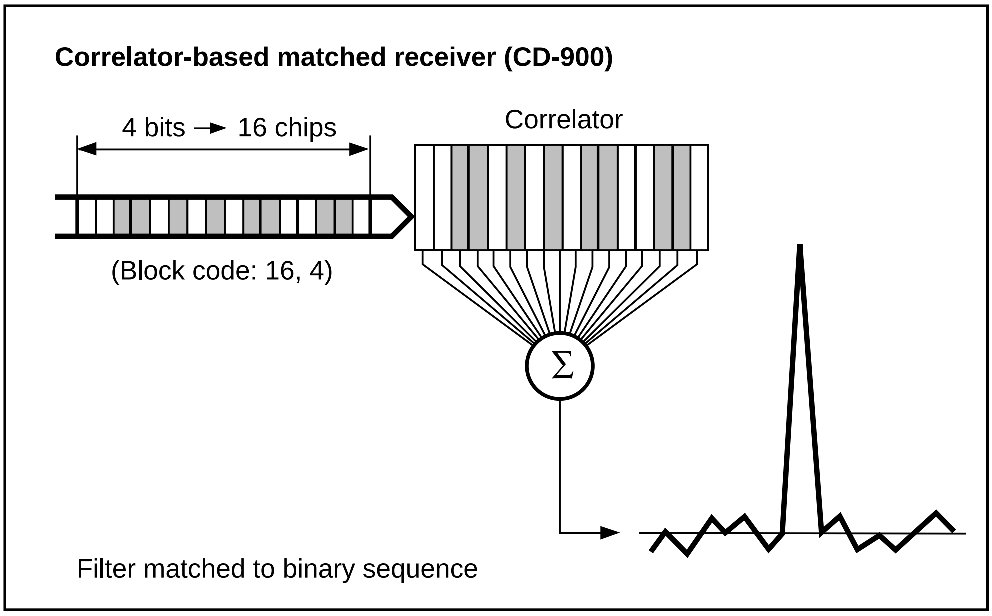


Fig. 1.11

The powerful channel coding produces a high coding gain, i.e. a substantial reduction in the required C/I with respect to co-channel interference. A further reduction in the required protection ratio in a rapid fading situation is obtained from a considerable diversity gain from frequency diversity. The reason is the wide modulation bandwidth which, for most propagation conditions, is much greater than the correlation bandwidth of the radio channel. The combined coding and diversity gains enable the local mean of the protection ratio to be brought down as low as $(C/I)_{min} \approx 4$ dB. This means that a cluster size of three is adequate.

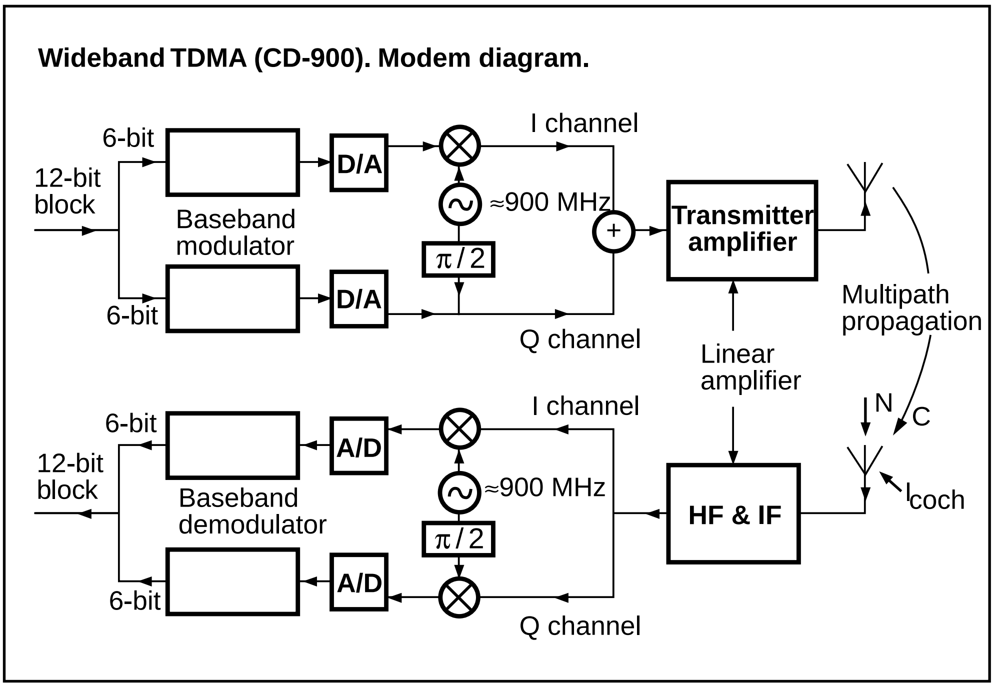


Fig. 1.12

To achieve reasonable spectrum efficiency despite the large bandwidth expansion, due to channel coding the wideband TDMA system uses linear modulation (QAM), i.e. a linear transmitter amplifier has to be used (see Fig. 1.12). A coded radio symbol carries 12 information bits, 6 bits on each of the I and Q channels. To obtain a sufficient modulation bandwidth both for accurate measurement of the radio channel impulse response (for setting of the channel equalizer) and for a high frequency diversity gain, a TDMA arrangement with 63 time slots was used (60 traffic channels and three channels for signalling), which were used jointly by the three sector cells belonging to each base-station site. Thus, each cell is allocated an average of 20 traffic channels. (Cluster size of 3 means that all base-station sites can use the same radio channels.)

Advanced digital signal processing is used for channel equalization (see Figs. 1.13 and 1.14). Each data burst, which comprises a number of blocks (radio symbols) each corresponding to 12 information bits, starts with a synchronization and training sequence. The wide modulation bandwidth allows accurate measurement of the radio channel's impulse response, h(t), and thus advanced channel equalization, which in an optimum way adds the signal power from the different propagation paths.

Channel equalization is implemented by introducing before the symbol detector a filter matched to the impulse function h(t) of the radio channel. This filter convolvs the received burst (excluding the training sequence) with h(T-t). This gives optimum addition of the signal power from propagation paths with different delays, thereby eliminating intersymbol interference and at the same time achieving a high frequency diversity gain. After this initial signal processing, detection takes place by determining from which of the 32 matched filters the highest absolute value is obtained at the sampling instant. (The symbol alphabet comprises 32 near-orthogonal symbols.) In addition, the polarity of the output signal from the selected filter is measured.

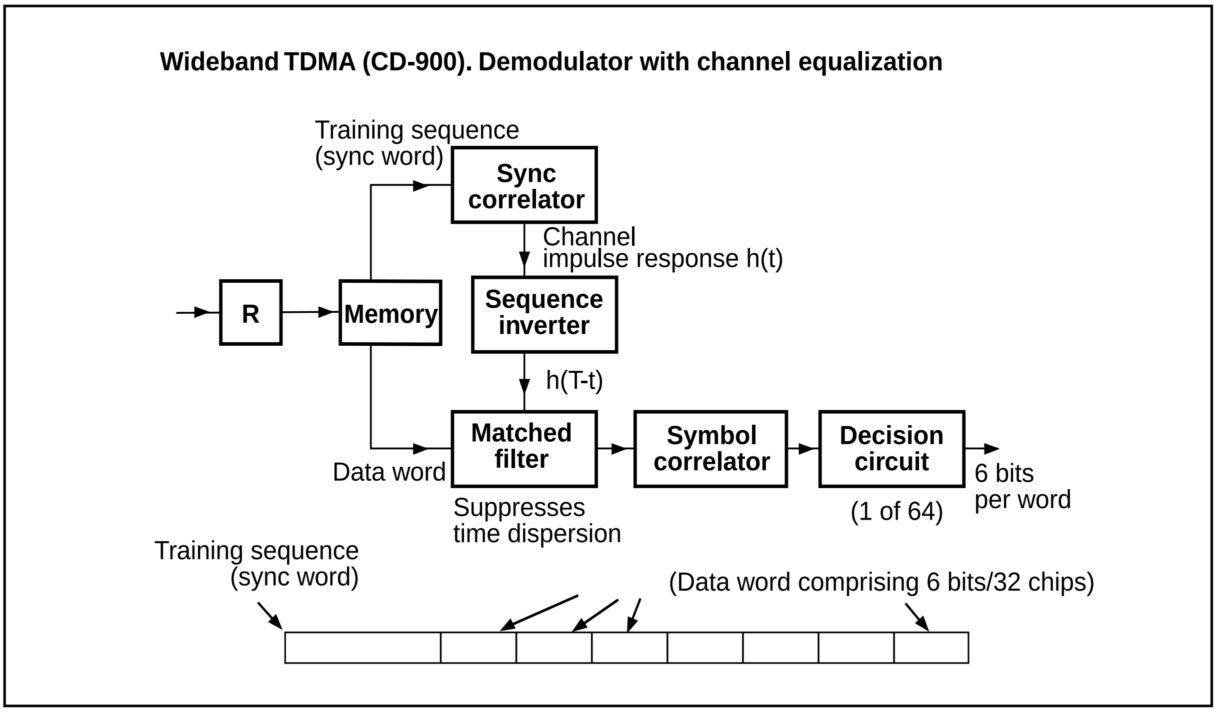


Fig. 1.13

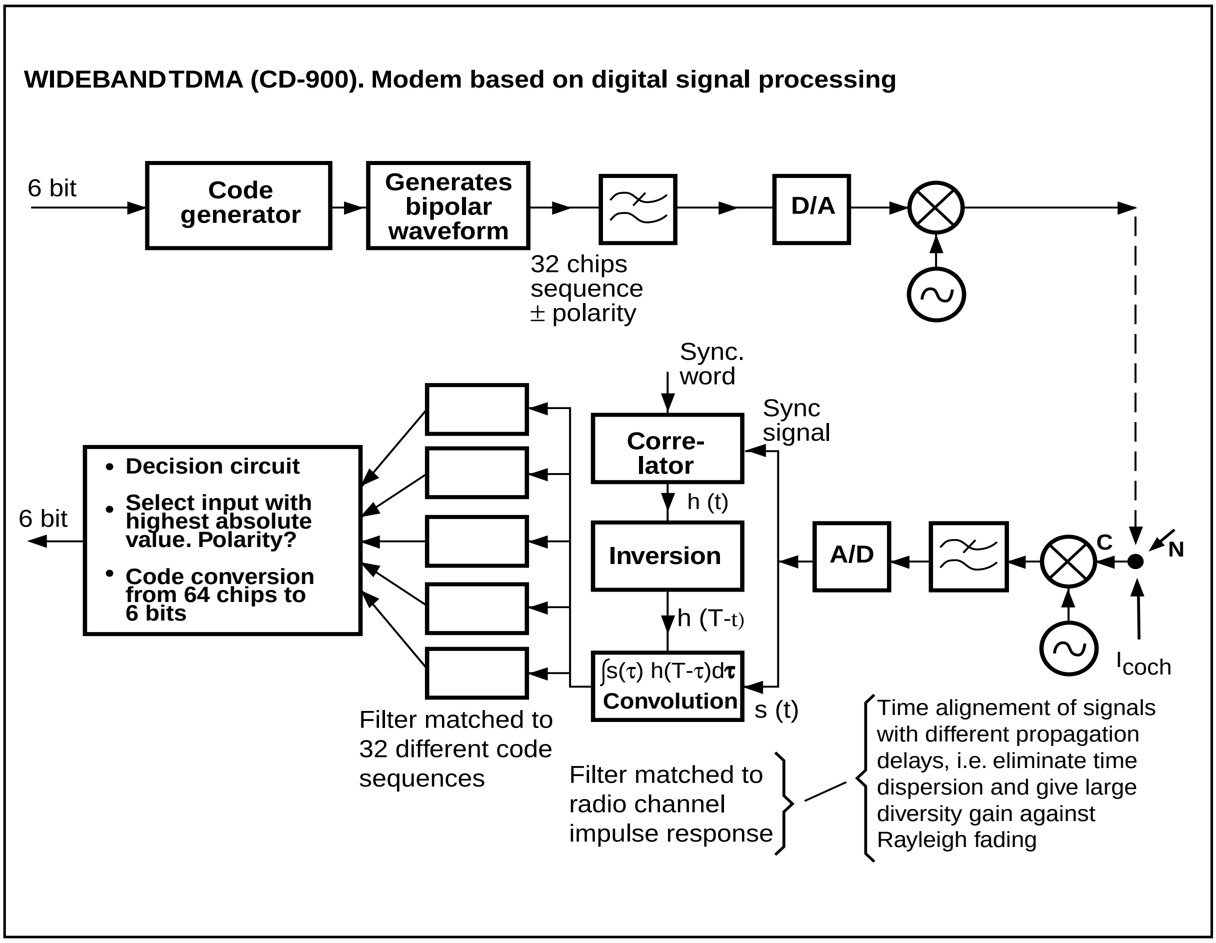


Fig. 1.14

Corresponding analogue signal processing (in SAW filters instead of digital processing) but with noncoherent detection was used in the Autotel system. Examples of the output signal from the sync correlator (i.e. the measured impulse response from the radio channel) have been published (see Fig. 1.15).

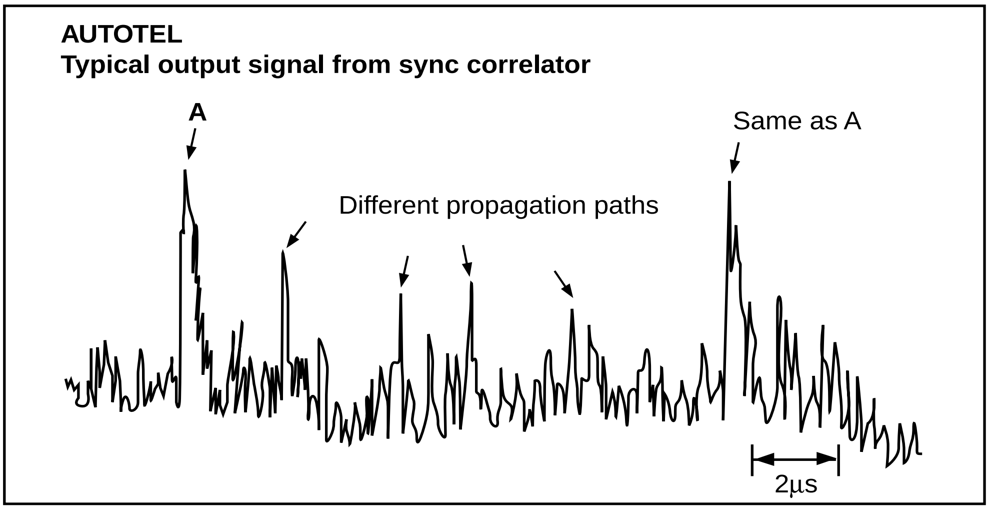


Fig. 1.15

The overall performance of the testbeds based in N-TDMA and W-TDMA was very similar. So why was narrowband TDMA chosen instead of wideband TDMA?

A comparison between narrowband and wideband TDMA based on the finding of the Paris evaluations indicates comparable spectrum efficiency and speech quality. The choice therefore had to be made on the basis of other system and implementation characteristics.

An advantage of wideband TDMA in some cases is the large number of speech channels per carrier and the fact that all base-station sites can use the same radio channels. This reduces the cost of the base-station equipment if every cell is allocated a large number of traffic channels, and also facilitates frequency planning of the individual networks. However, the wide channel spacing imposes considerable limitations on the gradual transfer of frequencies from the analogue to the digital mobile telephone system. Moreover, the cost per speech channel is high for small base stationsites that are located outside urban conurbations and only have a traffic volume for a small number of radio trunks.

Some other drawbacks with wideband TDMA are shown in Fig. 1.16. In its assessment, the GSM group attached great importance to the apparently greater technical risk inherent in wideband TDMA. Not only were highly linear transmitter amplifiers an unproven technology in mobile radio applications but the wide modulation bandwidth also imposed a heavy demand on high-speed digital signal processing for channel equalization and detection.

Why wideband TDMA was rejected?

- Inflexible frequency planning
- Too complex base stations in rural locations
- High peak power in portable terminals
- More complex, high-speed digital signal processing (technical risk, increased power drain)
- Linear modulation (technical risk with linear transmitter amplifiers)

Fig. 1.16

2. Overview of the radio subsystem for GSM

2.1 Introduction

This section provides an overview of the way in which a radio link operates after a traffic channel has been set up. There are two options: at present, full-rate speech coders having a data rate of 13 kb/s are used. However, the GSM has been designed for subsequent changeover to half-rate speech coders, by allowing both full-rate and half-rate traffic channels to be used. As well as speech, data transmission at different speeds can also take place over full-rate or half-rate traffic channels (this is dealt with in the next section).

During a call, the terminals and base need to exchange large amounts of information, especially to prepare for the next handover. Sometimes it may also be necessary to allocate a different radio channel in the same cell for the connection, e.g. if the channel is experiencing strong interference. The Associated Control Channels are used for this system signalling, utilizing either a Slow Associated Control Channel (SACCH) or a Fast Associated Control Channel (FACCH).

The FACCH is used during the actual channel-switching phase, during which a wealth of information needs to be transferred. The SACCH is used, for instance, for transferring from terminal to base measurements of the received signal levels from nearby cells. The information is used by the mobile switching center to determine the next cell and radio channel to which the call will be transferred on handover. In the outward direction, information is sent on the transmitter power to be used by the terminal. The SACCH has several additional functions as mentioned below. To indicate if a 57 bit sequence is used for signalling (FACCH) an associated one bit "stealing flag" is set.

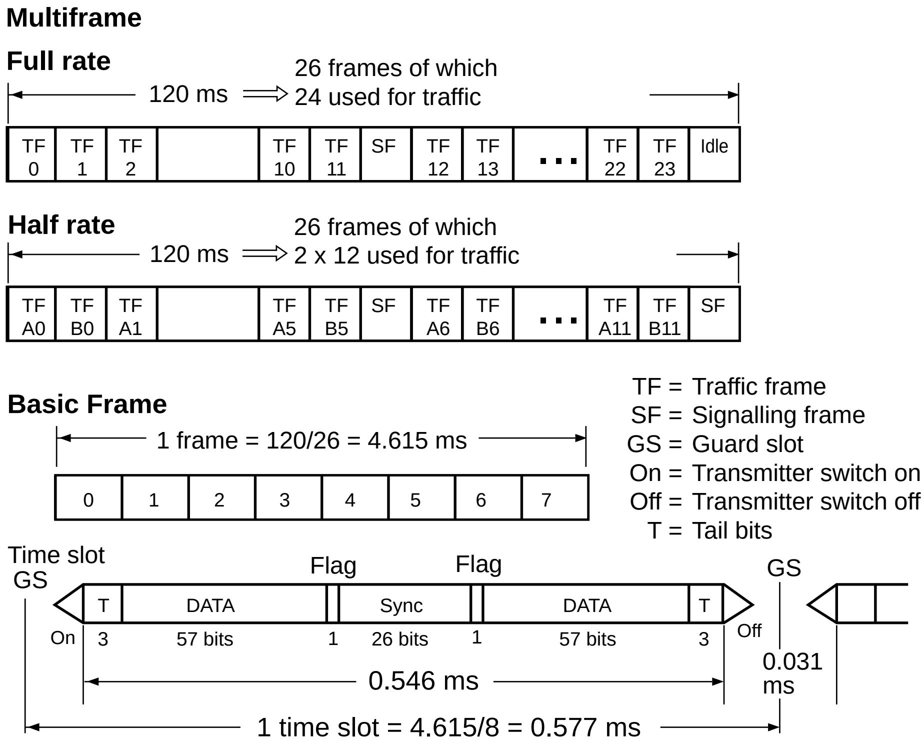
The most important transmission specification items ("air interface") are shown in Fig. 2.1. The frequency band comprises 2 x 25-MHz in a duplex arrangement with 124 duplex channels with 200 kHz channel spacing. This channel spacing allows a system data rate of about 270 kb/s with GMSK modulation and modest adjacent channel selectivity requirement (9 dB). 270 kb/s corresponds to a symbol length of 3.7 µs.

Radio transmission specification for GSM		
Frequency band:		890 - 915 MHz
(frequency duplex)		935 - 960 MHz
Channel spacing:		200 kHz
Modulation:		GMSK
System data rate:		271 kb/s
TDMA	Frame:	4.6 ms
	Time slots:	8 x 0.58 ms
Data rate (full-rate traffic channel):		22 kb/s
Speech coder:		Regular Pulse Exited LPC 13 kb/s
Diversity:		Channel coding
		Interleaving
		Frequency hopping
		Channel equalization

Fig. 2.1

2.2. TDMA structure for traffic channels

Each radio channel (carrier) is divided by a TDMA arrangement into 8 channels used for user traffic and system signalling. The Slow Associated Control Channel (SACCH) is implemented by multiplexing each physical channel (a certain time slot in a sequence of TDMA frames) between two virtual channels within a multiframe of 26 basic TDMA frames. Most of the time slots are used for traffic but, in some of the basic TDMA frames, the eight time slots are used instead for the SACCH. Each SACCH frame are associated with 8 traffic channels (one time slot per traffic channel). A multiframe has a time of 120 ms (see Fig. 2.2).



Gross data rate for full-rate physical channel: *114/0.004615 = 24.7 kb/s*

Traffic channel:

24 x 114 = 2,736 bits are sent during 120 ms.
Therefore, data rate = 2,736/0.12 = 22.8 kb/s

SACCH:

114 bits are sent during 120 ms. Therefore, data rate = $\frac{114}{0,12} \approx 1kb/s$

Fig. 2.2

If the traffic consists of full-rate channels (8 traffic channels per carrier), the time slots in a multiframe during 24 frames are used for user traffic and during one frame for signalling. One TDMA frame is not used. (This frame is used by the terminals to read the base identity ("BSIC") of carriers from other cells.) If the traffic comprises half-rate channels (16 traffic channels per carrier), each time slot in the frame can handle two traffic channels, which occupy the time slot alternately during 12 frames of a multiframe. The remaining two basic TDMA frames are used for two SACCH signalling channels. Each SACCH is associated with 8 traffic channels.

2.3. Structure of data bursts in a TDMA time slot

A time slot of length 0.577 ms is used as follows (see Fig. 2.2).

Guard slot, timing advance

To prevent data bursts from different terminals overlapping in the input to the base-station receivers, a guard slot with a time of 31 μs has been introduced. This is needed, above all, to cope with variations in the propagation time to terminals at different distances from the base. The guard slot corresponds to a two-way propagation path of about 4.5 km.

This is considerably less than the maximum range (35 km). Therefore, to prevent data bursts from different terminals overlapping in the input to the base-station receivers, the base-station instructs the terminals to insert a suitable delay between received and transmitted data bursts. The delay is adjusted such that a transmitted burst from the terminal reaches the base-station receiver at the right instant relative to the time-slot structure. The closer a terminal is to the base station, the greater will be the delay inserted. Thus, regardless of how far the terminal is away from the base, the bursts to the base receiver will always arrive in the middle of the intended time slot. The measurements made to determine the timing advance can also be used to calculate the distance between the terminals and the base. (During the first contact from a terminal to the base the terminal has not yet been instructed about the suitable timing advance. Therefore signalling bursts with much larger guard times are used).

On/Off switching of transmitters. Tail bits.

A transmitter pulse corresponding to a data burst must have rounded start and end. If not, additional widening of the spectrum relative to the modulation spectrum for GMSK will occur. A small part of the guard slot is used for this rounding of the transmitter pulse. To facilitate channel equalization, each burst start and ends with three tail bits (0,0,0). (The channel equalizer has to cope with time dispersion up to four symbol intervals. The three tail bits ensure that channel equalization can start and finish in a known state.)

Synchronization and training sequence

The impulse response of the radio channel can change drastically during a frame of 4.6 ms. This means that for each time slot, the receiver must carry out bit synchronization and set the channel equalizer. The impulse response can sometimes change even during a burst (if the terminal velocity is very high and especially when moving up to 1800 MHz), which means that if the setting of the channel equalizer was optimized at the beginning of the burst, equalization may be suboptimum for the last part of the burst.

To avoid the complication of having to adapt the channel equalizer to variations in the impulse response of the radio channel during a time slot, short slots are used and, in addition, the training sequence is placed in the middle of the burst. The setting of the channel equalier is based on a known bit sequence of 26 bits, which is also used for synchronization. This sequence can also be used to equalize the first part of the data burst, since the received data burst is stored in a buffer before channel equalization and detection are initiated.

To limit risk for synchronizing to a distant strong carrier 8 different training sequencies ("color codes") are used.

Transmission of user information. Fast Associated Control Channel

Each data burst comprises two user sequences of 57 bits each. A flag bit is associated with each 57-bit sequence and this denotes whether the sequence contains normal speech information or if the sequence is instead being used for system signalling (FACCH). (Fast signalling is needed during handover, for example. A short break in the speech transmission will hardly be noticeable, since the speech coder fills out the slot with information taken from the previous speech frame.)

2.4. Multiframe with SACCH

During a multiframe of 120 ms, 24 or 12 bursts will be allocated to a traffic channel (corresponding, respectively, to full-rate and half-rate traffic channels). In the former case, the average interval between bursts will be 5 ms, i.e. 200 bursts a second. The gross data rate (including channel coding) for a full-rate channel will be 22.8 kb/s (200 x 114), and for a half-rate channel 11.4 kb/s. In addition to the above, a Slow Associate Control Channel (SACCH) is also available.

A full-rate traffic channel shall carry the signal from a 13-kb/s speech coder. Channel coding increases the data rate from 13 to 22.8 kb/s. For each speech frame of 20 ms, 260 bits will be output from the speech coder and 456 from the channel coder. A 456 block is divided into 8 blocks of 57 bits, which are interleaved over time slots within 8 consecutive traffic TDMA frames (signalling frames and empty frames are skipped). Each burst carries information from two adjacent 20 ms sequences from the channel coder.

2.5. Duplex arrangement

The two time slots corresponding to a two-way traffic channel are mutually displaced in time (see Fig. 2.3). The figure corresponds to the case in which a base station not using frequency hopping has been allocated four carriers, each of which carries eight physical channels in a TDMA frame. The mutual displacement of the time slots for the outward and inward directions corresponds to a time duplex arrangement. No duplex filter is required in the terminals, therefore. Instead there is a fast T/R switch, which alternately connects the transmitter and receiver to the antenna.

A terminal receiver also has time during a TDMA frame to measure the carrier level of a signal from one of the six nearby cells (see Fig. 2.4). This gives valuable information for selection of the optimum base station for handover. (This procedure is discussed in the next section on system signalling.) To provide enough time for receive, transmit and listening in pace with the TDMA frames, the terminal's frequency synthesizer must be able to change frequency rapidly. This has influenced the specification of the TDMA frame length. An option is also frequency hopping, in which the utilized duplex channel is changed for each TDMA frame. To avoid collisions a coordinated hopping pattern must be used. The base informs the terminals about the hopping pattern on the Broadcast Control Channel (see section 3.2) and on the SACCH.

Physical channels of a base station with four radio channels (carriers). Total of 4 x 8 = 32 physical channels accessible

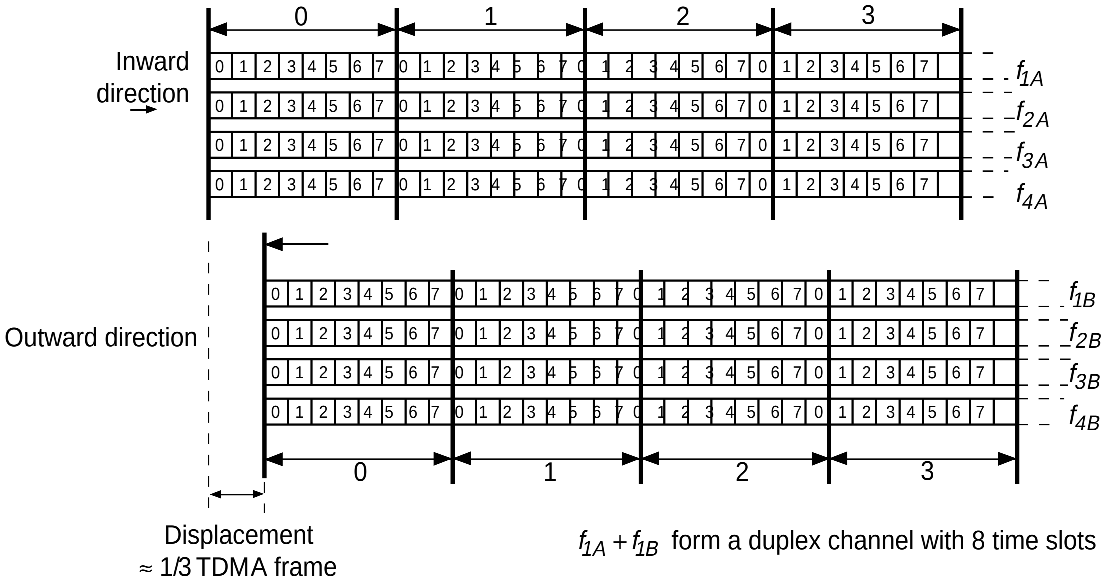


Fig. 2.3

TDMA, time duplex, frequency hopping and listening-in by terminal

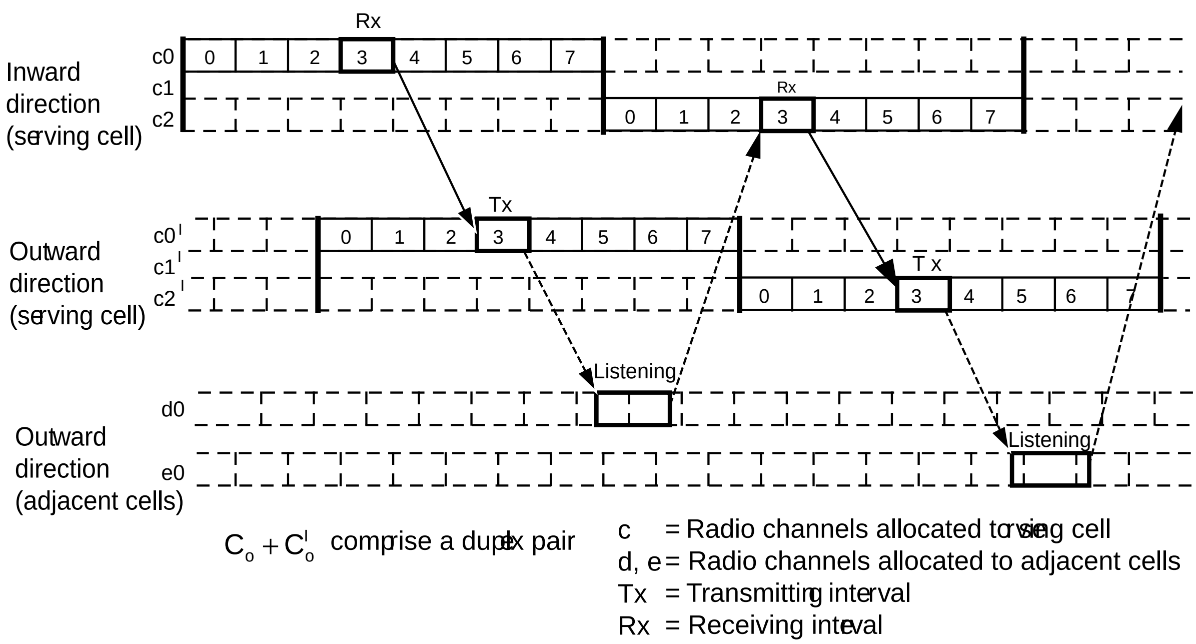


Fig. 2.4

2.6. Diversity against rapid fading

Instead of antenna diversity, the GSM uses a combination of channel coding, interleaving and coordinated frequency hopping. In addition, the modulation bandwidth is so great that additional frequency diversity (multi-path diversity) is obtained in conjunction with channel equalization if the radio channel is subject to fairly strong time dispersion (which results in frequency selective fading). Together, these give such high diversity and coding gains that the required protection ratio (the local mean across the rapid fading) will typically be 9 dB, which is compatible with a cluster size of 3 x 3.

Interleaving a **full-rate channel**, means that the 456 bits in a 20-ms speech frame are split up into 57 bits sequencies, which are spread over 8 TDMA frames, that is over 40 ms (see section 3.3.). If the duration of a fading dip is not more than a few milliseconds, meaning that only one time slot in a physical channel (one TDMA frame) is affected, the deinterleaving on the receiver side will change an error burst to a relatively random error sequence spread over 8 code words. Thus, 1/8 of the bits in each code word will be subject to a bit error rate of about 50%. The FEC is so powerful that error correction is possible. Greater interleaving depth cannot be used, as it would give rise to excessive transmission delay. Because the interleaving depth is only four for a half-rate channel, the diversity gain will be somewhat lower.

Fading dips that are longer than the channel coding with interleaving can cope with, occur over quasi-stationary propagation paths – something that affects portable terminals in particular (a fading dip could affect several consecutive TDMA frames). The situation can be improved considerably with the introduction of coordinated frequency hopping, whereby each physical channel is switched between different radio channels that can be chosen, for instance, from a 4-group. For each TDMA frame a new carrier frequency is used. The size of a typical frequency hop is usually large enough to give uncorrelated fading due to multipath propagation in the different frequency slots (see Fig. 2.4).

Another advantage of frequency hopping is that averaging occurs in respect of co-channel interference from different cells.

3. Detailed systems description GSM

3.1 Introduction

The previous section gave an overview of the use of the radio channels during a call. The setting up of a call requires extensive system signalling for **synchronization** of a terminal to the base station, for **registration** of the terminal and for allocation of a traffic channel. In preparation for handover, the terminals also measure the signal levels from adjacent-cell base stations and transmit this information to the base (Mobile Assisted Handover).

An alternative **multiframe** with a length of 51 TDMA frames is used for most of the system signalling, instead of the multiframe having a length of 26 TDMA frames as in the case of traffic channels. The reason why two different multiframe lengths have been chosen is explained in section 3.2.3. To incorporate two multi-frames of different lengths in the same overall structure, the TDMA hierarchy has been extended with a **superframe**. In addition, the requirement for secure encryption necessitated the introduction of a further higher level in the TDMA structure – the **hyperframe**. See Fig. 3.1. The hyperframe includes more than two million TDMA frames and has a duration of about 3.5 hours. The running number of a basic TDMA frame in a hyperframe is one of the parameters determining the encryption key (see section 3.6.3). An overview of the system signalling is given in section 3.2.

The achievement of considerable diversity gain is based on advanced **channel coding**. This is different for full-rate and half-rate traffic channels and for speech and data transmissions with different data rates. The signal channels are also protected by channel coding. An outline of a few of the many channel-coding cases that occur is given in section 3.3, where the **interleaving structure** for a full rate speech channel is also described.

The GSM system uses GMSK modulation, which implies moderate filtering of the modulation spectrum for MSK, on which a description of the **modem** can be based. MSK is related to 4-QAM but it includes modifications that complicate **channel equalization**. This can be made simpler if the modulation of the signal to the equalizer follows as closely as possible the basic 4-QAM arrangement. For this reason, further modifications have been made to the MSK modulator and demodulator arrangement. Modems and channel equalization are discussed in sections 3.4 and 3.5.

The radio subsystem of the GSM is part of a highly complex network structure, which includes mobile exchanges and the public telephone transport network with common channel signalling. Advanced network facilities are required to handle roamers (**mobility management**) and to protect the network and its users from unauthorized usage and listening in. Before a connection is established, a check is made to ensure that the terminal is authorized to use the network for the requested service (**Authentication**) and that the terminal has not been stolen (subscribers can report stolen units to the **Equipment Identity Register**). To achieve a very high level of security against listening in and unauthorized connection to the network, user traffic and some sensitive system signalling can be **encrypted**. A similar arrangement is used for authentication (electronic signature). An outline of this is given in section 3.6.

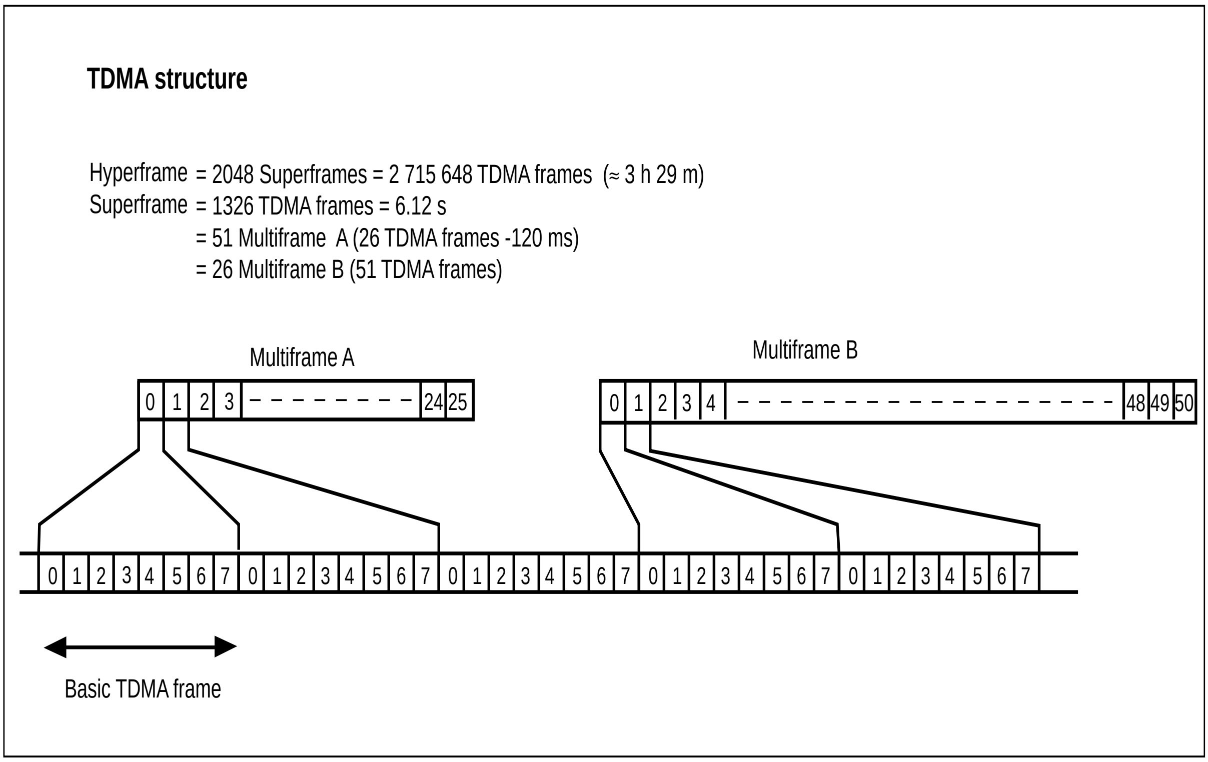

TDMA structure
Hyperframe = 2048 Superframes = 2 715 648 TDMA frames (≈ 3 h 29 m)
Superframe = 1326 TDMA frames = 6.12 s
= 51 Multiframe A (26 TDMA frames -120 ms)
= 26 Multiframe B (51 TDMA frames)
Multiframe A
0 1 2 3 24 25
Multiframe B
0 1 2 3 4 48 49 50
0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7
Basic TDMA frame

Fig. 3.1

3.2. Signalling, TDMA structure

3.2.1. Overview

As a preliminary to the analysis of the signalling procedures, a summary is given below of the different traffic and signalling channels. The slow and fast associated channels (SACCH and FACCH) have already been discussed in section 2 and are therefore not dealt with here.

One of the duplex radio channels (carriers) allocated to a cell is given the task of transmitting the system signalling. A large proportion of the system signalling is sent via broadcast signalling channels, i.e. channels accessible to all the terminals. The carrier that is used in a cell for signalling is therefore called the broadcast carrier. In the outward direction the base station transmits this carrier continuously at a constant power level. (If there is no information to be sent, dummy bursts are inserted.)

Of the eight time slots in each TDMA frame, time slot 0 on the broadcast carrier is used in the outward direction for broadcast or common control channels, and in the inward direction for call requests from terminals (Random Access Channel, RACH). Time slot 1 is used in both directions for the Stand-Alone Dedicated Control Channel (SDCCH). The SDCCH is assigned exclusively ("dedicated") to a terminal for signalling between base and terminal during the setting up of a call or to exchange messages between the base and the terminal, e.g. registering the terminal with a new cell or location area. Time slots 2-7 on the broadcast carrier are used for traffic channels.

Traffic channels	**Signalling channels on Broadcast Carrier**
• Full-rate:	**• Broadcast channels**
Speech (TCH/FS)	FCCH: Frequency Control Channel
9.6 kb/s (TCH/F9,6)	SCH: Syncronization Control Channel
4.8 kb/s (TCH/F4,8)	BCCH: Broadcast Control Channel
2.4 kb/s (TCH/F2,4)	
• Half-rate:	**• Common control channels (CCCH)**
Speech: (TCH/HS)	Inward direction:
4.8 kb/s (TCH/H4,8)	RACH: Random Access Channel
2.4 kb/s (TCH/H2,4)	Outward direction:
	PCH: Paging Channel
	AGCH: Access Grant Channel
• SACCH and FACCH	**• SDCCH: Stand-Alone Dedicated CCH**
(See section 2)	

Fig. 3.2

Terminal synchronization to base station

When a terminal is switched on or moves into the coverage area of a new cell, the first step is for the terminal to check and adjust itself to the local radio environment in the cell. This implies:

- Fine-tuning of the local oscillator frequency (master oscillator) to minimize the frequency error between the terminal and the base station
- Setting the counters that determine the complex TDMA structure, so that the terminal's TDMA timing (TDMA hyperstructure) coincides with that of the base station
- Determination of the network's ID code (several operators in a country may be sharing the GSM band), the cell's ID code and allocated radio channels and the broadcast carriers in adjacent cells

See also section 3.2.2.

Registering with the base station

Registration of the terminal so that it can be reached by incoming calls takes place both when the terminal is switched on and when it enters a cell belonging to another location area. The terminal notices this because the information is transmitted continuously on the BCCH.

As part of the registration procedure, the GSM network checks that the terminal is authorized for connection (Authentication) and may also check that the unit is not listed as stolen (enquiry to the Equipment Identity Register, EIR). Hereafter, the terminal monitors continuously the signalling channel (Paging Channel, PCH) over which the base station pages terminals belonging to that location area (that might comprise several cells).

If the terminal moves into another cell, synchronization to a new broadcast carrier takes place. If the cell belongs to a different location area (apparent from the signalling on the broadcast channel), re-registration takes place so that the network can transfer paging signalling on the PCH to the new location area. (The size of a location area is a trade-off between heavy registration signalling in small areas and heavy paging signalling in large location areas containing many cells.)

Signalling for registration is initiated by the terminal which sends a paging signal over the RACH, whereupon the network assigns an SDCCH to the terminal via a data message on the AGCH. Registration signalling on the SDCCH includes checks to ensure that the terminal is authorized to use the network (authentication, see section 4.2).

As described in section 3.6, successful registration of a roamer with a base station (new location area) results in the Visiting Location Register sending a message to the subscriber's Home Location Register with details of the subscriber's ID and where he should be paged for incoming calls. If necessary a cancellation message is sent to the VLR where the roamer was registered previously.

Call set up

Setting up of calls to or from a terminal requires extensive signalling for transfer of address information and allocation of a radio channel/time slot. Initially signalling is via the CCCH and, subsequently, the SDCCH. The outcome is that the call is assigned to a traffic channel. The procedure makes it possible for the network to determine in which cell within the location area the terminal is situated.

Setting up of a call **to a terminal** is initiated by the network paging the terminal over the PCH in all cells belonging to the location area. The terminal acknowledges the call on the RACH (slotted Aloha). The procedure makes it possible to determine in which cell of the location area the terminal is situated. Thereafter, the network sends a message to the terminal via the Access Grant Channel (AGCH), instructing it to switch over to a given SDCCH, which has an associated SACCH. The SDCCH is used for transmission of the calling and called-party numbers, for authentication, for sending encryption keys, etc. Finally, a traffic channel is allocated to the terminal. The SACCH is used for setting the transmission power.

For setting up a call **from a terminal**, the terminal sends a call request via the RACH. The network sends back details of the allocated SDCCH on the AGCH. Further signalling takes place as described above.

Handover decisions are based on a host of different radio parameters measured both by the terminal and the base-station equipment (MAHO, e.g. mobile assisted handover). The measured data on conditions at the terminal are transmitted to base via the SACCH. The following information is used by the MTX when deciding the best cell for handover:

Carrier level (RXLEV) and connection quality (RXQUAL), i.e. the bit error probability, for the connection between the terminal and the base. Averaging is done over 12 seconds.

Signal levels at the terminal receiver of broadcast carriers from nearby cells.

Distance between base and terminal. This parameter is obtained from the timing advance in the terminal for the transmitted time slots.

Interference level in the base receiver in idle time slots.

A variety of handover algorithms based on these data can be used to determine the handover instant. Some hysteresis is desirable so that repeated handovers back and forth can be avoided in marginal coverage areas.

Power control in the base and terminal transmitters reduces the average interference level due to co-channel interference. This gives better average speech quality, especially if frequency hopping is employed. Frequency hopping implies averaging of the C/I over several channels and, consequently, the cell planning does not have to be based on the worst case of interference.

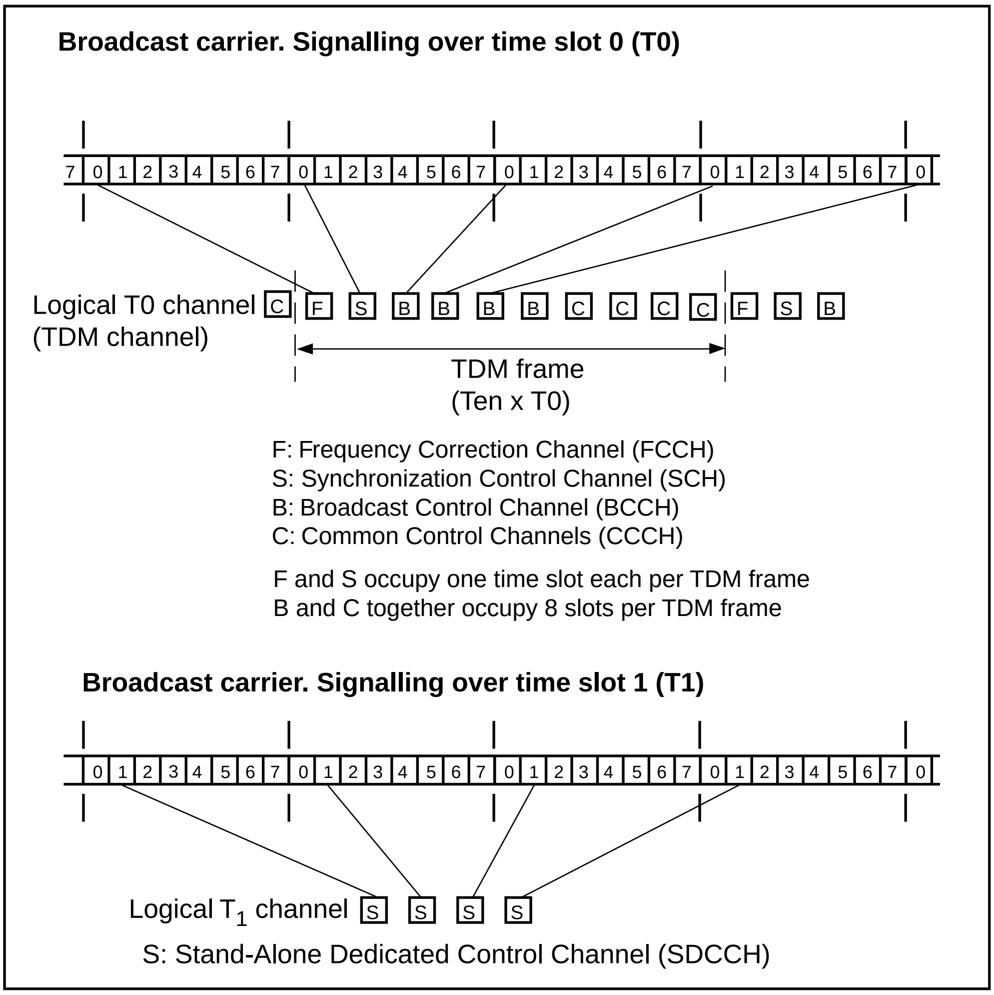
Power-control decisions are based on the carrier level and transmission quality at the terminal and base. If these values are unnecessarily high, power reduction will be initiated at the opposite end of the link. On the other hand if the carrier level or connection quality is too low, the transmitter power will be increased.

Discontinuous transmission and reception are used. Discontinuous transmission means that transmission is stopped during pauses in speech. This requires reliable detection by a voice activity detector (VAD) of gaps in incoming speech signal. The advantages of discontinuous transmission are lower average levels of co-channel interference and lower power consumption by the terminals. One drawback , however, is that totally silent intervals are perceived as disturbing. It is therefore necessary to generate a rough approximation of the background noise level during pauses in speech (comfort noise). Because the data rate required to describe the background noise level is low, the capacity of the SACCH is adequate for this.

With **discontinuous reception**, if the transmission quality is too low during a 20-ms speech frame, output of the detected signal is suspended. If it is only a single, isolated speech frame that is disrupted, the previous frame will be repeated. However, if several consecutive speech frames are disrupted, several repetitions of a previous frame will cause strong quality degradation. For each repetition after the first one, the output level is progressively reduced down to 0, and comfort noise is inserted instead.

3.2.2. Adapting a terminal to the radio environment in a new cell

One radio carrier (the broadcast carrier) at each base station is allocated for system signalling. This radio channel is excluded from any frequency hopping and carries a continuous signal at a constant transmission power level. Time slot 0 in the TDMA frame is used in time multiplex for several different signalling channels. The TDM structure includes ten number 0 time slots, and this is repeated five times in a B multiframe, which comprises 51 TDMA frames. One time slot in the TDM frame is occupied by the FCCH, one by the SCH and the others are divided between the BCCH and the CCCH (see Fig. 3.3).



Figur 3.3

When a terminal has no stored data on the radio environment at the time it is switched on, the first step it takes is to scan all GSM carriers (124 available in the first GSM specification) and to record the field strength. The terminal then returns to the strongest channel to check if it is a broadcast carrier. If not, the terminal checks the next strongest channel and so on. This procedure enables the terminal to find the strongest broadcast carrier and, hence, it has preliminarily selected the best cell, to be connected to.

The terminal then performs fine-tuning of the frequency and synchronization to the base TDMA structure by first receiving the FCCH (Frequency Correction Channel) and, thereafter, the SCH (Synchronization Channel). In addition to synchronization data with respect to the hyper TDMA structure, the SCH also transmits the Base Station Identity Code (BSIC), which consists of the network code and the base station's colour code. (If the terminal were to seize a carrier belonging to another operator, it would have to scan for a new broadcast carrier.) The terminal then receives information on the BCCH. The BCCH continuously transmits information on the identity of the network (operator) and cell, and on the channel allocations for the cell and the broadcast carriers for the six adjacent cells.

3.2.3. Measuring the signal levels from adjacent cells

The terminal will have received information over BCCH from the base station to which it is connected in respect of the frequencies for the broadcast carriers from the adjacent cells. As described in section 2, there is time during each TDMA frame for the terminal to measure the level of a carrier from an adjacent cell. To ensure that readings are reliable, averaging must be carried out over the rapid fading. Therefore, several measurements are made of each carrier before the mean values are sent over the SACCH to the base. The terminal also needs to identify the cell from which the carrier is being transmitted (during extreme propagation conditions, the carrier from a remote cell may be strongest).

Identification of the measured carrier is done by noting the BSIC transmitted on the SCH, which is placed in time slot 0 on the broadcast carrier. The terminal measures the identity of a carrier during a time slot in the last, idle TDMA frame in the multiframe. (This applies to full-rate traffic channels. On half-rate channels, one of the free time slots allocated to the other half-rate channel is used.)

A complication here is that the base stations may not be mutually synchronized. This means that, to start with, the terminal has to listen to the BCCH on the broadcast carrier from an adjacent cell for an entire TDMA frame to be certain of capturing a 0 time slot. This alone is not enough, however, since the 0 time slot is time multiplexed between several signalling channels. The SCH uses only one in ten of the 0 time slots.

For this reason, two different multiframes are used. The inward and outward traffic channels use multiframe A (26 basic TDMA frames), whereas the broadcast channels use multiframe B (51 frames). This means that the idle TDMA frame that is used for listening will slide over the TDMA frames in multiframe B, thus ensuring that, after a number of A multiframes, the terminal will have reached the correct 0 time slot in the broadcast channel of an adjacent cell.

3.2.4. Structure of signalling frames

The structure of some of the signalling frames mentioned above is shown in Fig. 3.4. The **frequency-correction frame** (FCCH) consists of an unmodulated carrier pulse with a given small frequency shift relative to the nominal carrier frequency. The **synchronization burst** (SCH) has a longer synchronization sequence than the normal TDMA frame, to ensure that very reliable time synchronization is achieved. The **paging burst** (RACH) from the terminal to the base is shorter than the others, because the terminal may not yet been informed by the base on the suitable timing advance which is needed to compensate for different propagation delays. To prevent paging bursts from remote terminals from spilling over into the next time slot, a much larger guard slot than normal is needed.

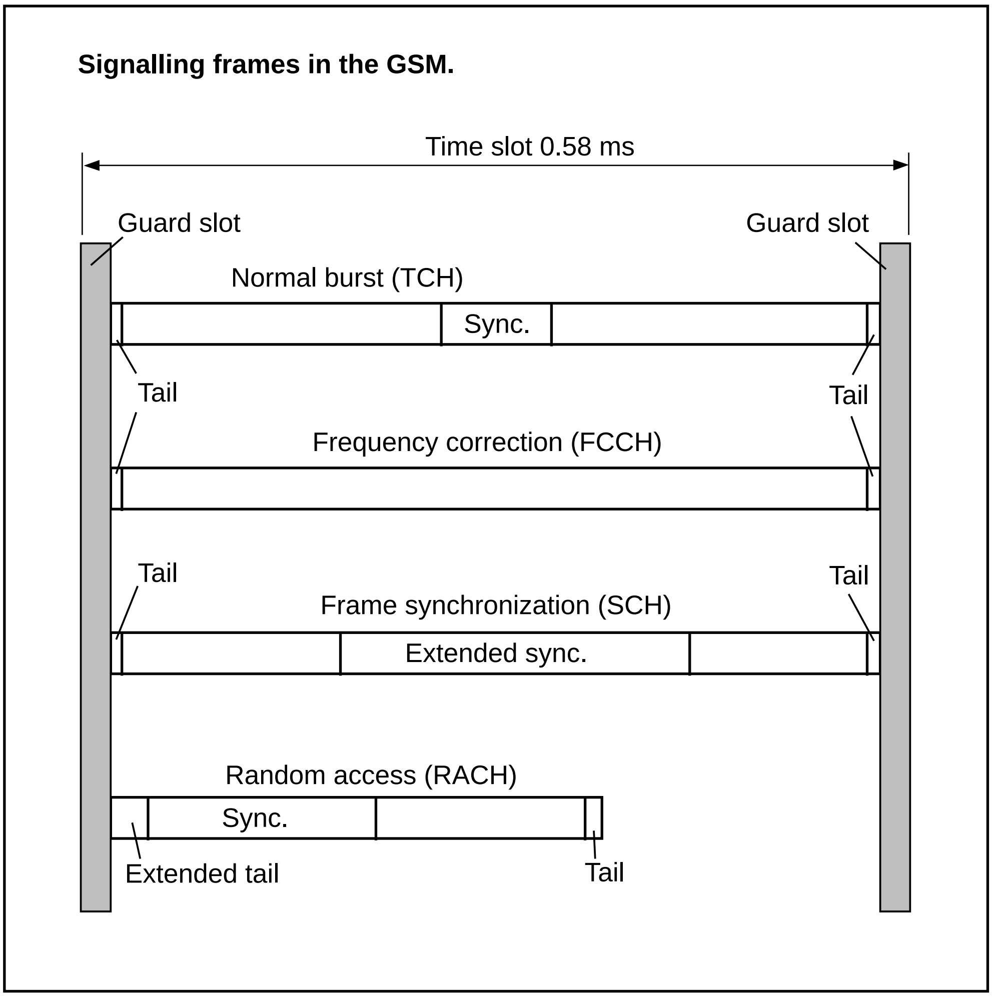


Fig. 3.4

3.3. Channel coding and interleaving

3.3.1. Traffic channels

Traffic channels are either full-rate or half-rate, and both can be used for speech or data channels with different data rates. It would take too long to look at all the cases but three types of traffic channels are discussed below:

a) Full-rate traffic channel for speech (TCH/FS)

b) Full-rate traffic channel for 4.8 kb/s (TCH/F4.8)

c) Half-rate traffic channel for 4.8 kb/s (TCH/H4.8)

As mentioned above, the structure of a full-rate traffic channel corresponds to the transmission of 26 TDMA frames in 120 ms. Twenty-four of the frames are used for traffic, which means that on average 200 traffic frames per second are transmitted (one time slot every five milliseconds). The time for one time slot in a TDMA frame is 0.577 ms, during which time 114 (2 x 57) bits are sent. In the case of a half-rate traffic channel, 12 traffic frames are sent in 120 ms, i.e. 100 time slots per second. Thus, a full-rate traffic channel transmits 22.8 x 10^3 b/s (200 x 114), and a half-rate traffic channel 11.4 kb/s.

3.3.2. Speech transmission (TCH/FS)

The speech coder in the first-generation GSM outputs a binary signal in the form of sequences of 260-bits at a rate of 50 blocks per second (20 ms per block). The bits are classified according to their sensitivity to transmission errors.

Class Ia includes bits for which transmission errors result in a strongly disrupted output signal from the speech decoder. If transmission errors occur in the class Ia group (despite FEC channel coding), the 20 ms frame will be replaced by the preceding frame from the speech coder ("frame erasure"). To enable error detection, three parity bits are inserted into the class Ia group (Cyclic Redundancy Check, CRC).

Assigned to class Ib are bits for which transmission errors result in fairly large degradation of the speech quality. Class Ia (including the parity bits) and class Ib bits are then combined with four tail bits (since a convolutional code with constraint length of 5 is used). This gives a block comprising a total of 189 bits, which is coded in a 1/2 rate convolutional coder. The coder outputs 378 bits (2 x 189).

The remaining bits from the speech coder are assigned to class II. Because these are relatively non-critical as regards the impact of transmission errors on speech quality, they are not given any channel protection through channel coding. Thus, the channel-coder arrangement outputs a total 456 bits every 20 ms, which matches to the capacity of a full-rate channel (see Fig. 3.5)

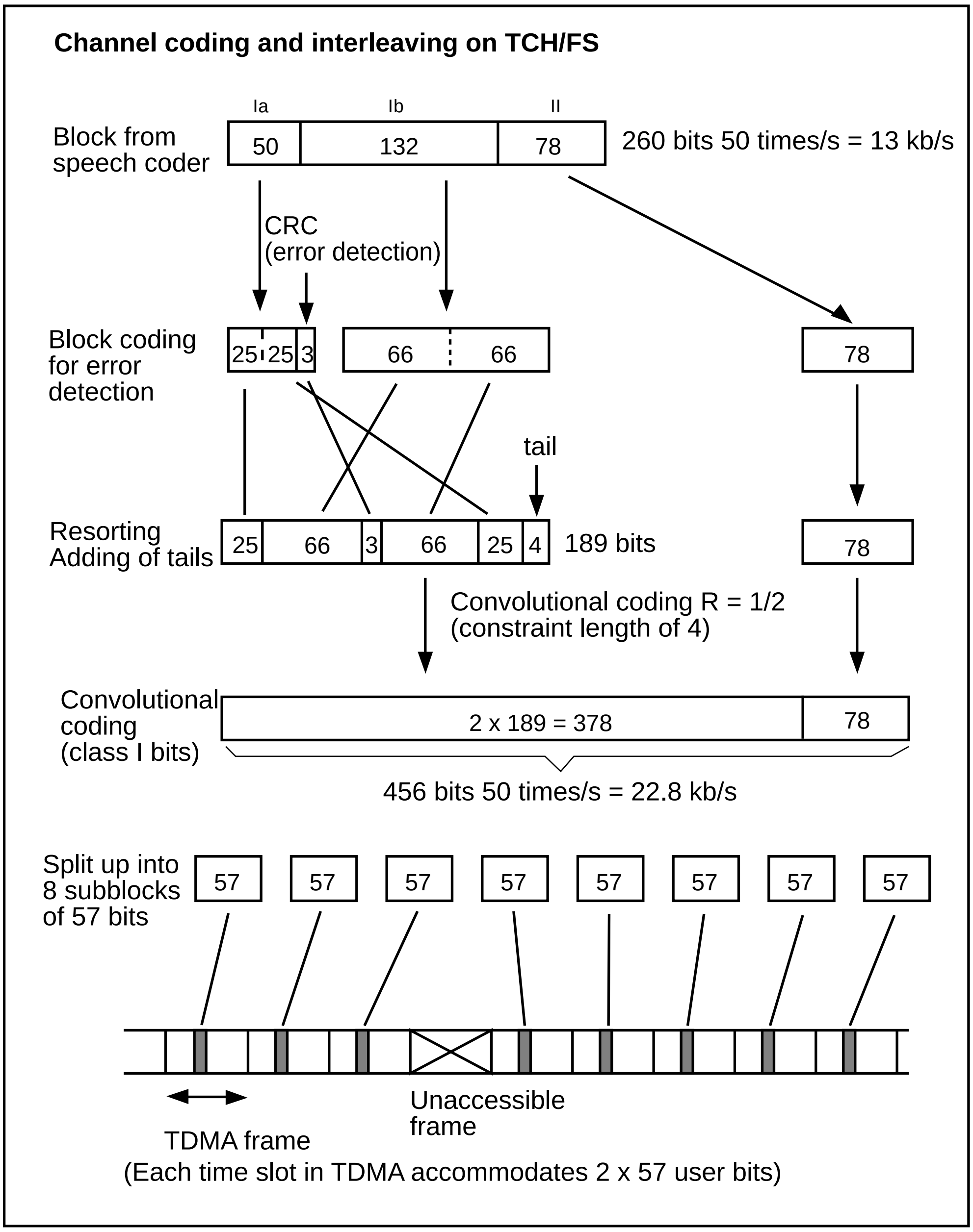


Fig. 3.5

The reason for adding four tail bits is as follows. In a convolutional code, which in this case is decoded by a Viterbi arrangement, the decoding is based on calculation of the Euclidean distance between the received signal and different paths through the trellis. Apart from the current radio symbol, several previous and subsequent symbols contribute to the distance. Therefore, if there is a sudden break in transmission after the last information symbol in the block, the trailing symbols are given less decoding information and, consequently, have a higher error rate than the other bits (see also Module DT12).

The error rate can be substantially reduced if enough tail bits are added to reset the memory cells in the channel coder to zero. In this way, a known final state in the trellis is obtained in the same way that there is a known starting state. Because the coding can be based on known start and end states, the number of paths through the trellis will be lower at the start and end. This means that the bit error rate will be lower for the first and last bits in the block input to the channel coder (see Fig. 3.6). This is exploited by placing class Ia bits at the beginning and end of the word.

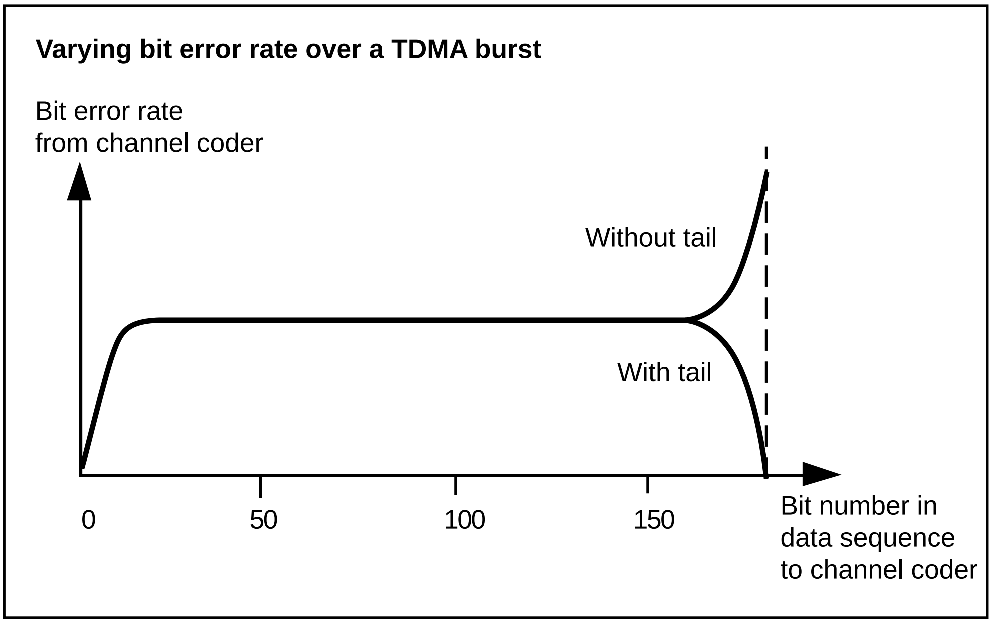


Fig. 3.6

A block from the channel coder comprising 456 bits is then divided into 8 subblocks of 57 bits each. These are inserted into the allocated time slot (physical channel) in eight successive TDMA traffic frames by an interleaving arrangement. (Besides 24 traffic frames, a multiframe also contains one signalling frame and one idle frame. The non-traffic frames are skipped.) Since there are 114 user bits in each time slot, each time slot will therefore contain information from two adjacent speech frames. The interleaving and TDMA frame arrangement creates a time delay of about 40 ms, which is the maximum interleaving delay permissible for speech. (Additional delays in the speech and channel coders bring the total transmission delay up to about 75 ms).

3.3.3. Transmission of 4.8-kb/s data on TCH/F

Apart from the data signal itself, 1.2 kb/s is needed for synchronization and addressing in the fixed network. Accordingly, 6 kb/s need to be transmitted. The incoming data signal is split up in blocks of 120 bits every 20 ms. Prior to convolutional coding at a rate of 1/3 and a constraint length of 5, waists of 4 bits is inserted between subblocks comprising 15 bits (see Fig. 3.7). In addition, a 4-bit tail is added. The waist arrangement reduces the errors in nearby bits in the same way as described for tails. The waist and tail extend the block to 152 bits. The number of bits obtained after channel coding is 456, which corresponds to a gross data rate on the traffic channel of 22.8 kb/s (50 x 456).

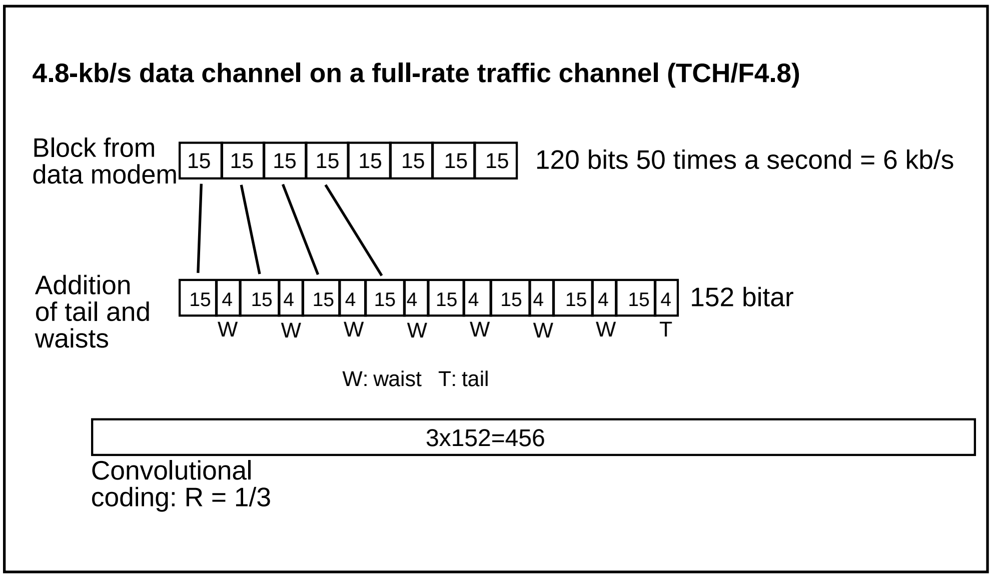


Fig. 3.7

An interleaving arrangement is introduced after the channel coding. The output from the channel coder is 456 bits, 50 times a second. These are divided into 19 words of 24 bits, and the words are distributed among the time slots allocated to the traffic channel in 19 adjacent traffic frames (the signalling frame and idle frames in each multiframe are skipped). This results in an interleaving delay of more than 100 ms. A longer delay is permitted in data transmission than in two-way speech.

3.3.4. Transmission of 4.8-kb/s data on TCH/H

In the transmission of 6 kb/s over a half-rate traffic channel, the margin available for channel coding is much smaller than in the last case. A half-rate traffic channel accepts 456 bits every 40 ms, and 6 kb/s corresponds to 240 bits being transferred 25 times a second. Channel coding consists of the following steps (see Fig. 3.8):

- Addition of 4 tail bits to the input block of 240 bits
  (to fit a convolutional code with a constraint length of 5)
- Convolutional code with R = 1/2 generates 488 bits (2 x 244)
- 32 bits are removed by puncturing, leaving 456 bits

Interleaving involves dividing a block of 456 bits into 16 subblocks, which are placed in 16 consecutive traffic time slots in the half-rate traffic channel (signalling frames are skipped).

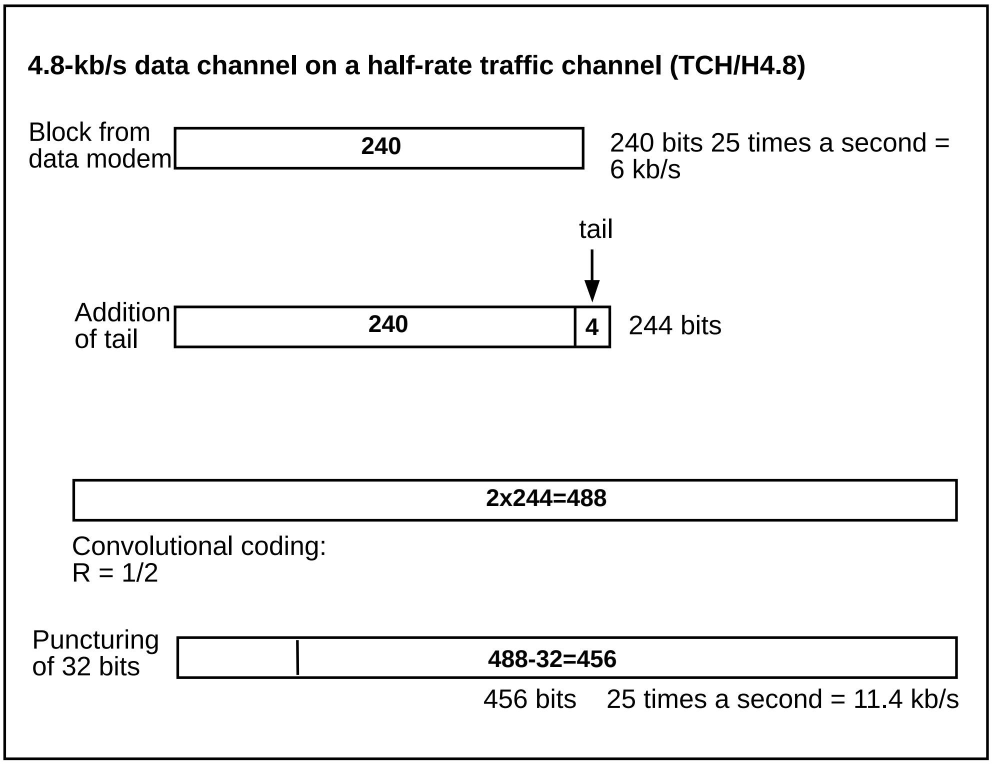


Fig. 3.8

3.4. Radio modem

The GSM uses GMSK modulation with $BT_b = 0.3$, where B is the gaussian filter's 3-dB bandwidth and $1/T_b$ is the data rate of the input signal to the modulator. The specification stipulates that the modulation spectrum shall be attenuated by 30 dB at a 200-kHz distance from the carrier frequency. Filtering of the modulation spectrum in a filter having a Gaussian characteristic and $BT_b = 0.3$ has little effect on the main part of the spectrum (see Fig. 3.9). This means that the filtering has little effect on the waveform. Apart from a considerably steeper spectrum flank, the characteristics of GMSK is very close to MSK. We shall therefore study the simpler MSK type of modulation.

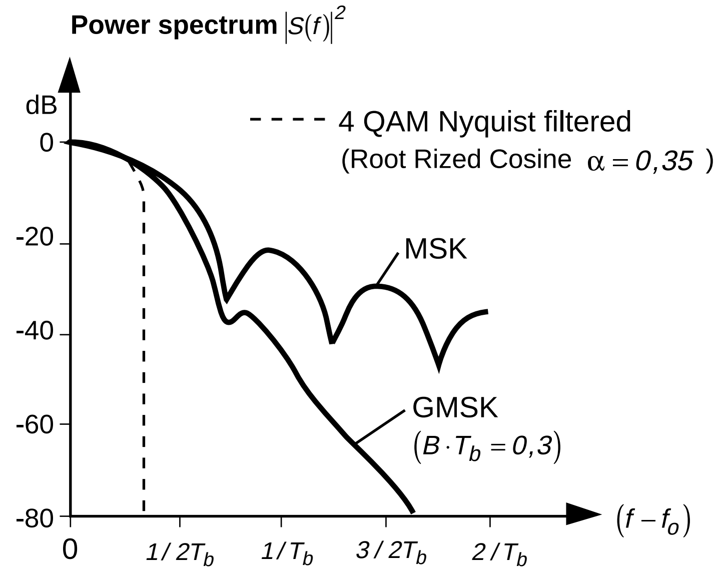


Fig. 3.9

MSK has been described in Module DT4 either as a special variant of 4-QAM or as ortogonal 2-FSK with phase continuity in the transitions between successive symbols. One symbol in the FSK symbol pair is displaced by $+\Delta\omega$ relative to an suppressed carrier frequency, and the other by $-\Delta\omega$. $\Delta\omega$ is the lowest value giving an orthogonal symbol pair. On practical implementation, the close relationship to QAM is utilized in the modulator and detector.

The first point to note is that since MSK can be interpreted as a special variant of 4-QAM, the detector characteristic will be identical to that in normal QAM. This, in turn, results in the same receiver sensitivity as in antipodal signalling, e.g. 2-ASK. The bit error rate for an optimum, matched MSK receiver with a constant signal level (i.e. no fading) will be: $p_b = Q\left(\sqrt{2E_b / N_o}\right)$

MSK differs from normal Nyquist-filtered 4-QAM in the following characteristics:

- The transmitter filter does not have a Nyquist-related transfer function. To obtain constant-envelope MSK-type modulation, a transmitter filter is used whose impulse response, $h_t(t)$, comprises a half sine period of duration 2 T_b.

  $$h_t(t) = \sin\left(\pi \cdot \frac{t}{2T_b}\right)$$

  - The excitations of the transmitter filters in the I and Q channels are mutually displaced by one bit period, T_b, which corresponds to incoming bits being fed alternately to the I and Q channels. This gives O-QAM, which also helps to give constant envelope. (DT3, fig. 16).

- The polarity of the antipodal symbols in the I and Q channels is determined by the value (polarity) both of the input bit and of the previous bit to the modulator. This correlation between the current and the immediately preceding symbol complicates the channel equalization. It is therefore eliminated by suitable precoding.

The modulator for the **normal version of 4-QAM** is outlined in Fig. 3.10.

4-QAM basic version. Modulator.

Transmitter filter related to Nyquist filtering, e.g. Root Rized Cosine

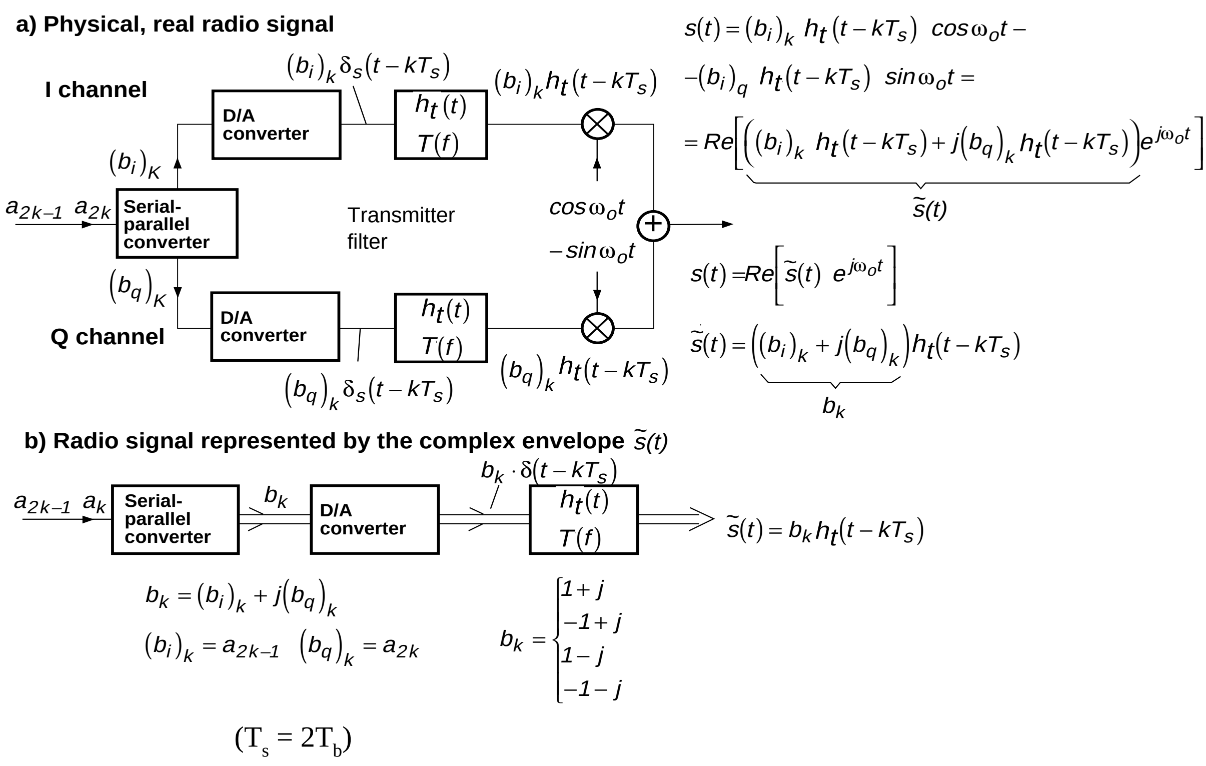


Fig. 3.10

Two bits at a time are input to a serial-parallel converter, which, at the same time, outputs a pair of binary symbols that determine the excitation of the I and Q channels. The transmitter filter is related to Nyquist filtering in order to obtain the narrowest modulation spectrum. In a) in the figure, the physical, real radio signal, *s(t)* is generated. In b), the complex envelope, $\tilde{s}(t)$, is generated instead. The relationship between *s(t)* and $\tilde{s}(t)$ is:

$$s(t) = Re\left[\tilde{s}(t)e^{j\omega_o t}\right]$$

In b), the I and Q channels have been combined in a complex channel, where the I channel constitutes the real part and the Q channel the imaginary part.

An incomplete model of an MSK modulator, in which MSK is interpreted as **Offset QAM** (OQAM), is shown in Fig. 3.11. Excitation of the I and Q channels takes place at different times, and the impulse response for the transmitter filter consists of a half-sine period. However, the model does not include the correlation between successive symbols, which is related to the ”±” sign in front of the a_k coefficients. Only two successive symbols from the continuous stream of symbols generated are shown in the figure.

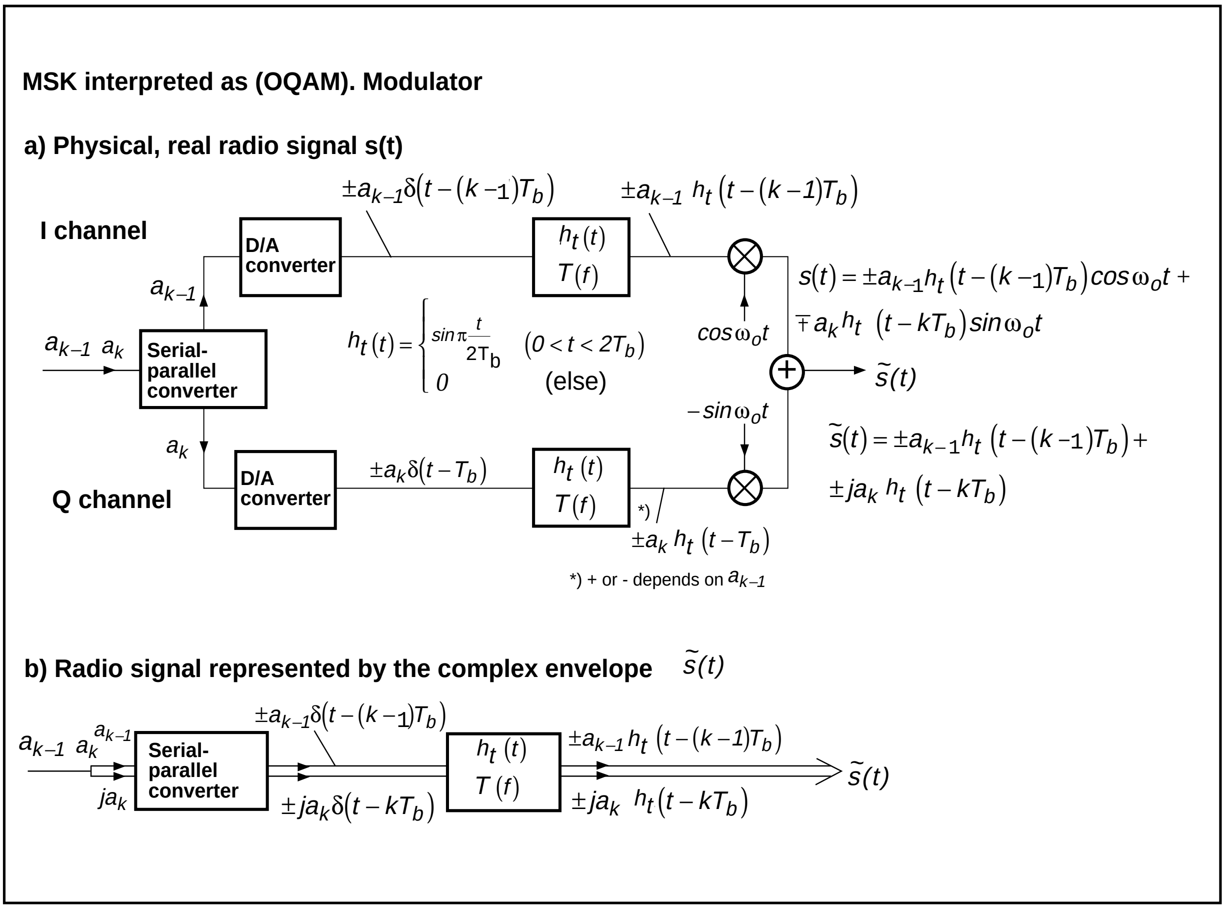


Fig. 3.11

The demodulator for MSK is outlined in Fig. 3.12. The MSK-modulated input signal, s(t) is divided into an I signal and a Q signal. These are filtered through the receiver filter, R, and then sampled in the normal QAM procedure. The diagram can be simplified by having the complex envelope represent the radio signal and using a complex baseband signal, as in b). The received symbols will then be alternately real and imaginary. The receiver filter needs to be matched to the transmitter filter, i.e. $h_r(t) = h_t(-t)$, for the best receiver sensitivity (lowest E_b/N_0) (Strictly, $h_r(t) = h_t(T-t)$, where T shall be large enough so that the causal condition is satisfied.) .

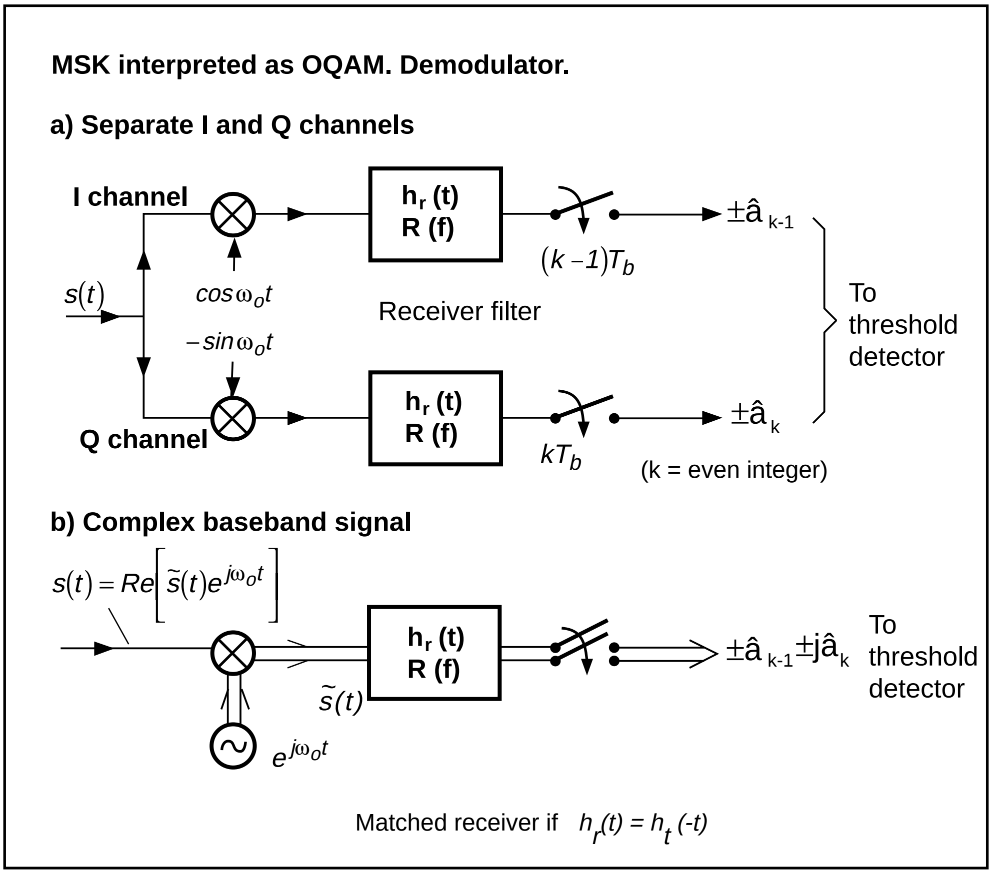


Fig. 3.12

To get a complete diagram for the MSK modulator, the correlation between adjacent symbols must be included. As can be seen in Fig. 13 in Module DT4, the polarity of the antipodal Q symbol generated is determined both by whether $+\Delta\omega$ or $-\Delta\omega$ is to be generated (i.e. whether the current input symbol is +1 or -1) and by the polarity of the preceding symbol on the I channel. If $+\Delta\omega$ is to be generated, the polarity of the current symbol and the preceding symbol must be opposite. If they have the same polarity, $-\Delta\omega$ will be generated. The same applies to the choice of polarity for a symbol on the I channel.

These relationships, together with the fact that alternate symbols must be real (input to the I channel) and imaginary, lead to the complete diagram for the MSK modulator shown in Fig. 3.13, a. The symbol, b_n, which determines the polarity of an output I or Q symbol, is obtained by multiplication of the preceding symbol, b_{n-1}, by ja_n, where a_n is the bipolar bit input to the modulator ($a_n=\pm1$).

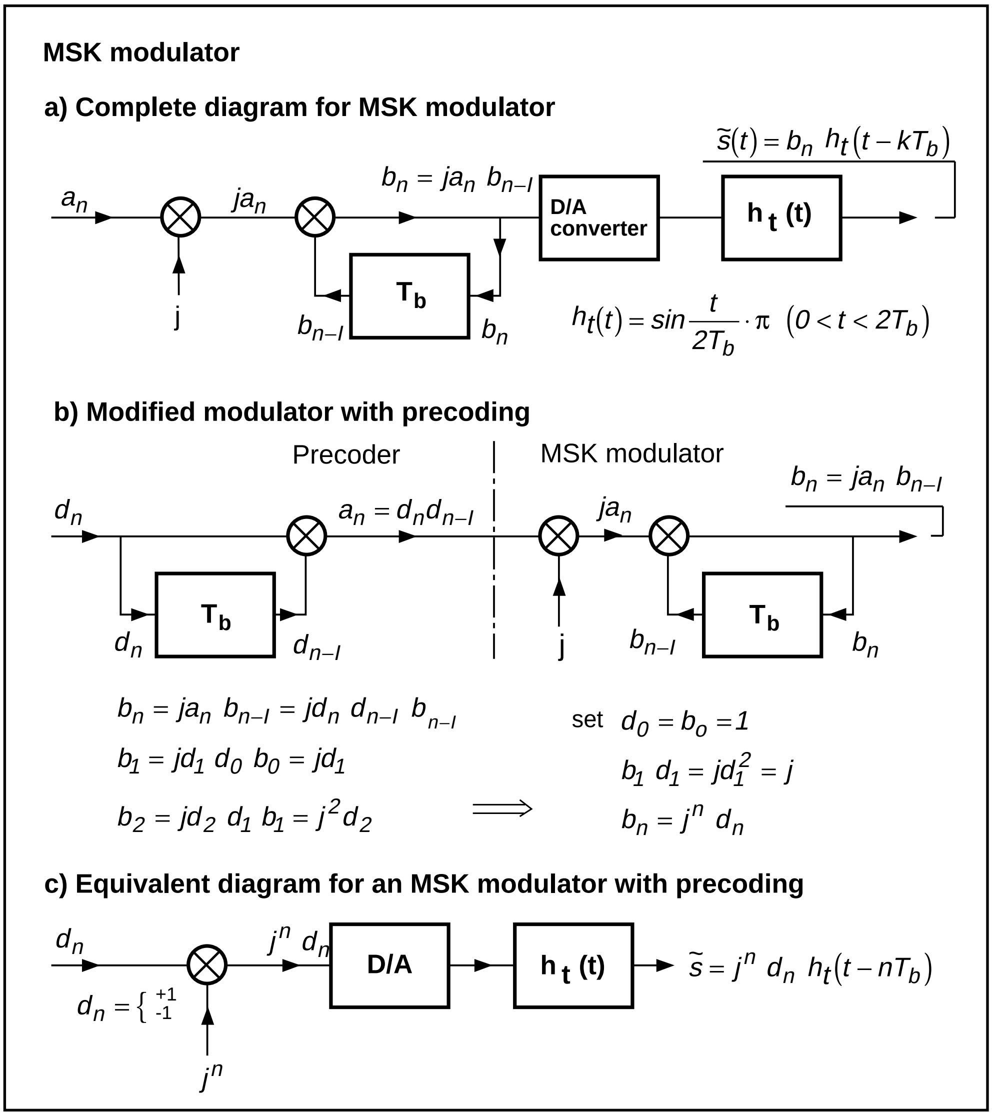


Fig. 3.13

To simplify channel equalization, the correlation between adjacent I and Q symbols should be removed by precoding (see Fig 3.13, b). Precoding involves multiplying the current input bit, d_n, to the precoder by the preceding input bit, (d_{n-1}). As shown in the figure, the relationship between symbol b_n, which determines the excitation of the transmitter filter, and the input baseband bit, d_n, is:

$$b_n = j^n d_n$$

The factor, j^n, results in the b symbols beeing fed alternately to the I and Q channels. A simplified equivalent diagram for an MSK modulator with precoding is given in c) in Fig. 3.13.

To further facilitate channel equalization, the receiver should also be modified so that the rotation $\pi/2$ in the I-Q plane for each symbol is removed from the input signal to the channel equalizer. This is achieved by multiplication of the complex baseband signal from the sampler by $-j^n$ (see Fig. 3.14).

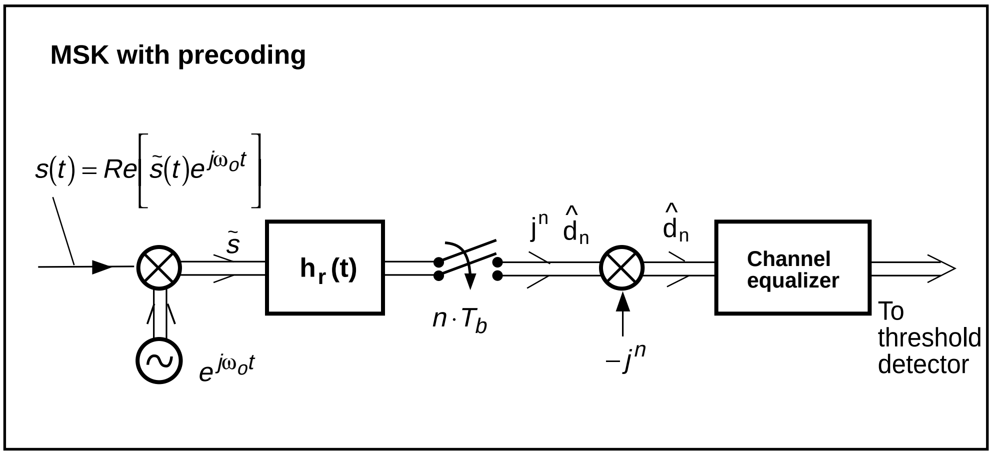


Fig. 3.14

After these modifications, we get the equivalent QAM transmission diagram for analysis of the channel equalizer as shown in Fig. 3.15

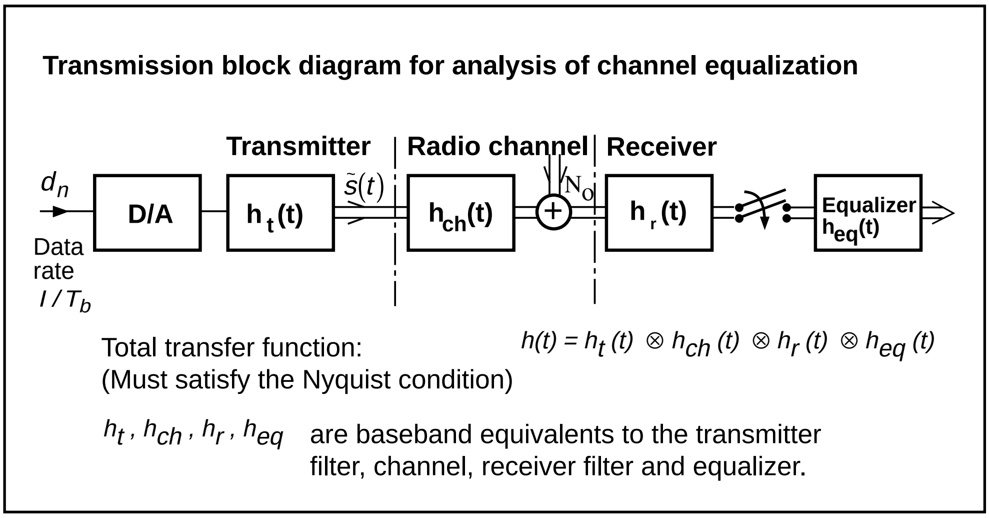


Fig. 3.15

3.5. Channel equalization

The system data rate of the GSM is 271 kb/s, i.e. the duration of a symbol is 3.7 µs. The specification for the GSM stipulates that the system must be able to handle time dispersion up to 16 µs (16 µs equalization window), which corresponds to four symbols. The transmission performance is influenced by the sampled impulse response of the radio channel. Time dispersion over 16 µs implies that the impulse response of the channel can cover five sampling points (see Fig. 3.16). This is the time discrete impulse response to be processed in the channel equalizer.

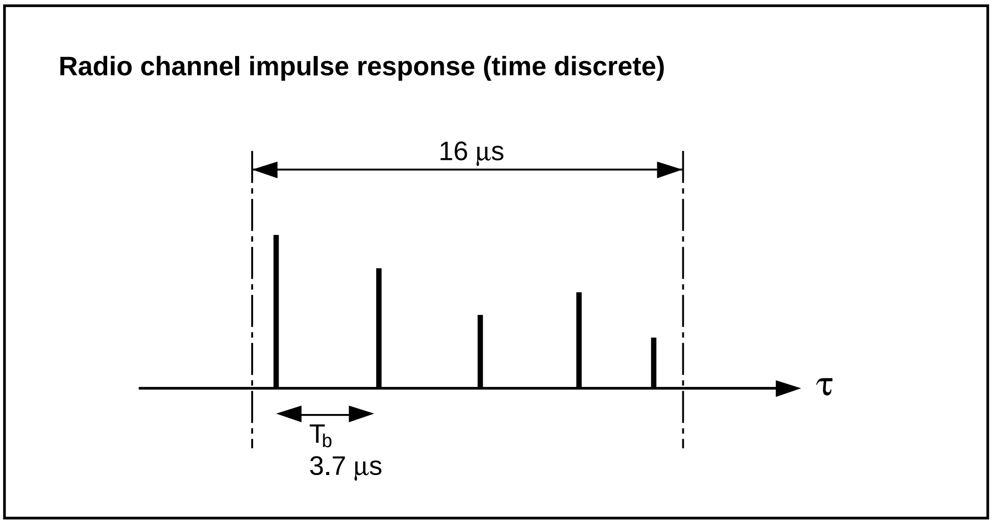


Fig. 3.16

Because of the relatively short data burst in a TDMA time slot (0.55 ms), the variation in the impulse response of the radio channel during this interval can nearly be disregarded, especially as the training sequence is placed in the middle of the burst. Thus, the setting of the channel equalizer that is determined by the training sequence applies to the entire data burst. (The rapid changes in the channel's impulse response are due to independent fading of the signal components with different propagation delay times. The distance between two fading dips is at least half the wavelength, i.e. approx. 0.15 m. For a terminal travelling at 90 km/h (25 m/s), there will therefore be 17 fading dips a second. One complication in going over from 900 MHz to 1800 MHz is doubled doppler frequency for a certain terminal velocity. This results in a reduction in the correlation time, i.e. larger variations of the impulse response during one data burst. (The time correlation function is the Fourier transform of the doppler spectrum.)

Two types of channel equalizers have been used:

a) Combination of linear equalization and decision feedback

b) Maximum Likelihood Sequence Estimation (MLSE) based on a Viterbi arrangement.

Although the two options are fairly comparable in terms of performance and complexity, the most widely used is the Viterbi arrangement.

The result of a simulation of the performance using a simple two-ray model of the time-dispersive radio channel is shown in Fig. 3.17.

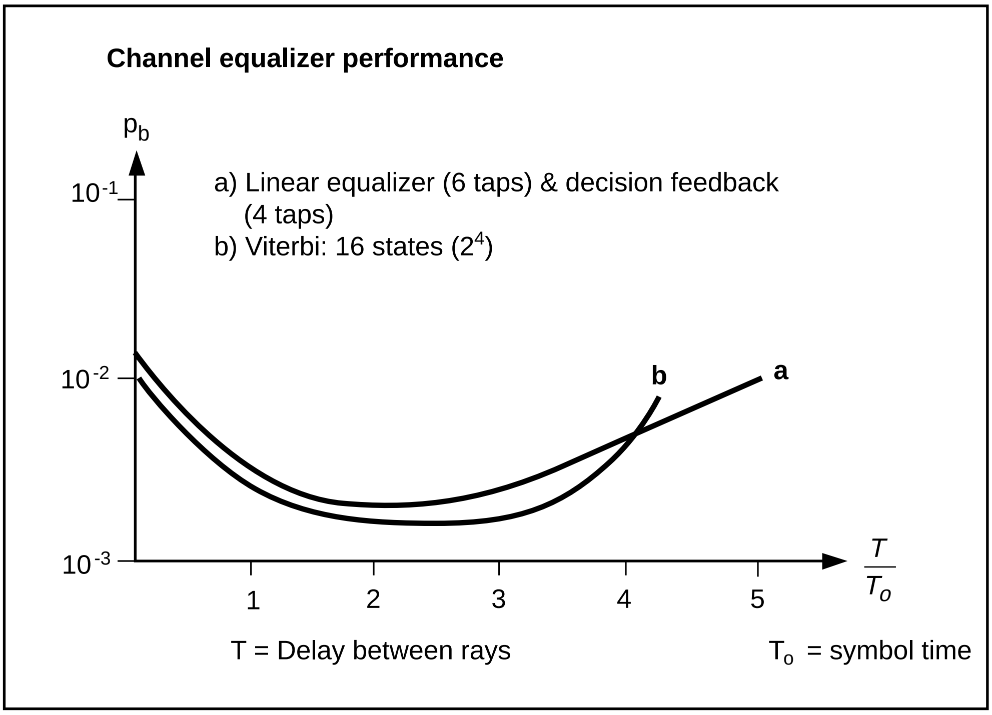


Fig. 3.17

No diversity gain is obtained when there is a zero delay between the two propagation paths. As the delay, T, increases, there is an initial fall in the bit error rate, since the channel equalizer can distinguish between the two propagation channels with independent Rayleigh fading and combine the signals so that diversity is obtained. When the delay is more than four symbol periods, the time dispersion window that the equalizer can handle is exceeded. Any further increase in the time difference between the two rays will result in a rapid deterioration in performance.

In **MLSE**, the demodulator determines the most probable input sequence (maximum likelihood) considering the received sequence and the measured impulse response of the propagation channel. The most likely transmitted sequence is that having the shortest Euclidean distance to the received sequence. The comparations are based on simulated sequences which are determined both by the transmitted sequences and the time dispersion (impulse response).

(If the measured impulse response coincides exactly with the real one, and the noise and interference are negligible, then the distance between the received output sequence and that simulated in the receiver corresponding to the transmitted sequence will obviously be nil.)

In the comparison procedure all possible transmitted sequences must pass through a network that simulates the total radio channel (modulator with transmitter filter, propagation channel with time dispersion, and modulator with receiver filter). The principle of this arrangement is shown in Fig. 3.18.

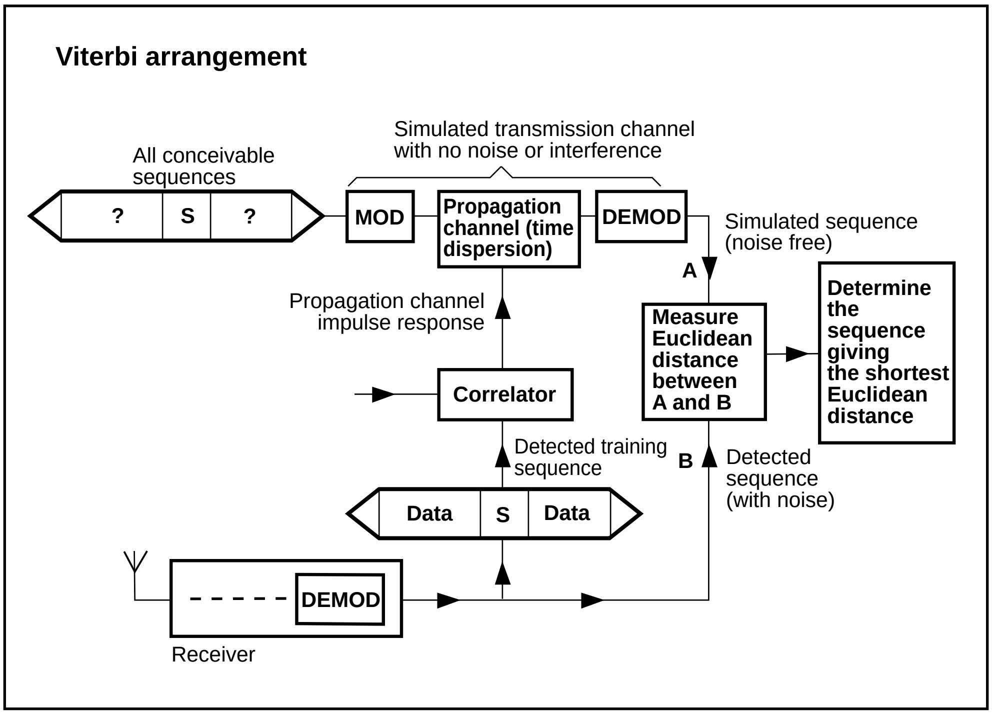


Fig. 3.18

In principle, in the MLSE comparison all possible transmitted sequences must be generated in the receiver and the distance to the received sequence determined. However, this is evidently not possible, as the number of sequences is almost infinite (2^{114}, since a data burst contains 114 user bits). Nonetheless, the comparison can be made step by step using the Viterbi procedure. Each step involves considering all possible sequences during 16 µs, which corresponds to the time-discrete impulse response, with a maximum width of 5 samples.

The procedure can be interpreted as Viterbi coding of a convolution code. The number of states in the coder is determined by the number of memory cells n. n is the constraint length -1. If there are n memory cells, the number of states will be 2^n (all permutations of a group of n input bits). The constraint length is equal to the extent of the time dispersion. In accordance with the GSM specification, the time dispersion extends over 5 sampling points, which means that the trellis will have 2^4 (16) nodes (states). Each node has two incoming and two outgoing paths, as with MSK there are only two symbols in the alphabet (M = 2).

The complexity of the equalizer increases sharply with an increase in the number of states. During the design of the GSM, it was decided that the equalizer would be too complex if more than 16 states were required. This, therefore, had a decisive influence on the design specification. A trade-off was made between the maximum time dispersion the system could handle and the maximum system data rate. In extreme cases (such as in the Swiss Alps), time dispersion may exceed 16 µs, which means that measures to reduce the time dispersion have to be adopted (such as suitable antenna siting and use of directional antennas).

As well as frame synchronization, the synchronization sequence is also used to measure the impulse response of the radio channel. The sequence consists of a total of 26 bits. The sixteen bits in the middle form a word having perfect, cyclic autocorrelation characteristics.

(Eight different sequences with good, mutual cross-correlation characteristics are available. To reduce the likelihood of strong co-channel interference disturbing the channel equalizer, nearby co-channel cells are assigned different synchronization sequences.)

The autocorrelation function will be zero outside of the peak for zero time difference. This allows accurate measuring of the impulse response of the radio channel, provided that the calculation can be made of one complete period of the 16 bit sequence. For this condition to be met, the training sequence is extended by five bits (the length of the channel's impulse response) on either side of the middle part through repetition (lengthening) of bits taken from the other side of the primary 16-bit sequence.

3.6. The fixed network

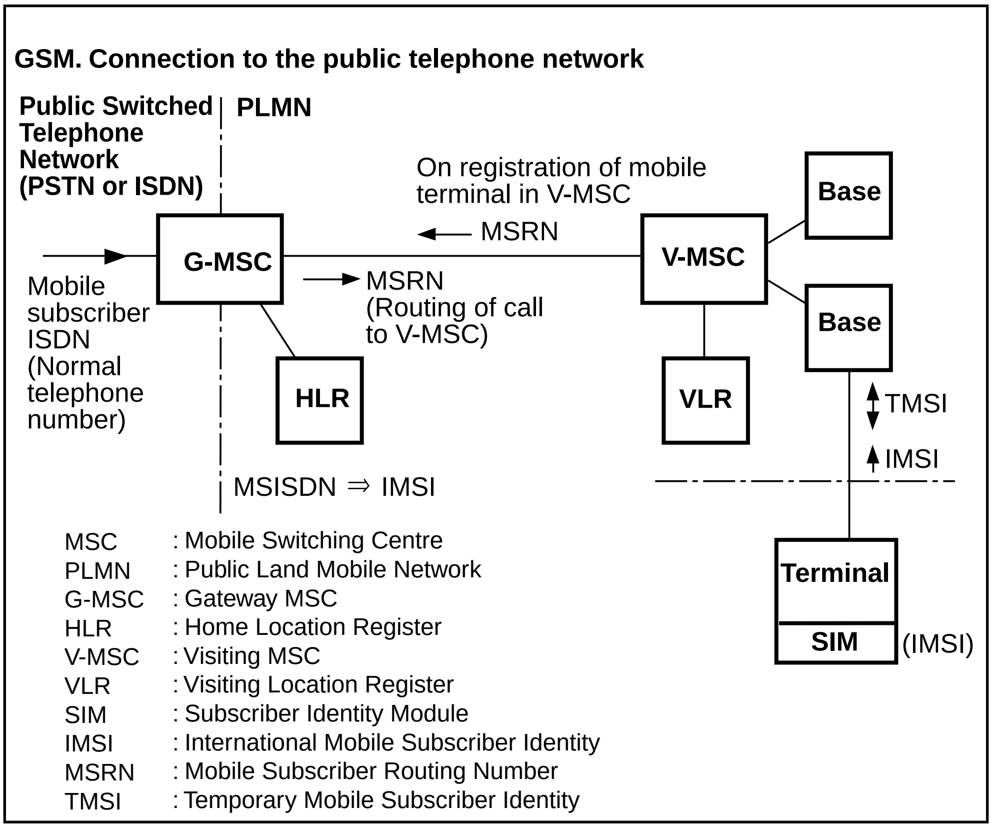
3.6.1. Structure

The Public Land Mobile Networks (PLMN) run by different operators constitute islands in the Public Switched Telephone Network (PSTN). When the PSTN initiates a call to a mobile terminal belonging to a PLMN, the call request is fed to the interface between the PSTN and the PLMN. The interface consists of the operator's Gateway Mobile Switching Centre (G-MSC). Details on all the subscribers belonging to the PLMN are contained in the Home Location Register (HLR) database.

In the simplest case, there will only be one HLR in the mobile network that is connected to the G-MSC. An incoming call is therefore routed direct to the HLR, which contains a host of subscriber data, e.g. the subscriber's mobile network number, which mobile services shall be provided to each subscriber, security codes to prevent unauthorized access to the network, and the location of roamers registered in another MSC with its associated Visiting Location Register (VLR).

3.6.1 Calls to mobile terminals

The first step when a call request is made to a mobile subscriber is to send a data message to the mobile network (operator) serving the subscriber. This initial request follows the standard calling procedure on the PSTN (or ISDN) number. Every mobile subscriber is thus allocated a normal phone number (Mobile station ISDN number) comprising a country code and a network operator code NDC (National Destination Code) plus a subscriber number issued by the operator. By means of the country code and the NDC, the fixed network (e.g. the PSTN or ISDN) establishes a connection with the operator's gateway (G-MSC) (see Fig. 3.19).



Figur 3.19

The MSC transfers the call to the HLR for the called subscriber. Here, the number is replaced by the subscriber’s International Mobile Subscriber Identity (IMSI), which is also stored in the Subscriber Identity Module (SIM) or, if not used, permanently coded into the terminal. The IMSI has the same form as a normal telephone number, i.e. a country code, mobile network code and the subscriber number.

In the case of a non-roaming subscriber, the terminal will be within a location area belonging to the G-MSC (assuming that the operator uses only one HLR), in which case paging can be initiated immediately. If the terminal is roaming within the area covered by the operator, the call request must be processed further within the operator’s PLMN. The HLR associated with the G-MSC will hold information on the V-MSC with which the terminal is currently registered. This information will have been sent to the HLR by the VLR with which the terminal has registered (through the IMSI) and which has allocated an MSRN (Mobile Subscriber Routing Number) to the terminal. This temporary number comprises a country code, a trunk code for the VLR and a temporary subscriber number. Using the MSRN, the G-MSC is able to connect the call to the V-MSC, which then pages the terminal in accordance with the procedure described in section 3.2.1.

3.6.3. Security against unauthorized access

Because a mobile radio link is more vulnerable to listening-in and unauthorized usage than is the public telephone network, the following subsystems features are available:

Encryption of messages and some system signalling associated with a given subscriber. Temporary subscriber numbers are also used for the identification of subscriber during the signalling procedure, to prevent an unauthorized party from being able to locate a subscriber by intercepting the IMSI. The IMSI is replaced by the TMSI (Temporary Mobile Subscriber Identity) as soon as possible in different signalling procedures over radio links.

Authentication. This involves checking that a terminal wishing to make a connection to the network is authorized and has subscribed to the requested service. Encryption procedures are used for transmission of electronic signature (signed response) to prevent unauthorized parties from sending false authorization data.

Checking the (manufacturer's) **serial number** of a terminal unit. The network can ask a subscriber to provide this number so that it can be checked against a database (Mobile Equipment Identity Centre) containing details of equipment reported as stolen or as not functioning properly.

Encryption and authentication are based on matching keys (K_I) stored both in the subscriber SIM and in the authentication database (Authentication Centre, AUC) connected to the subscriber's home switching center (H-MSC) and HLR. In addition to the IMSI, the SIM (Subscriber Identity Module) contains the encryption and authentication keys, and the algorithms for computing the enciphering sequence and the response to the authentication request from the network. The corresponding information on all the subscribers belonging to the HLR are held in the AUC database. The algorithms are the same for all subscribers but the encryption K_I is individual. K_I is the secret keys used by the encryption and authentication system (see also the session-key procedure described in Module S1).

In the terminal the signed response (SRES) sequence to be transmitted for authentication is formed in the terminal. The authentication algorithm combines the key (K_1) with a random sequence (RAND), which is different for each call. The RAND is transmitted to the terminal, which has access via the SIM to K_1 and the authentication algorithm. The terminal can therefore generate the signed response (SRES) and send it to the network. The terminal will only be accepted by the network if the SRES sent as reply by the terminal is identical to the sequence that has been computed by the AUC and stored in the VLR. The VLR can request the SRES from the HLR (see Fig. 3.20).

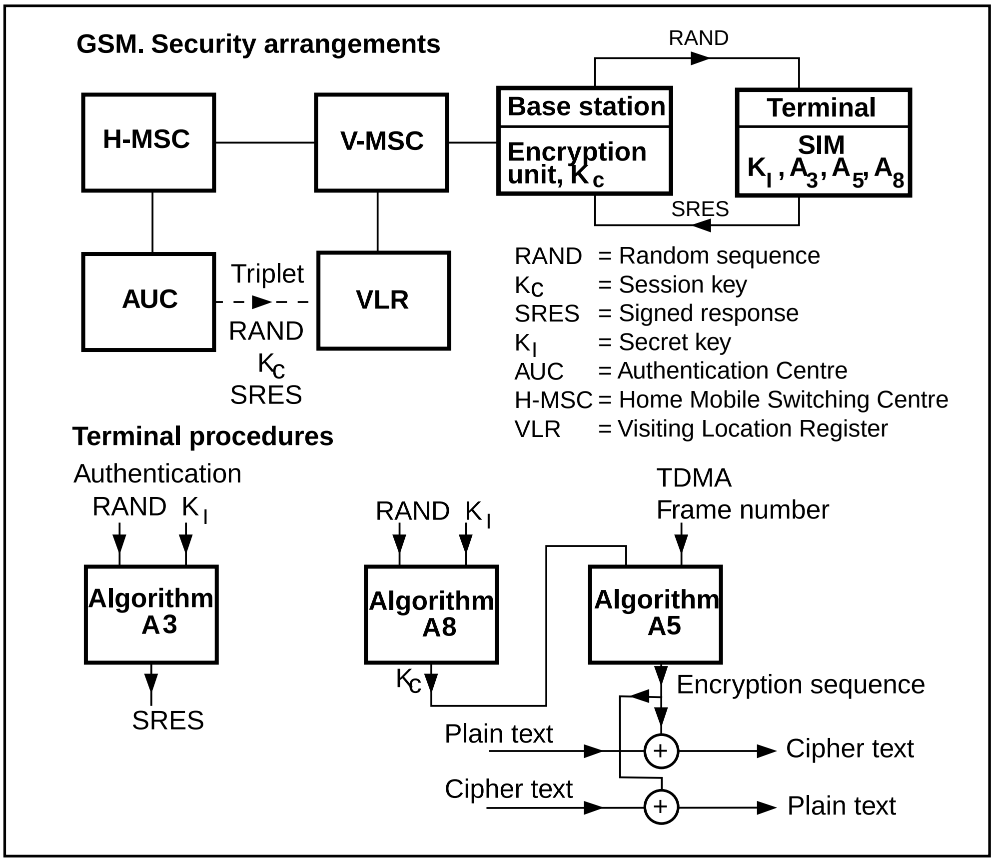


Fig. 3.20

The final encryption sequence that is added to the plain text to produce the cipher text is obtained in two steps. First, the session key, K_c, is obtained by combining the random sequence as above with the encryption key, K_1 (the secret key) using the encryption key algorithm (A8). The final encryption sequence is then obtained by combining K_c with the current frame number (in the hyper frame) using the encryption algorithm (A5).

On the fixed network, the authentication and encryption processes require access to a triplet of parameters: SRES, RAND and K_c. Since the RAND must be different every time it is used, The HLR (with support from the AUC) computes several triplets with different RANDs for each terminal. On receipt of a request from the VLR, the HLR transmits a number of triplets for a roaming terminal. For encryption/decryption of outward and inward traffic, the VLR (or HLR if the terminal is not a roamer) sends the current encryption keys, K_c, to the relevant base stations. The arrangement with different K_c keys for each call means that the only units that know the secrete key K_1 is the AUC and the terminal.

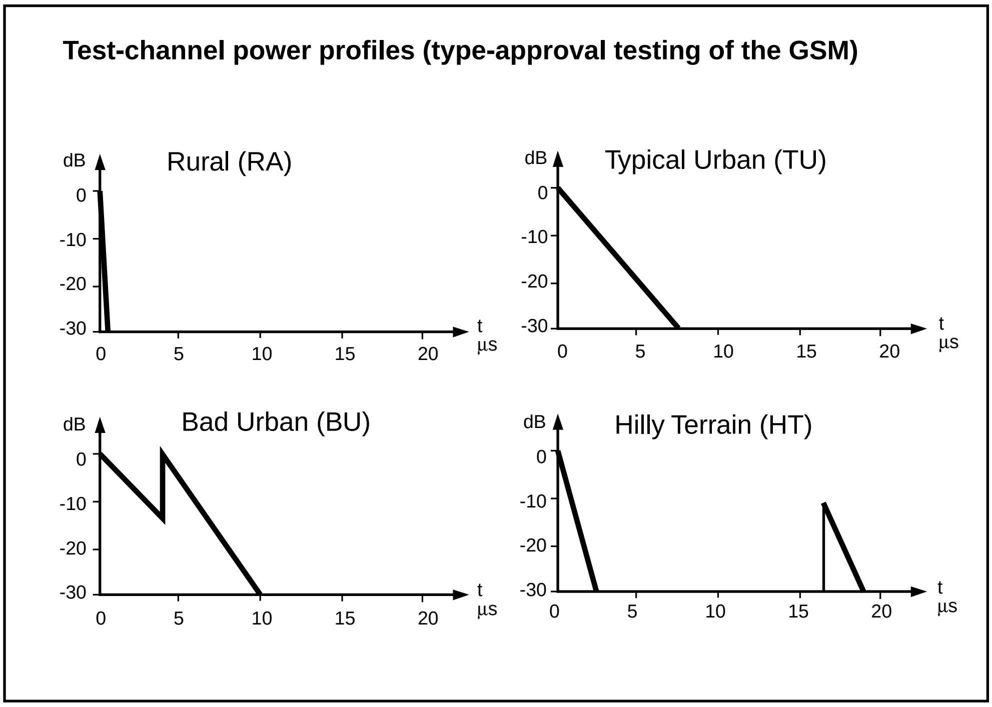
Only the actual data sequence is encrypted – not the training sequence. Encryption and decryption is implemented through by Modulo-2 addition of the identical encryption sequences to the transmitted and received signals.

To prevent stolen or defective equipment being connected to the network, a special database (Mobile Equipment Identity Centre) is set up. The manufacturer gives every terminal unit a serial number (Mobile Equipment Identity), which is stored in a memory that is difficult for unauthorized persons to access and modify. The terminal will transmit this number when requested to do so by the network.

4. Radio performance

The GSM specification stipulates the highest permissible bit error rate for different traffic and signalling channels for different combinations of time dispersion, C/I and C/N. Thus, type approval testing involves a multitude of measurements. Only some of these are dealt with here.

The measurements are made through laboratory tests, in many cases using a fading and time dispersion simulator. The simulator generates a transfer function (impulse response) for the propagation channel with several fading signal components with varying delays. The parameters that describe the rapid fading (Rayleigh or Rice, maximum Doppler frequency) can be set for each subsignal. The channel simulator can be set to a number of standardized time dispersion models (see Fig. 4.1). It is assumed in the majority of test cases that the subsignals with varying propagation times are subject to independent Rayleigh fading. The specification stipulates that the channel equalizer shall be able to cope with time dispersion within a 16-μs window.



Figur 4.1

Type testing includes measurement of the noise-limited sensitivity (transmission performance for E_b/N_0 = 8 dB) and the performance with co-channel interference of (C/I_{co} = 9 dB) and adjacent channel interference of (C/I_{adj} = -9 dB). The type-approval specification gives the permissible degradation for a number of traffic and signalling channels for these three test cases. A few of the test specifications are shown in Fig. 4.2. Other test cases are the bit error rate at the demodulator output (i.e. without the aid of channel coding) at a high C/N (C = -115 dBW), with no interference and with the TU propagation channel.

Type approval test criteria for the GSM

		Noise-limited sensitivity C=-132dBW E_b/N_o =8dB		Co-channel interference C/I_{co} =9dB Adjacent channel interference C/I_{adj}=–9dB			
		TU50 Frequency hopping:		TU3 Frequency hopping:		TU50 Frequency hopping:	
		Without	With	Without	With	Without	With
TCH/FS	FER	3%	2%	21%	3%	6%	3%
	RBER Ib	0.4%	0.3%	2%	0.2%	0.4%	0.2%
	RBER II	8%	8%	4%	8%	8%	8%
TCH/F 4.8	BER	10^{-4}	10^{-4}	3	10^{-4}	10^{-4}	10^{-4}
TCH/F 2.4	BER	$2 \cdot 10^{-4}$	10^{-5}	3	10^{-5}	$3 \cdot 10^{-5}$	10^{-5}
SDCCH	FER	10%	4%	22%	4%	10%	4%

TU50 = Typical Urban, 50 km/h
FER = Frame Erasure Rate
RBER = Residual Bit Error Rate (Ib and II: see section 3.2.2).

Fig. 4.2

Several parameters are related to the speech transmission quality: the residual bit error rate (RBER) is specified for the speech blocks during which error correction for class 1a bits is works satisfactory, i.e. the CRC decoder has verified the block. The frame erasure rate is the number of blocks in which class 1a bits from the convolutional coder contain errors (detected by the CRC).

The performance criteria shown in Fig. 4.2 are confined to those applicable to when the channel simulator is set to the Typical Urban model (TU). Two terminal speeds have been assumed for the cases with co-terminal and adjacent channel interference: 3 and 50 km/h. The criteria are also given for the cases with and without frequency hopping. As is clearly apparent, frequency hopping gives a considerable improvement in performance at the low terminal speed (3 km/h).

5. References

1. R. Steele: Mobile Radio Communication
   Pentech Press, London, 1992

2. M. Rahnema. Overview of the GSM System and Protocol Architecture
   IEEE Communication Magazine, April 1993

Appendix 1.
Evolution of GSM for data services

Content

1. Introduction

The original GSM system was optimized for speech transmission. Due to the rapidly increasing market need for data services, a gradual evolution has taken place. The original specification (phase 1) included a framework that made possible moderate-speed circuit-switched data channels (up to 9.6 kb/s net user data + 2.4 kb/s network signalling) and short-message service based on packet transmission. The detailed specification of these add-on services (phase 2) was completed 1996. The GSM standardization is now in phase 2+. This phase includes improved speech coders and above all advanced data transmission services: high-speed circuit-switched data service (HSCSD) and the general packet radio service (GPRS).

HSCDS allows higher user rates by given a user the option of using several time-slots per TDMA frame.

GPRS is a connection-less packet transmission service, including the multi-slot arrangement from HSCDS. The ETSI standardization activities were completed late 97 and sent out for review during 1998.

The next step, which is still in the study phase, gives a further increase in system data rates by introducing improved, more frequency-economic modulation arrangements. See section 4 about EDGE.

All the extensions are based on the original GSM specification, i.e. the original structure of signalling and hand-over arrangements is used as far as possible and these are used during the call set-up procedure. During a call it is also possible to switch between the original GSM services and the new data services. Some modifications are necessary in connection with HSCDS and GPRS. These changes in the detailed arrangements are not taken up here, mainly the new facilities, needed for connection-less packet transmission, are discussed.

This appendix mainly covers GPRS.

2. HSCSD

The maximum user data rate can be increased by allocating several (n) times slots to one user. At the BSC (Base Station Controller), the incoming data stream is divided up into n independent channels (full rate traffic channels, called HSCSD channels), which are combined in the terminal after channel decoding and ARQ for each channel. The same applies to the other direction. The maximum user data rate is determined by the 64 kb/s links between BSC and MSC. (Presently only one such link is used for each connection.)

The net data rate is influenced by the rate of the channel coding. If fairly large and varying delays can be accepted, ARQ (backward error control) can be introduced to further decrease the error rate (non-transparent service). In principle, each time slot is used for one of the data transmissions modes (TCH/F9.6 or TCH/F4.8) according to the original GSM specification. In addition a raw rate of 14.4 kb/s is introduced by reducing the amount of channel coding. The maximum rate per user over the air interface is nx14.4 kb/s, if one user is allocated n slots. (The total data rate per timeslot, incl. channel coding is 22 kb/s.)

The normal GSM terminals cannot receive and transmit simultaniously. That means that the maximum number of traffic slots that can be allocated in either the inward or outward direction, is limited by the need to:

- include slots to transmit control signals in the other direction
- measure the level of signals from near-by cells (for MAHO)
- introduce guard times to allow for settling times of frequency synthezisers (after frequency change).

The maximum number of slots that can be used in either direction is therefore 4, but the total number of traffic timeslots (for both directions) per basic TDMA frame should not be more than 5.

These restrictions can be eliminated, if full frequency-duplex terminals are introduced and no frequency-hopping is used. In that case one user can be allocated up to 8 time slots (but the total user data rate should not exceed 64 kb/s).

3. GPRS

3.1 Overview

Independent packet routing within the public landmobile network is supported by a new logical network node GSN (GPRS Support Node). The Gateway GSN (GGSN) is a logical interface to external packet data networks. The SGN (Serving SGSN) is responsible for the delivery of packets to the terminals within its service area. A much simplified block diagram of the complete network is shown in figure A1-1.

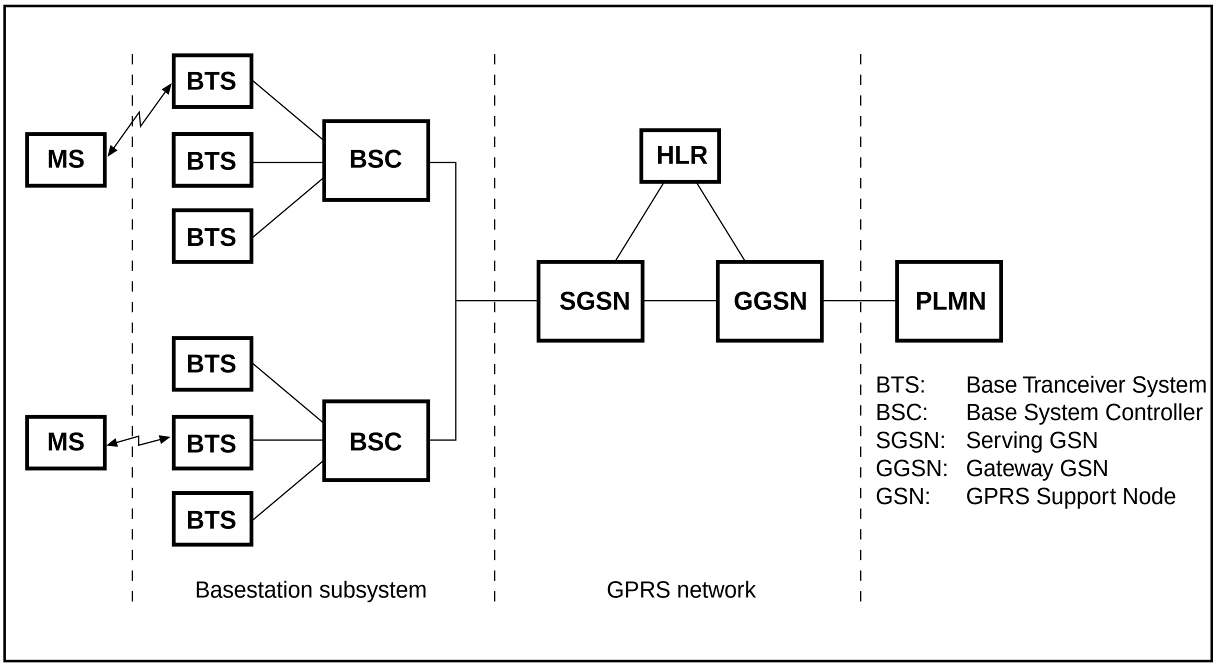


Figure A1-1

An hierachical protocol structure according to the ISO/OSI reference model is used to handle the data transmission. See figure A1-2. Large information blocks from the SNDCP (Subnetwork Dependent Convergence Protocol) are segmented and placed in LLC frames (LLC: Logical Link Control). Different frame lengths are possible, the maximum permitted length is 1600 octets. Each frame contains parity bits for ARQ on the LLC level (FCS: Frame Check Sequence) and a frame header (FH) with routing information.

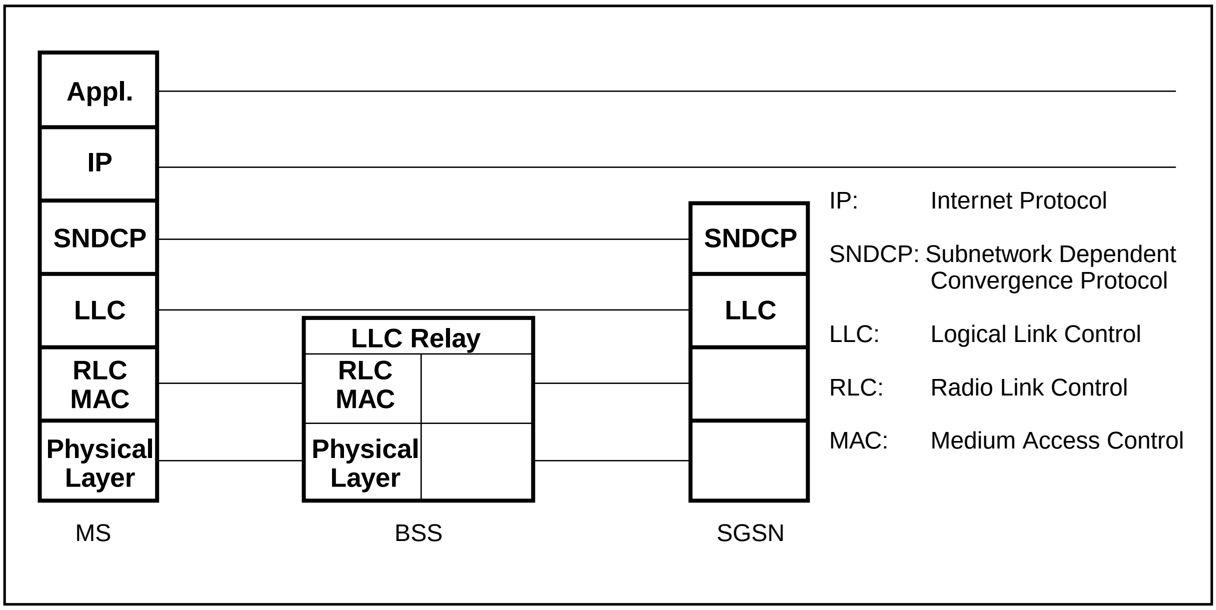


Figure A1-2

The next layer is the RLC layer (Radio Link Control). See figure A1-3. A LLC frame is broken up in a number of radio blocks, which also comprises a header (BH:Block header) and parity bits for selective ARQ (BCS: Block Check Sequence). Two types of radio blocks are used: data blocks and signalling blocks. Both type of blocks start with a MAC header, comprising USF, T and PC fields. The USF (Uplink State Flag) is used in connection with the reservation of radio blocks on the inward traffic channel. The T field is a flag, which tells if a block is used for data transfer or for signalling. The PC field is used in connection with the power control.

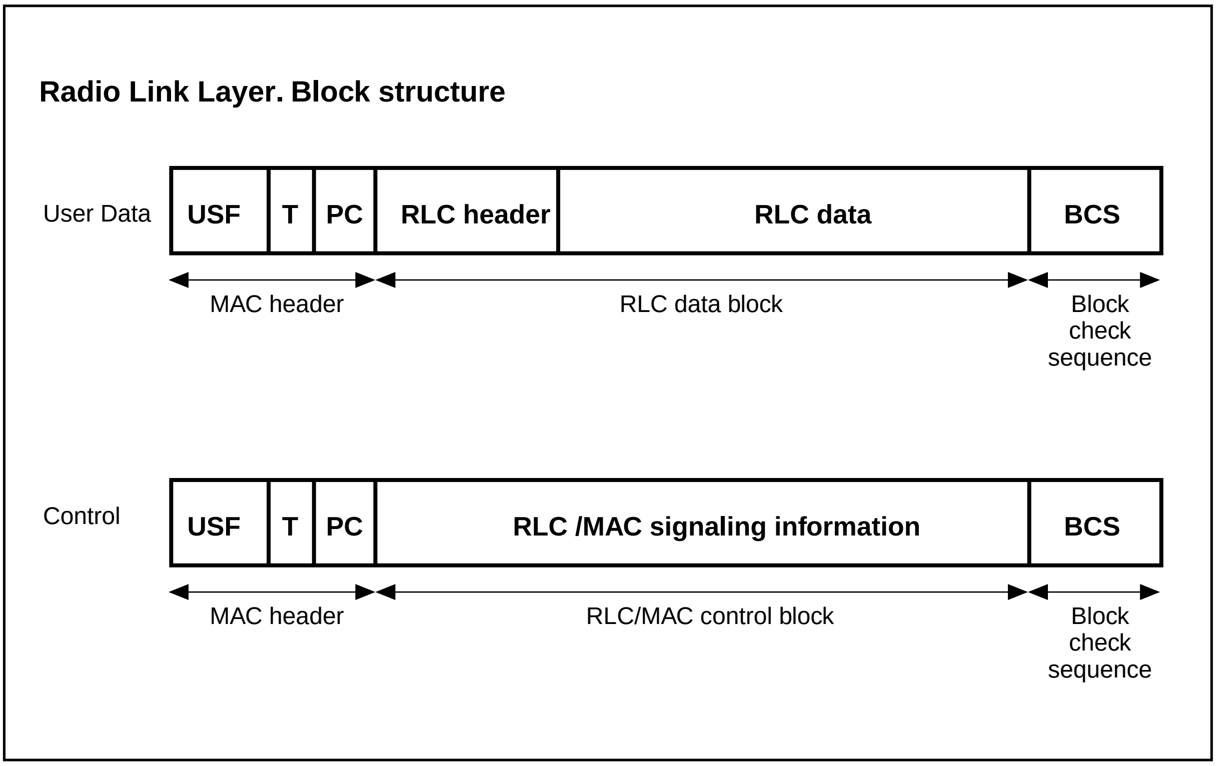


Figure A1-3

The relations between the data sequences in the three layers are shown in figure A1-4.

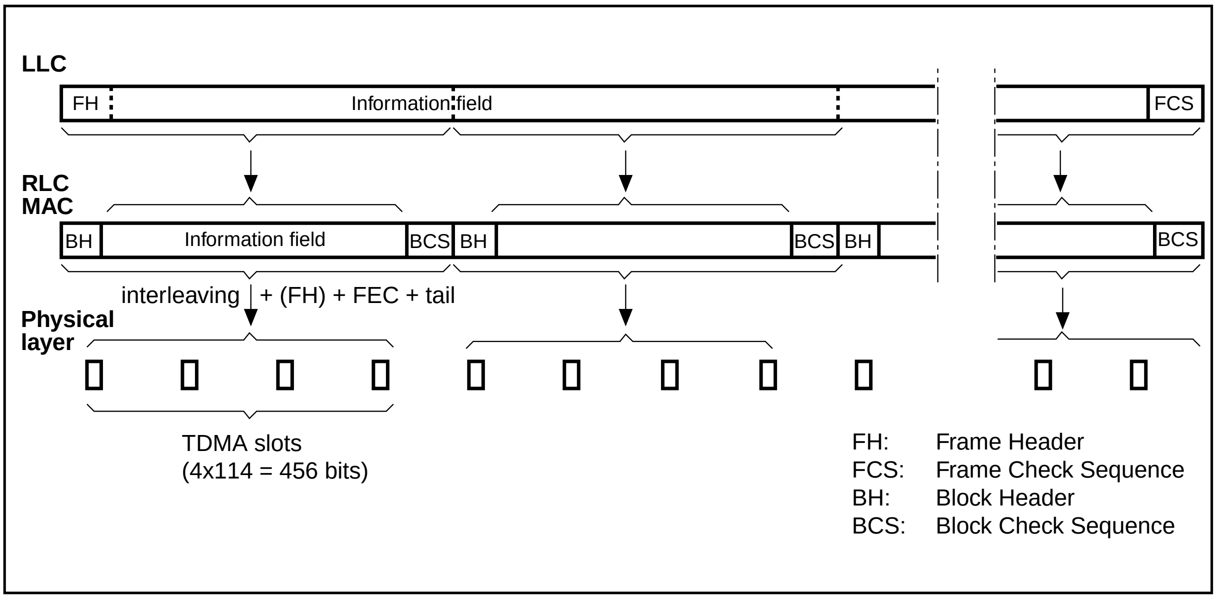


Figure A1-4

The RLC blocks are numbered (TFI: Temporary Flow Identifier) and the receive side can request retransmission of erroneus blocks. Besides ARQ based on the BCS, FEC coding can be applied. See section 3.3.

The radio blocks are fed to the Physical layer, which is based on slots in the normal 8-slot TDMA frame. Each ot the 8 slot positions constitutes one Packet Data Channel, which is multiplexed between traffic channels and different signalling channels.

One radio block is allocated 4 slots in successive frames, i.e. one block comprises 4x114 = 456 bits. Most of the slots in a Packet Traffic Channel (PTCH) are used for data transfer (PDTCH: Packet Data Traffic Channel) but a small percentage of the slots is used for signalling (PACCH: Packet Associated Control Channel). The arrangements corresponds in principle to the signalling structure of the original GSM.

The PTCH is shared between several simultanous connections (sessions) under Media Access Control (MAC), see section 3.2. If multi-slot is used, one connection is served by more than one PTCH.

In principle, the same signalling facilities as in the orignal GSM are used for timing advance, power control and MAHO. Example of extended facilities are:

- possibilities for several hand-over protocols (one alternative is that the terminal by itself determines suitable handover)
- paging groups to permit Discontinuous Reception DRX (the terminal has to listen to pages only during certain time intervals and can go to sleep in between).

The network control is decentralized with the RLC and MAC performed by the BSC units in the BSS. The handover is also placed in this layer. Radio blocks can be destroyed during handover, but that is handled by ARQ at the LLC layer. Some of the functions of the LCC layer are placed at the BSC (LCC Relay).

3.2 Media Access Control (MAC)

To permit efficient sharing of a common channel resource between several sessions, the principle is that reservation of time slots (groups of four) for Radio blocks is only given as long as there are information stored in the buffer memory on the transmit side. As soon all the buffered data has been successfully transferred (incl. possible ARQ retransmissions) the channel allocation is released and can be used by other connections. When a new burst of data arrives to the buffer, a new reservation of time slots must be made. The network might even interrupt long data sequencies, in order to reduce waiting times and a more fair access to the medium, when operating close to the capacity limit.

The MAC is more complicated for uplink data transfer (multiple access). In the downlink direction the base controller has full information about requests for transmission capacity and can store them in a common queu (single access).

Uplink transmission

The BSC controls the inward data traffic by means of Uplink State Flags (USF), which have 3 information bits. Of the 8 available flags, one is used to mark the slots which make up the inward access channel (Aloha) (USF = Free). The other 7 flags (R1 - R7) are used to reserve 4-group of time slots for terminals. An example of the allocation procedure is given in figure A1-5.

Mobile-originated traffic

DMA fr	1 2 3	4 5 6 7	8 9 10 11	12 13 14 15	16 17 18 19	20 21 22 23	24 25 26 27	28 29 30 31	32 33 34 35	36 37 38 39	40 41 42 43	44 45 46 47	48 49 50 51	......
TS No Downlink			ChRes A-B											
TS 0	F	F	F	F	F	F	F	F	F	F	F	F	F	
Uplink	PA PA													
Downlink									ACK-B				ACK-A	
TS 1	R1	R1	R1	R2	R2	R1	R1	R1	R1	R1	R2	R1	R1	
Uplink					A1	A3	B2				A3			
Downlink									NACK-A					
TS 2	R1	R1	R1	R2	R1	R2	R1	R1	R1	R1	R1	R1	R1	
Uplink					A2	B1	A4							
Downlink														
TS 3	R1	R1	R1	R1	R1	R1	R1	R1	R1	R1	R1	R1	R1	
Uplink														

F: Flagcode "free" (TSO uplink available for paging from terminals)
R1, R2: Flagcodes which point to assigned packet slots

Figure A1-5

Time slot 0 (uplink) in the figure is used as Packet Random Access Channel PRACH). Two terminals make request (PA) during frame 1 and 3 repectively. The PRACH is a contention channel, with a risk of collision of simultanious requests from several terminals. To minimize the risk for repeated collisions, a more refined procedure is used than in original GSM. It is a protocal of the “persistance“ type, see module S5, section 3.3.

If channel capacity is available, time slots are allocated to the terminal (immediate assignment), including the maximum number of radio blocks that could be transmitted. This information is sent on the PAGCH (Packet Access Grant

Channel). Also a two-step reservation process can be used. In the second step the terminal comes back with a Packet Resource Request, and the base replies with a Resource Assignnment. This signalling is performed on the two-way Packet Associated Control Channel (PACCH). Compare the signalling in the original GSM. If no channel capacity is available for the moment, the terminal is informed by a Packet Queuing Notification.

In the example, the BCS makes an immediate assignment by sending back a reservation for the two calls (A and B) on times slot 0 (downlink) as a radio block during frames 8 to 11.

USF = R2 is used to mark reservations for call A, and USF = R1 is used for call B. The reservation message also includes information about which frequency slot and timeslot within a frame shall be used. In this case, both calls will be serviced by timeslot 1 and 2. The terminals therefore will monitor the signalling sent on these slots. The reservation for radioblocks is made by transmitting the corresponding USF flags over the downlink channel during the block position before the allocated position.

The first two allocations for call A are during frames 16 to 19. Therefore during frames 12 to 15 USF = R2 is transmitted on the downlink channels TS1 and TS2. Call A is also allocated frames 20 to 23 on TS1 and 24 to 27 on TS 2. Therefore USF = R2 is also transmitted downlink during frames 16 to 19 on TS1 and during frames 20 to 23 on TS 2:. In the same way call B is allocated two timeslots on TS1 and TS2.

The two B blocks were transmitted successfully, and therefore a positive acknowledgement (ACK-B) is transmitted back during frames 32 to 35 on a PACCH (Packet Associated Control Channel). On the other hand, block A3 was not transmitted successfully. Therefore a negative acknowledgement (NACK-A) is transmitted back during frames 32 to 35. The NACK-A message indicates that blocks A3 must be retransmitted (selective ARQ) and also gives additional reservation for the retransmission. A3 is retransmitted during frames 40 to 43. The transmission was succesful this time and ACK-A is transmitted during frames 48 to 51.

Downlink transmission

For downlink data transfer, the first step is a Packet Paging Request on a paging channel (PPCH). The terminal replies with a Packet Channel Request. The rest of the set-up procedure corresponds to what has been described above for inward data transfer.

Point-to-multipoint

Data can also be transmitted from the network to several terminals (point-to-multipoint or broadcast). To set up this connection a special channel is used (PNCH: Packet Notification Channel) to inform the terminals concerned about the resource assignment for the packet transfer.

Multiframe

The basic TDMA frames are grouped in multiframes comprising 52 basic frames. The main reason for the multiframes is for discontinous reception, i.e. the terminals have to listen to pages from the BSC only during a specified part of each multislot.

3.3 Channel coding for the PDTCH

An adaptive channel coding arrangement is used with the possibility to dynamically select one of four channel coding procedures. These are called CS-1 to CS-4. In addition to the coding of the information sequence also the USF is protected by a block code.

CS-1 is a combination of a rate 1/2 convolutional code and a long BCS (Block Check Sequence), resulting in a net data rate of 9.05 kb/s. CS-4, which only contains a short BCS and no FEC, has a net data rate of 21.4 kb/s. The convolutional code that is used for CS-1, is punctured to get the higher rate codes CS-2 and CS-3, which give net data rates of 13.4 and 15.6 kb/s respectively. To reduce the degradation due to error bursts, interleaving over four time slots is used. In addition frequency hopping may be introduced. This is especially motivated with quasi-stationary terminals.

When there are block errors, the radio blocks, which contain errors, are identified by their TFI numbers (Temporary Frame Identity). The TFI:s are sent back to the transmitter side, so that corresponding blocks can be retransmitted. The numbering is modulo 128 and 64 blocks can be handled by the ARQ protocol. Positive acknowledgement is sent back for successfully received blocks. As gradually the block errors are cleared, the error control window can be moved forward. Radio blocks without errors are fed to the LLC layer, so that LLC frames can be generated on the receive side.

The optimum code arrangement, i.e. the code which gives the highest throughput, depends on the ber from the data demodulator. With extremely low ber, the number of retransmissions becomes very low even without FEC. Therefore CS-4 gives the highest throughput. When the quality of the radiochannel degrades more and more, it is optimum to introduce successively more FEC. The reduction of the basic data rate (without the influence of ARQ) is more than compensated for by increased probability of successful block transmission. Already CS-3 gives most of these advantage - CS-2 and CS-1 give small additional gains at very marginal radio channels. Typical relations between channel quality and throughput for the different coding alternatives are indicated in figure A1-6. The figure corresponds to a propagation channel with short fading dips (fairly high terminal speed or low speed in combination with frequency hopping).

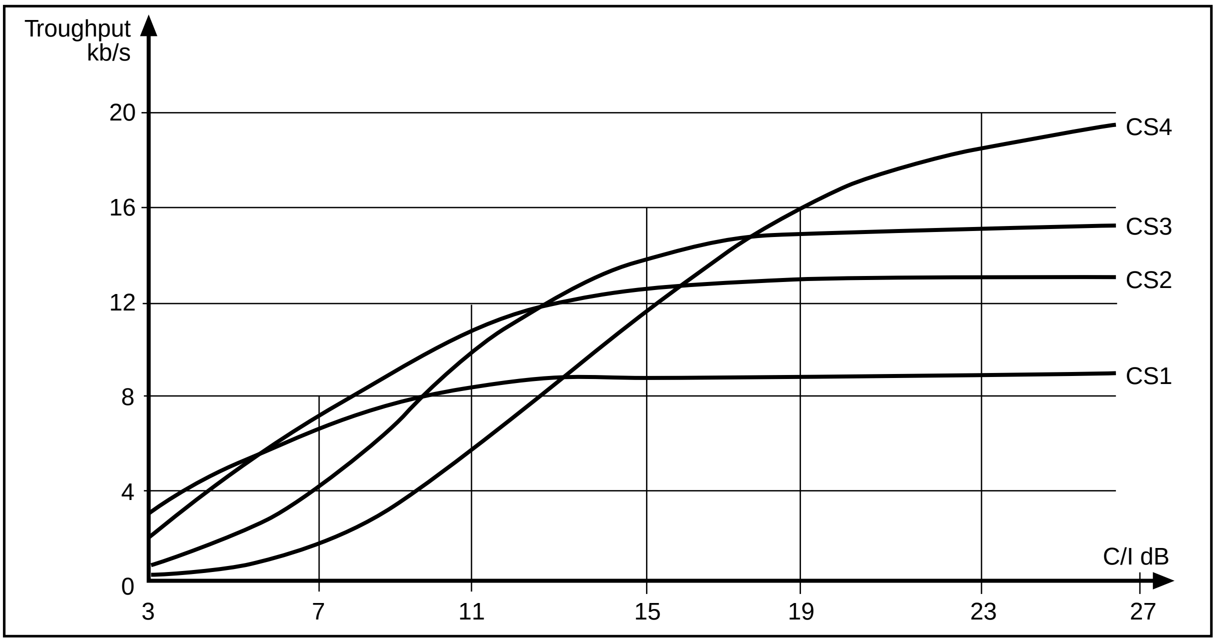


Figure A1-6

4. Edge

Edge stands for Enhanced Data rate for GSM Evolution. A further evolution of GSM/GPRS is to introduce more spectrum-efficient modulation. It is an extension of the GPRS arrangement, i.e. also multi-slot is used. It is to be used in an adaptive arrangement based on packet transmission. A dynamic selection of suitable combinations of channel coding and modulation is used depending on the the quality (C/I and E/N_0) of the radio channel. The net transmission data rate (including the effect of channel coding) will be reduced, when the quality of the radio channel goes down. The throughput will in this case be further reduced as more packets must be retransmitted (ARQ).

Several alternative modulation arrangements have been proposed and studied. One possibility is to use Offset 4QAM (similar to what is used for DAMPS), another Offset 16QAM. O- 4QAM gives in principle the same sensitivity as GMSK but results in 33% higher symbol data rate. O-16QAM gives a further increase of the data rate by a factor of 2, but results in 4-6 dB less sensitivity.

The final proposal for an international standard, sent to ITU in May 98, is based on an agreement between study groups in Europe and USA, i.e. applicable both to GSM and DAMPS (IS-136). 200 kHz channel spacing will be used also for evolved DAMPS by combining several 30 kHz channels. The modulation will be either GMSK or linear 8PSK. With 8PSK a gross rate of 69.2 kb/s will be used.

Simulations have been performed to estimate the maximum throughput for packet transmission using 8 time slots. Assuming a maximum normalized delay of 0.15 sec/kb/s, the following results were obtained.

For a cluster size of 9 (normal case for GSM) and a normalized offered load of 0.325 b/s/Hz/site, 30% of the users achieve an average bit rate of 384 kb/s and 97% 144 kb/s.

Similar results were obtained with a 3-cluster arrangement (normal case for DAMPS with less spectrum available).

Author professor **Sven-Olof Öhrvik**
in cooperation with Ericsson Radio Systems AB
unit ERA/T, Core Unit Radio System and Technology
Publisher **Ericsson Radio Systems AB**
T/Z Ragnar Lodén

Ericsson Radio Systems AB
Torshamnsgatan 23, Kista
S-164 80 Stockholm, Sweden
Telephone: +46 8 757 00 00
Fax: +46 8 757 36 00

EN/LZT 123 1246/1 R6