Principles of modulation and channel coding for digital broadcasting for mobile receivers

This article explains the benefits of using a system called Orthogonal Frequency Division Multiplexing to overcome the adverse effects of severe multipath propagation, such as occurs in mobile reception. The signal is demodulated with the aid of a Fast Fourier Transform technique.

Consideration is given to the digital coding arrangement, and it is concluded that a concatenation of a convolutional code and a Reed-Solomon code gives excellent results.

The feasibility of implementing such a system for the domestic market is briefly discussed.

1. Introduction

The broadcasting channel for mobile reception has a particularly hostile transmission environment, simultaneously containing multiple paths, interference and impulsive parasitic noise. We will not deal here with the problem of channel modelling; this subject has already been treated in [1] and in other references [2, 3]. It will simply be taken that the impulse response of the channel can be represented as the sum of Dirac pulses having different delays. Each of these pulses will be subject to a multiplying factor of which the amplitude generally follows a Rayleigh law and which corresponds to the effect of local diffraction in the region near the mobile receiver.

Typically, this pulse train extends over several microseconds [4, 5], so that the transmission of high bit-rates (several Mbit/s) is a risky process.

The techniques generally used to resolve this problem are based on an extension of the duration of the symbols by increasing the dimension of the
symbol alphabet. Among these techniques, the most widely used is that of spectrum-spreading [6, 7], where the emitted symbols have a wide spectrum and a narrow autocorrelation function, thus permitting discrimination between the different paths composing the received wave. The principle is to define an alphabet of \( M \) pseudo-random orthogonal sequences and to transmit one of these sequences, which is equivalent to \( \log_2 M \) bits of information. Unfortunately, the constraints of intercorrelation between different sequences [8] lead to an important increase in the bandwidth which is far from compensated by the increase in the number of bits emitted per symbol. That is why the spectral efficiency of this type of system is always relatively low. A concrete example can be found in the domain of digital cellular radiotelephony [9].

Another approach, of which the general philosophy is described in [1], consists in splitting the information to be transmitted into a large number of elementary sub-channels each carrying a low bit-rate. This transforms a highly-selective wide-band channel into a large number of non-selective narrow-band channels which are frequency-division multiplexed (FDM). This allows the problem of intersymbol interference to be solved by increasing the symbol duration in the same ratio as the number of sub-channels. There remains the problem of fading, the amplitude of each of the sub-channels following a Rayleigh law or a Rice-Nakagami law [10] if there is a direct path.

The use of a coding system adapted to the fading nature of the channel permits the performance to be considerably improved. Among the codes which could be envisaged, convolutional codes are particularly efficient and special attention is paid to them in this article. Their performance can be further improved by using concatenated codes [11] where the convolutional code plays the role of the inner code. This technique, which was initially developed for space applications, proves to be particularly effective in the domain of transmission to mobile receivers.

In this article, we describe the principles of a frequency-multiplexing system (FDM) which is particularly effective from the point of view of the spectrum, and which is known as OFDM (Orthogonal Frequency Division Multiplexing) [12, 13, 14, 15]. We then describe methods of demodulation applying the fast Fourier transform (FFT) to this type of signal. We then consider the problem of the optimum decoding of a code associated with OFDM modulation in a Rayleigh selective channel. Having presented the performance obtained in these conditions by means of convolution codes or concatenated codes, we analyse these results and draw a certain number of conclusions concerning the consequences for the source coding. Finally, having examined the complexity of implementing the demodulation and decoding arrangements, we consider the feasibility, in the domestic mass-production context, of systems based on these principles.

2. OFDM multiplexing system

2.1. Spectral arrangement

In the applications envisaged for broadcasting to mobile receivers, the object is usually to receive only a part of the broadcast information. This is particularly so in the case of digital sound broadcasting, where it would rarely be useful to receive simultaneously more than one sound programme among the \( L \) programmes emitted.

Moreover, the modulation systems described hereafter form part of the category of systems known as wideband; that is to say, where the product of the available bandwidth \( W \) and the spreading \( \tau \) of the pulse response of the channel is large compared with unity. In these conditions, it is unlikely that fading would simultaneously affect the whole of the channel.

The two foregoing remarks suggest a frequency arrangement as shown in Fig. 1. Each programme comprises \( M \) carriers spaced uniformly throughout the allocated band, the \( L \) programmes sharing the available frequency resource. This arrangement allows the carriers associated with a programme to have the maximum spacing and thus to ensure the best possible conditions of independence.

![Figure 1](image)

At first sight, it appears natural to consider the spectra of the \( N = LM \) digital carriers as being separate. A set of \( M \) matched filters can then be used to separate the emitted carriers. However, this solution has two disadvantages:
- the fact that the carrier spectra are separate does not allow the spectrum to be filled uniformly, so there is a loss of spectral efficiency;
- from a technological point of view, it is difficult to imagine a set of matched filters as the input stage of the receiver, especially if \( M \) is large.

An alternative solution consists in tolerating an overlapping in the spectra of the emitted signals, provided that these satisfy certain orthogonality conditions which ensure the absence of inter-symbol interference at the sampling time. These conditions of orthogonality are a generalization of the first Nyquist criterion where the frequency dimension is added to the temporal dimension and they form the basis of OFDM multiplexing, of which the general principles are described below.

2.2. General principles of OFDM

Let \( \{ f_k \} \) be the set of carrier frequencies under consideration, with:

\[
f_k = f_c + k / T_s , \quad k = 0 \text{ to } N - 1
\]

where \( T_s \) represents the time duration of the symbol.

A base of elementary signals is then defined as \( \psi_{j,k}(t) \) with \( k = 0 \text{ to } N - 1, \ j = -\infty \text{ to } +\infty \):

\[
\psi_{j,k}(t) = g_k(t - jT_s) \\
\text{with:} \begin{cases} 
0 \leq t < T_s : g_k(t) = e^{2\pi i f_k t} \\
\text{otherwise} : g_k(t) = 0
\end{cases}
\]

The signal spectra \( g_k(t) \) mutually overlap, as illustrated in Fig. 2.

It can easily be seen that the set of signals \( \psi_{j,k}(t) \) satisfies the orthogonality conditions:

\[
j \neq j' \text{ or } k \neq k' : \int_{-\infty}^{+\infty} \psi_{j,k}(t) \overline{\psi_{j',k'}}(t) dt = 0
\]

and \( \int_{-\infty}^{+\infty} |\psi_{j,k}(t)|^2 dt = T_s \)

\( \overline{\cdot} \) represents the complex conjugate and \( \| \cdot \| \) the modulus of a complex number.

We then postulate a set of complex numbers \( \{ C_{j,k} \} \) having values taken from a finite alphabet and representing the emitted data signal.

The associated OFDM signal can then be written:

\[
x(t) = \sum_{j = -\infty}^{+\infty} \sum_{k = 0}^{N-1} C_{j,k} \psi_{j,k}(t)
\]

and the decoding rule can be deduced as:

\[
C_{j,k} = \frac{1}{T_s} \int_{-\infty}^{+\infty} x(t) \overline{\psi_{j,k}(t)} dt
\]

Fig. 3 gives an example of the construction of an OFDM signal in which the emitted symbols have their values from an alphabet \( 1 + i, 1 - i, -1 + i, -1 - i \) (4-phase modulation, 4 PSK). When the emitted symbols are independent and have equal probabilities, the corresponding spectral density of the OFDM signal can easily be calculated. An example is given in Fig. 4 for the case \( N = 32 \). It should be noted that this value is used here only for clarity of the diagram, and the values envisaged in practice are considerably higher. It may be noted that, even if the secondary lobes have a high amplitude, their width depends only on \( T_s \) and their relative value therefore decreases as \( N \) increases.

The spectrum of an OFDM signal then tends asymptotically towards an ideal rectangular spectrum. In the case of 4 PSK of the different carriers, the asymptotic spectral efficiency is 2 bits/Hertz. The OFDM multiplexing technique is thus equivalent to 4 PSK with no roll-off in the temporal domain.

2.3. Utilisation of a guard interval

In the presence of inter-symbol interference caused by the transmission channel, the properties of orthogonality between the signals are no longer maintained. One can approach asymptotically towards a solution of the problem of channel selectivity by increasing indefinitely the number of carriers \( N \). However, this method is limited by the temporal coherence of the channel (Doppler effect) or simply by technological limitations of the phase noise of the oscillators.

Another solution consists in deliberately sacrificing some of the emitted energy to eliminate the
The associated OFDM signal can then be written:

\[ x(t) = \sum_{j=-\infty}^{+\infty} \sum_{k=0}^{N-1} C_{j,k} \psi_{j,k}(t) \]

from which the decoding rule can be deduced as:

\[ C_{j,k} = \frac{1}{T_s} \int_{-\infty}^{+\infty} x(t) \psi_{j,k}^*(t) \, dt \]
Compared with the original OFDM system (without guard interval), the mismatching between the emitted signals and the pulse response of the receiving filter corresponds to a loss of 10 log $T_s' / T_s$. This loss can in practice be kept below 1 dB. This disadvantage is largely compensated by the system advantages in a real transmission channel.

2.4. Analysis of the behaviour of an OFDM system, with a guard interval, in the presence of multipath propagation

The transmission channel $c$ is modelled in accordance with the relation:

$$c(t) = y(t) = [x * h](t)$$

where $h(t)$ represents the impulse response of the channel and "*" denotes the convolution product. For the purpose of the following analysis, it is sufficient to state that the impulse response $h(t)$ has a finite duration $\tau$ less than $\Delta$ and that the channel varies slowly compared with the duration $T_s$.

The reader interested in a more-detailed modelling of the channel can refer to various published material (for example, [2, 3, 10]).

We will describe the channel by a series of discrete values $H_{j,k}$ representing its complex response at the frequency $f_k$ and at the instant $JT_s'$. If we consider the time interval $[JT_s', JT_s + T_s]$, the received signal no longer depends on the values of the symbols of the form $C_{j-p,k}$, $p > 0$. The memory of the channel is less than the guard interval, so the behaviour is as if the sinusoids composing the OFDM signal were emitted at an infinite time previously (see Fig. 5).

The received signal can then be written:

$$jT_s' \leq t < jT_s + T_s \quad y(t) = \sum_{k=0}^{N-1} H_{j,k} C_{j,k} s_j(t)$$

In applying to the received signal $y(t)$ the decoding rule previously indicated, we thus obtain simply the transmitted $C_{j,k}$ modified by a multiplying factor corresponding to the channel response at the instant $jT_s'$ and at the frequency $f_k$:

$$H_{j,k} C_{j,k} = \frac{1}{T_s} \int_{-\infty}^{T_s} y(t) s^*_j(t) dt$$

In other words, the addition of a guard interval has totally removed the effects of channel selectivity. Another consequence of the introduction of the guard interval is a modification of the spectral density of the emitted signal. This is illustrated in Fig. 6, which applies to the same conditions as those of Fig. 4, with a guard interval $\Delta = T_s/4$.

![Figure 5](image_url)

OFDM signal with a guard interval $\Delta = T_s/4$.

![Figure 6](image_url)

Example of the power spectral density of the OFDM signal with a guard interval $\Delta = T_s/4$ (number of carriers $N = 32$).

3. Demodulating of the OFDM by a partial FFT

At this stage of the description, the structure of the OFDM receiver does not differ much from that of an FDM receiver: for each of the emitted carriers $f_k$, a filter matched to the signal $g_k(t)$ is necessary. Fortunately, the orthogonality of the signals $g_k(t)$ allows the number of operations to be carried out to be considerably reduced. The structure of the receiver described below is based on the use of an FFT algorithm. In fact, because the receiver deals
a priori with only one programme among \( L \), the algorithm need not be completely executed. This is why it is convenient to call it a "partial FFT".

3.1. Sampled representation of the received signal

Before any processing, it is necessary to convert the continuous signal \( y(t) \) into a train of samples. We are then interested in the time interval \([0, T_s]\), bearing in mind that the channel memory is assumed to be absorbed by the guard interval. The index \( j \) is thus implicitly equal to 0 and will be omitted in the rest of this chapter, so as to simplify the notation. The received signal \( y(t) \) is then written as:

\[
0 \leq t < T_s: \quad y(t) = \sum_{k=0}^{N-1} H_k C_k e^{2i\pi k j / N}
\]

Let us put \( T = T_s / N \)

This signal is translated by means of a local oscillator to the frequency \( f_s + 1/2T \). The corresponding complex signal thus obtained can be written:

\[
\tilde{y}(t) = e^{-i\pi T / T_s} \sum_{k=0}^{N-1} H_k C_k e^{2i\pi k t / T_s}
\]

Then, in sampling the signal \( \tilde{y}(t) \) at a frequency \( f_s = 1/T \):

\[
\tilde{y}(nT) = (-1)^n \sum_{k=0}^{N-1} H_k C_k e^{2i\pi kn / N}
\]

let us put: \( y_n = \frac{(-1)^n}{N} \tilde{y}(nT) \) and \( Y_k = H_k C_k \)

so that: \( y_n = \frac{1}{N} \sum_{k=0}^{N-1} Y_k e^{2i\pi kn / N} \)

\( \{ y_n \} \) thus appears as the inverse discrete Fourier transform (IDFT) of \( \{ Y_k \} \), where:

\[
Y_k = \sum_{n=0}^{N-1} y_n e^{-2i\pi kn / N}
\]

To avoid any problem of aliasing, the sampling frequency \( f_s \) should be more than twice the maximum frequency of the signal \( y(t) \). This requirement is theoretically satisfied by choosing \( f_s = 1/T \). Nevertheless, to take account of filtering constraints, it is in practice necessary to limit the number of carriers effectively emitted to a number \( N' \) which is strictly less than \( N \) (see Fig. 7).

3.2. Processing by means of partial FFT

The calculation of the values \( Y_k \) supposes the carrying out of a DFT over \( N \) complex points in a time less than \( T_s \). There are various algorithms permitting the efficient generation of a DFT known generically by the term "DFT" (Fast Fourier Transform).

Let us now consider the transmission of \( L \) programmes each having \( M \) carriers \( (N = LM) \), among which only \( N' \) are actually emitted \( (N' = LM') \). In the following discussion, among the set of \( N \)

![Figure 7](image)

Reduction of the number of carriers effectively emitted with respect to the Nyquist limit \( (N = 32, N' = 28) \).
carriers we will not distinguish the virtual carriers of the carriers actually emitted. As the receiver has to deal only with \( M \) carriers among the \( N \) emitted, the basic idea behind the simplifications described below is not to carry out sequentially the FFT followed by programme selection, but, instead, to combine them so as to use the programme selection as a procedure for decimating the FFT.

We will adopt the usual conventions by putting (see in particular [16]):

\[
W = e^{-\frac{2\pi}{N}}
\]

and by representing the multiplication of a complex number by \( W^k \) in the form of a graph:

\[
A \xrightarrow{W^k} X = AW^k
\]

as well as representing a DFT over two complex points by the graph:

\[
A \quad X = A + B \\
B \quad Y = A - B
\]

The calculation of the DFT:

\[
Y_k = \sum_{n=0}^{N-1} y_n W^{nk}
\]

can be broken down as follows:

\[
k = Lr + s \quad s = 0 \text{ to } L - 1 \quad r = 0 \text{ to } M - 1 \\
n = Mr + m \quad m = 0 \text{ to } M - 1
\]

Each programme is associated with a particular value of \( s \) and corresponds to \( M \) carriers spaced by \( 1/MT \).

The following result is obtained:

\[
Y_{s,r} = \sum_{m=0}^{M-1} W^{Lmr} W^{ms} \sum_{l=0}^{L-1} y_{l,m} W^{ml}
\]

\[
q_{s,m}
\]

The usual procedure consists in calculating the \( q_{s,m} \) by \( M \) DFT over \( L \) points. To the extent that we are interested only in one value of \( s \), the calculation of the \( q_{s,m} \) only represents \( N \) complex multiplications (of which some may be trivial). To this, it is necessary to add \( M \) multiplications by the "twiddle factors" \( W^{ms} \), then a DFT over \( M \) points.

Figure 8
Example of a partial FFT
\( (L = 5, M = 6) \)
The points which are actually calculated are enclosed by a heavy line

DFT
5 points

3 points

174
Figure 9
Example of a partial FFT
($L = 8, M = 4$)
Decimation in frequency (DIF) algorithm
The points which are actually calculated are enclosed by a heavy line.

Figure 10
Example of a partial FFT
($L = 8, M = 4$)
Decimation in time (DIT) algorithm
The points which are actually calculated are enclosed by a heavy line.
The latter operation can be done by means of an FFT, if $M$ can be factorised. Fig. 8 gives an example for $L = 5$ and $M = 6$.

Most of the FFT algorithms are based on values of $N$ which are powers of 2. This allows the use of highly repetitive structures based on elementary operators called "butterflies" (for example, of the type radix 2 or radix 4), and which are well adapted for cabled implementations.

The decimation procedure described above can obviously be extended to these cases. One example is given in Fig. 9 for $L = 8$ and $M = 4$, based on an algorithm known as decimation in frequency (DIF). The algorithm uses a butterfly of the form:

$$A \quad \begin{array}{c} W^k \hline \end{array} \begin{array}{c} X = A + B \\ \hline Y = (A - B) W^k \end{array} \quad B$$

Another variant consists in using an algorithm for decimation in time, for which the graph is finally nothing other than the preceding graph inverted. A similar decimation procedure is then applied to it (see Fig. 10). The algorithm then uses a butterfly of the form:

$$A \quad \begin{array}{c} W^k \hline \end{array} \begin{array}{c} X = A + BW^k \\ \hline Y = A - BW^k \end{array} \quad B$$

While these two algorithms are equivalent in the case of a complete FFT, that is with $N/2 \log_2 N$ operators of the butterfly type having one complex multiplication and two additions, they are no longer so in the case of a partial FFT. In this case, the algorithm for decimation in time (DIT) starts by trivial multiplications by 1 or $-1$ during the two first stages of butterflies, while in the case of the algorithm for decimation in frequency, the trivial multiplications constitute the latter stages of the FFT.

In the case of a partial FFT, the algorithm for decimation in time is thus much more favourable than the algorithm for decimation in frequency, as shown in the table below.

It may be noted that the ratio of the number of non-trivial multiplications for each processed point, which is a measure of the necessary computing power, varies between 0.3 and 0.45 for the DIT algorithm.

Compared with classical wide-band modulation with digital filtering at the receiver, this result is extremely favourable. In other words, the partial FFT permits one to obtain the benefits of a wide-band OFDM approach while limiting the computing power to a value comparable with that necessary in a narrow-band modulation scheme.

4. A posteriori maximum likelihood decoding in OFDM

4.1. The need for channel coding

The OFDM system in suppressing the effects of channel selectivity, leads us to a Rayleigh or a Rice-Nakagami model of the channel if there is a direct path. The values $Y_{j,k}$ from the partial FFT permit us to find the emitted symbols $C_{j,k}$ by using the relationship:

$$C_{j,k} = \frac{Y_{j,k}}{H_{j,k}}$$

where $H_{j,k}$ is the channel response at the frequency $f_k$ and at the instant $jT_s$.

In the absence of noise, the emitted symbols can be recognized without error, and the phenomena of residual error-ratio caused by channel selectivity, such as those described in [1], disappear. However, the decrease of error-ratio as a function of the ratio $E_b/N_0$ is extremely slow in a Rayleigh channel. This characteristic has often been considered as inherent in a transmission channel for mobile receivers, in particular for narrow-band radiotelephony systems. To pass in such a system from an error-ratio of $10^{-2}$ to $10^{-4}$ in 2 PSK with coherent demodulation, the $E_b/N_0$ ratio must be increased by 30 dB! This implies the need for large operational security margins and also a high resistance to errors of the

| Number of non-trivial complex multiplications for a partial FFT with $L = 16$ |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| $N = $ 256 512 1024 2048 4096 |
| DIT 80 176 384 832 1792 |
| DIF 256 528 1088 2240 4608 |
coding source. As an example, it is common to see sound-coding systems for radiotelephony at 16 k/bits specified up to an error-ratio of 10⁻². In other words, the coding source in this situation must compensate for the absence of a channel coding system.

The approach described below results from a completely different philosophy, where the functions of source coding and channel coding are clearly separated. Instead of “progressive degradation” of the source coding system, we have a channel coding system which is “virtually error-free”. We shall see that this approach allows the necessary operational margins to be considerably reduced, while freeing the coding source systems from any constraint with regard to error sensitivity.

### 4.2. Weighting of the symbols \( C_{j,k} \)

We assume that the emitted symbols \( C_{j,k} \) are related by a code the nature of which it is not necessary to define at this stage of the analysis. The channel is modelled by the relation:

\[
Y_{j,k} = H_{j,k} C_{j,k} + N_{j,k}
\]

where \( N_{j,k} \) is a complex Gaussian noise with:

\[
N_{j,k} = N_{I,j,k} + iN_{Q,j,k}
\]

and

\[
E( N_{I,j,k} )^2 = E( N_{Q,j,k} )^2 = \sigma_{n,j,k}^2
\]

where \( E(\cdot) \) represents the mathematical expectancy. The criterion of a posteriori maximum probability consists in maximizing with respect to \( |C_{j,k}| \) and under the code constraint the probability density:

\[
p(\{Y_{j,k}\mid H_{j,k} \}, \{C_{j,k}\})
\]

Let us then maximize the expression:

\[
\prod_j \prod_k \exp \left( -\frac{\| Y_{j,k} - H_{j,k} C_{j,k} \|^2}{2 \sigma_{n,j,k}^2} \right)
\]

or its Neperian logarithm:

\[
- \sum_j \sum_k \frac{\| Y_{j,k} - H_{j,k} C_{j,k} \|^2}{2 \sigma_{n,j,k}^2}
\]

If, in the applications considered below, the emitted symbols \( C_{j,k} \) have a constant modulus, this is equivalent to maximizing:

\[
\sum_j \sum_k \text{Re} \left( \frac{Y_{j,k} H_{j,k}^* C_{j,k}^*}{\sigma_{n,j,k}^2} \right)
\]

where \( \text{Re}(\cdot) \) represents the real part of a complex number.

### 4.3. Application to 2 PSK and 4 PSK modulation

We will now assume that the carriers multiplexed in OFDM are 2 PSK or 4 PSK modulated. Let us put:

\[
C_{j,k} = A_{j,k} + iB_{j,k}
\]

with:

\[
A_{j,k} = \pm 1 \quad \text{and} \quad B_{j,k} = 0 \quad \text{in 2 PSK or} \quad \pm 1 \quad \text{in 4 PSK}
\]

The maximum likelihood decoding consists in maximizing with respect to all of the series of values \( A_{j,k} \) and \( B_{j,k} \) of the code:

\[
\sum_j \sum_k \left\{ \text{Re} \left( \frac{Y_{j,k} H_{j,k}^*}{\sigma_{n,j,k}^2} \right) A_{j,k} + \text{Im} \left( \frac{Y_{j,k} H_{j,k}^*}{\sigma_{n,j,k}^2} \right) B_{j,k} \right\}
\]

The real and imaginary parts of \( Y_{j,k} H_{j,k}^*/\sigma_{n,j,k}^2 \) appear as weightings of the bits \( A_{j,k} \) and, where applicable, \( B_{j,k} \) (in 4 PSK), which have to be taken into account in maximum likelihood decoding to calculate the branch parameters in the Viterbi decoder matched to the given code.

The calculation of these weightings makes it necessary to know the variance of the noise \( \sigma_{n,j,k}^2 \) and the channel response \( H_{j,k} \).

The analysis of the noise is particularly simple with an OFDM system. It is sufficient to interrupt the signal during a time interval of duration \( T_s \) to make a spectral analysis of the noise. Note that it is not necessary that the noise be white. In particular, if the signal is affected by a narrow-band interferer, the weighting has the effect of “deleting” the corresponding carriers in the same way as fading of these same carriers. This property is specific to the OFDM system, and thus renders it extremely attractive for channels which are highly disturbed by noise or industrial interference.

The arrangement for recovering the carrier used in coherent demodulation provides us directly with an estimation of \( H_{j,k} = \rho_{j,k} e^{i\phi_{j,k}} \), or at least of \( \phi_{j,k} \). If the modulus \( \rho_{j,k} \) is not known, it is necessary to replace in the calculation of the parameters the expression:

\[
Y_{j,k} H_{j,k}^* / \sigma_{n,j,k}^2
\]

by the expression:

\[
Y_{j,k} e^{-i\phi_{j,k}}
\]

We may note that this simplified weighting, which is sub-optimum in a Rayleigh channel, is optimum in a Gaussian channel, where \( \rho_{j,k} \) and \( \sigma_{n,j,k}^2 \) are constant.
5. Performance of the codes in a Rayleigh channel

5.1. Criteria for choosing the codes

The channel model which we are seeking to match to the coding system has two essential particularities:

- The modulus $p_{i,k}$ affecting the emitted symbols follows a Rayleigh law. From this point of view, the main interest of the convolutional codes is that taking this weighting into account does not cause a significant increase in the complexity of the decoding. Another alternative consists in utilizing codes in short blocks which can be decoded by a soft decision. However, for a given complexity they have an inferior performance compared with convolutional codes*. In both cases, it is necessary to ensure the maximum independence between the received samples.

- The modulus $p_{i,k}$ is highly correlated with respect to the index $j$ (temporal coherence related to the Doppler spectrum) and the index $k$ (frequency coherence related to the impulse response of the channel $h(t)$). This correlation tends to create, in the absence of coding, groups of errors. This second property leads us towards codes which are capable of correcting bursts of errors. The codes having these properties are of the algebraic type and do not permit the use of soft decision with a "reasonable" complexity.

These two properties thus lead to relatively contradictory indications for choosing the codes. However, we may note that Reed-Solomon codes allow these two properties to be exploited simultaneously by associating an interleaving system matched to the length of the code symbols and a system for deleting the affected symbols in the case of a deep fade [17]. Nevertheless, this "all or nothing" weighting is far from being the optimum.

A relatively exhaustive analysis of the possible options leads one to consider convolutional codes as being the best possible choice as long as the necessary conditions of independence can be ensured. With this assumption, the decoder emphasizes the weighting of the received samples and deliberately masks the statistics in grouped errors, although these potentially constitute a source of complementary information.

5.2. Use of an interleaving arrangement

We will now consider the use of convolutional codes processed by soft decision. The necessary conditions of independence are assured by an interleaving arrangement. This interleaving is done in two dimensions:

- interleaving in time, which is very useful for mobile reception, even at slow speed. This interleaving can be very deep (several hundreds of milliseconds), because the constraints at the level of processing delay in the receiver are much less severe in digital sound broadcasting than in radiotelephony;

- frequency interleaving, which is essential for fixed reception or "stationary mobile". Note that with OFDM the number of carriers per programme can be very large. In fact, the real independence of these carriers depends in particular on the product of the bandwidth used $W$ and the standard deviation of the distribution of the delays. From this point of view, the existence of multiple paths can be considered as an advantage.

Paradoxically, the most critical case, which causes the two systems indicated above to fail, is that of a stationary mobile in a situation where there is a single path with total diffraction leading to a deep fade throughout all the band. In practice, this situation may correspond to the case of a vehicle stationary in a forest.

5.3. Performance of the convolutional codes

The convolutional codes considered in the following simulations have the common characteristic of a constraint length $K = 7$ and their efficiency varies from $1/4$ to $1/2$ [18,19] for conventional codes and from $5/6$ to $8/9$ for punctured codes [20,21]. Their detailed description is given in Annex I.

Fig. 11 gives the performances obtained by these different codes for a Gaussian channel. As a reference, the error-rate curve in the absence of coding is represented by a dotted line.

The two following figures give the performances of the same codes in a Rayleigh channel, based on the assumption that the interleaving arrangement effectively guarantees the conditions of independence required by the Viterbi decoder. Fig. 12 gives the performance obtained with a weighting matched to the Gaussian channel (sub-optimum weighting), while Fig. 13 gives the same results with an optimum weighting. The difference is about 1 dB. Subsequently, only the case of optimum weighting will be considered.

* Note, however, that the parity bit codes of the type (9,8) for example, give a good performance/complexity ratio, while the soft decision decoding is trivial (a trellis with two nodes).
Figure 11 (left)
Bit error-ratio performance of convolutional codes in a Gaussian channel with coherent demodulation.
The letter indicates the respective convolutional code (see Annex 1).

Figure 12
Bit error-ratio performance of convolutional codes in a Rayleigh channel (sub-optimum weighting) with coherent demodulation. The letter indicates the respective convolutional code (see Annex 1).

Figure 13
Bit error-ratio performance of convolutional codes in a Rayleigh channel (optimum weighting) with coherent demodulation. The letter indicates the respective convolutional code (see Annex 1).

5.4. Use of concatenated codes (convolution + Reed-Solomon)

Concatenated codes were introduced by Forney [11]. One of the most effective techniques, which has been used for planetary missions [22,23], is the combination of an internal convolutional code and an external Reed-Solomon code, as shown in the
schematic diagram of Fig. 14. The second disinterleaving arrangement (as well as the interleaving system associated with the emission) has the object of cutting the bursts of errors at the output of the Viterbi decoder into symbols which are compatible with the Reed-Solomon decoder and of interlacing them in such a way as to guarantee the optimum conditions of independence for the latter decoder.

Reed-Solomon codes have the general form \((n,k) = (2^t - 1, 2^t - 1 - 2t)\) where \(t\) represents the number of corrected symbols having \(j\) bits. For the following simulations, we have taken \(j = 8\) with \(n = 255\).

**Gaussian channel**

Here we consider a family of concatenated codes based on an internal convolutional code chosen from those given in Annex 1, and an external Reed-Solomon code with a variable correction capacity \(t\). \(R\) will designate the over-all efficiency of the code, that is the product of the efficiency of the convolutional code and that of the Reed-Solomon code (1-2t/255).

**Rayleigh Channel**

Figs 16a, b and c give the same results as in Figs 15a, b and c, but for a Rayleigh channel. It may be noted that the concatenated code \(C\) permits an error ratio of \(10^{-9}\) to be attained for a ratio \(E_b/N_0\) of 4.5 dB. The difference between a Gaussian channel and a Rayleigh channel is therefore reduced to 2 dB.

5.5. Utilisation of concatenated codes (convolution + CSRS)

Although the concatenated codes previously described offer remarkable performance, their implementation in the context of the mass production domestic market is difficult because of the complexity of decoding the Reed-Solomon codes. An alternative solution consists in utilising CSRS codes (cyclically shortened Reed-Solomon) [24, 25, 26] which are, like the Reed-Solomon codes, at a maximum distance separable \((d_{min} = n - k + 1)\). The essential interest of these codes is that they do not require processing.
Figure 15
Performance of concatenated codes (convolution plus Reed-Solomon) in a Gaussian channel with coherent demodulation

a) Ratio $E_b/N_0$, necessary for an error-ratio of $10^{-4}$, as a function of the overall efficiency $R$.
b) Ratio $E_b/N_0$, necessary for an error-ratio of $10^{-9}$, as a function of the correction capacity ($T$) of the Reed-Solomon code.
c) Bit error-ratio as a function of the ratio $E_b/N_0$.

The letter indicates the convolutional code used as the internal code or respective concatenated code (see Annex 1)
Figure 16
Performance of concatenated codes
(convolution plus Reed-Solomon)
in a Rayleigh channel with coherent demodulation

a) Ratio $E_b/N_0$, necessary for an error-ratio of $10^{-9}$, as a function of the overall efficiency $R$.
b) Ratio $E_b/N_0$, necessary for an error-ratio of $10^{-9}$, as a function of the correction capacity ($t$) of the Reed-Solomon code.
c) Bit error-ratio as a function of the ratio $E_b/N_0$.

The letter indicates the convolutional code used as the internal code or respective concatenated code (see Annex 1)
Figure 17
Performance of concatenated codes
(convolution plus CSRS)
in a Gaussian channel with coherent demodulation
a) Ratio $E_b/N_0$, necessary for an error-ratio of $10^{-3}$, as a function of the overall efficiency $R$.
b) Ratio $E_b/N_0$, necessary for an error-ratio of $10^{-9}$, as a function of the correction capacity $(t)$ of the CSRS code.
c) Bit error-ratio as a function of the ratio $E_b/N_0$.

The letter indicates the convolutional code used as the internal code or respective concatenated code
(see Annex 1)
Figure 18
Performance of concatenated codes
(robolution plus CSRS)
in a Rayleigh channel with coherent demodulation

a) Ratio $E_b/N_0$, necessary for an error-ratio of $10^{-3}$, as a function of the overall efficiency $R$.
b) Ratio $E_b/N_0$, necessary for an error-ratio of $10^{-3}$, as a function of the correction capacity ($t$) of the CSRS code.
c) Bit error-ratio as a function of the ratio $E_b/N_0$.

The letter indicates the convolutional code used as the internal code or respective concatenated code (see Annex 1)
Galois field extension GF (2^s), but only over GF (2). The implementation of the decoding is considerably simplified. The codes considered below have the form (n, n − 2t), where t is the number of corrected symbols of j bits. For the simulations described hereafter, we have taken j = 12 et n = 336.

Gaussian channel

The results obtained with convolutional codes concatenated with CSRS codes are given in Figs 17a and b for the same conditions as for Figs 15a and b for the Reed-Solomon codes. The results are very similar to those previously obtained. In the same way, we have selected a certain number of concatenated codes of which the characteristics are given in detail in the Annex. These codes are represented by circles in Fig. 17b, and their performance with respect to error ratio is given in Fig. 17c.

Rayleigh channel

Figs 18a, b and c give the same results as in Figs 17a, b and c, but for a Rayleigh channel.

Note that the convolutional code C concatenated with a CSRS code (336, 288) has the same performance as the "NASA" code, but its implementation is simpler.

5.6. Soft decision with differential demodulation

The simulations so far presented are based on the assumption of an optimum weighting of the received samples, and consequently on coherent demodulation of the OFDM signal. The results show that a virtually error-free channel can be obtained with a ratio E_b/N_0 of 4.5 dB when utilising the concatenated code C (either starting from a Reed-Solomon code or from a CSRS code). In terms of E_b/N_0 (E_b designating theenergy per transmitted bit, instead of per useful bit), the corresponding value is less by 1 dB. Considering that this is an average value, this means that, in terms of the instantaneous value this ratio can become highly negative.

It can easily be seen that coherent demodulation in these conditions is particularly difficult. Taking account both of the problems of locking out of the carrier recovery loop and also of the error made in its estimation, it can reasonably be expected that there is a significant degradation in the performance.

Differential demodulation is an alternative solution, of which the essential interest resides in the very great simplicity of its implementation and its absence of inertia after a deep fade. The price to be paid is evidently a degradation in performance, which nevertheless remains acceptable, and which is in reality minimal if account is taken of the practical limitations of coherent demodulation.

We have seen in Section 4.2 that the optimum demodulation of a code using symbols with a constant modulus is equivalent to maximising with respect to C_{j,k}:

\[ \sum_j \sum_k \text{Re} \left( Y_{j,k} H_{j,k}^* C_{j,k}^* / \sigma_{j,k}^2 \right) \]

Differential demodulation consists in using for the row j a simplified estimator from the channel deduced from row j - 1:

\[ H_{j,k} \approx \frac{Y_{j-1,k}}{C_{j-1,k}} \]

Case of 2 PSK

A differential coding system is defined by the relationship:

\[ c_{j,k} = \frac{C_{j,k}}{C_{j-1,k}} \]

with: \( c_{j,k} = a_{j,k} + ib_{j,k} \), \( a_{j,k} = \pm 1 \) and \( b_{j,k} = 0 \)

The weighting of the bit \( a_{j,k} \) to be used at the level of the Viterbi decoder then becomes:

\[ \text{Re} \left( \frac{Y_{j,k} Y_{j-1,k}}{\sigma_{j,k}^2} \right) \]

Case of 4 PSK

The differential coding system described previously is not applicable in 4 PSK, because the symbols \( c_{j,k} \) thus defined have values 1, i, -1, -i, that is, a different alphabet from that of the symbols \( C_{j,k} \). Nevertheless, we can easily adapt the procedure to this case by modifying the differential coding system by putting:

\[ c_{j,k} = \frac{C_{j,k}}{C_{j-1,k}} (1 + i) \]

with \( c_{j,k} = a_{j,k} + ib_{j,k} \), \( a_{j,k} = \pm 1 \) and \( b_{j,k} = \pm 1 \)

The weightings of the bits \( a_{j,k} \) and \( b_{j,k} \) to be used at the level of the Viterbi decoder then become respectively:

185

\[ \text{Re} \left( \frac{Y_{j,k} Y_{j-1,k}}{(1-i) \sigma_{j,k}^2} \right) \] and \[ \text{Im} \left( \frac{Y_{j,k} Y_{j-1,k}}{(1-i) \sigma_{j,k}^2} \right) \]
5.7. Performance of the codes with differential demodulation

To illustrate the performance obtained with differential demodulation, we have taken the example of the convolutional code C with efficiency 1/2 and of its concatenated version with a CSRS code (336, 288).

**Gaussian channel**

Fig. 19 illustrates the performance of the convolutional code C in a Gaussian channel with differential demodulation for 2 PSK and 4 PSK. As a reference, the performances of 2 PSK and 4 PSK without coding and with differential or coherent demodulation are recalled. It may be noted that the coding brings 2 PSK and 4 PSK more or less equal with differential demodulation, while in the absence of coding, 2 PSK has a definite advantage.

Fig. 20 illustrates the performance of the concatenated code in a Gaussian channel for the same conditions. Here again, 2 PSK and 4 PSK offer comparable performances. In these two cases, the price to be paid for the simplicity of differential demodulation is about 3 dB, compared with coherent demodulation.

**Rayleigh channel**

Fig. 21 illustrates the performance of the convolutional code C in a Rayleigh channel with differential demodulation of 2 PSK and 4 PSK. As a reference, the performances of 2 PSK and 4 PSK without coding and with differential or coherent demodulation are recalled. Fig. 22 gives the same results for the concatenated code. In these two cases, the degradation caused by differential demodulation compared with coherent demodulation is about 3 dB.

6. Analysis of the results

All the results explained above bring out clearly the interest of a system combining OFDM multiplexing and coding. Compared with other known systems for digital broadcasting towards mobile receivers, the essential advantages of this approach consist in the following.

186
6.1. The spectral efficiency

This results principally from the use of OFDM which, associated with 4 PSK, permits a spectral efficiency of 2 bits/Hertz to be approached asymptotically. This is of course the basic efficiency, which must be corrected by the efficiency of the associated code. In practice, the most interesting codes are the concatenated codes of type $C$ to type $H$, of which the efficiencies lie between 0.4 and 0.8, and the spectral efficiency lies between 0.8 and 1.6 bits/Hertz. This variety of usable codes permits the choice of a compromise between spectral efficiency and power efficiency, which is the best adapted to the particular situation. But, whatever may be the choice of code, the efficiency obtained is much higher than that which can be obtained by normal spread-spectrum techniques. Fundamentally, this is because, with OFDM, the waveforms are emitted in parallel, whilst in spread-spectrum techniques, a single waveform from the available alphabet is emitted.

6.2. Power efficiency

This depends principally on the coding arrangement. The more powerful the coding arrangement the smaller becomes the difference separating the Rayleigh channel and the Gaussian channel. For example, by using a concatenated code of type $C$, and taking account of the corresponding spectral efficiency, it is seen that with coherent demodulation the performance with respect to values of $E_b/N_0$ is around 3 dB from the Shannon limit in a Gaussian channel and around 5 dB in a selective Rayleigh channel. This can be taken as the state of the art.

Note that OFDM plays an essential role in the correct functioning of the coding system. On one hand, this technique permits an optimum weighting of the samples at the decoding level, which is not the case for other techniques used to combat multiple-path propagation, such as equalisation with feedback decision or spectrum spreading. On the other hand, the wide-band character of the multiplex guarantees the independence necessary for the good operation of the decoder, even in the case of a stationary mobile receiver.

The respective roles played by the convolutional code and the Reed-Solomon or CSRS code are very different. The convolutional code in a sense “inte-
grades” different Rayleigh laws along its trellis thus making the performance tend towards that obtained with the same code in a Gaussian channel. The Reed-Solomon or CSRS code achieves the same objective and gives to the error-ratio curve the shape of a “wall” near the Shannon limit. Moreover, the structure itself of the errors changes, since it becomes a matter of “erroneous blocks” and these are detected in a quasi-certain manner. But, above all, it leads us to abandon even the concept of error ratio, the system swinging from perfect operation to non-operation in less than 1 dB.

7. Consequences on the source-coding problems

Apart from an analysis of the technical performance, it is important to draw from the results certain information on the overall design of a digital transmission system, including the problems of source coding.

7.1. Error protection and source coding

The systems of source coding are usually based on a number of more-or-less implicit assumptions regarding the nature of the transmission channel. The most widespread assumption is that the channel is binary symmetrical where the errors appear in a random way; the degradation with respect to the noise is then certainly abrupt compared with the case of an analogue system, but there is nevertheless a certain progression in the degradation.

This modelling in fact corresponds to that of a modulation system where there is no channel coding. In these conditions, the natural tendency is to devote a part of the available bit-rate to protection against errors of the coded signal.

The results that we have shown demonstrate that the available bit-rate must not be considered as an untouchable parameter but, on the contrary, the bit-rate, spectral efficiency and power efficiency can be mutually exchanged. Moreover, it is imperative that most of the redundancy be situated at a level very near the modulation to permit a posteriori maximum likelihood decoding. In the preceding examples, the convolutional codes should be considered as a means of creating waveforms which are efficient in terms of the minimum distance rather than a redundancy at the binary level.

We have observed that the search for solutions near the Shannon limit leads us to extremely abrupt error-ratio curves. In these conditions, an error-protection system at the level of the source coding can be effective only in the transition region, thus giving a maximum gain of a few tenths of a decibel. The redundancy thus consumed would be better utilized in exchanging it for power efficiency by means of a modification of the channel coding system.

It can therefore be stated that the search for an optimum solution at the level of channel coding and modulation leads us to abandon any error protection at the level of the source coding. If the transmission channel has two states — perfect operation or non-operation — any correction strategy is doomed to failure and has no effect other than uselessly consuming bit-rate.

7.2. Progressive degradation or a virtually error-free channel

Many engineers seek to reproduce with digital systems a behaviour of the analogue type, in optimizing the source coding algorithm such that it is as tolerant as possible to transmission errors. In fact, it is generally considered that a system with progressive degradation allows us to provide a service which is degraded, but nevertheless usable, down to levels where, with a system having a threshold, the service would be totally interrupted.

The analysis of the preceding curves (see in particular Fig. 22, curves C and O) show that, on the contrary, an efficient coding system not only permits the transmission of data virtually without error up to very high noise levels, but it also provides that the point of cross-over with the performance of a system without coding is situated at a value of error-ratio (about $10^{-5}$) for which a service is interrupted irrespective of the source-coding system. Progressive degradation therefore results in transmitting a degraded signal only when it could be transmitted in a perfect manner. This is why it is necessary to abandon the idea of a “progressive degradation” and to tend towards a “virtually error-free channel” near the Shannon limit.

7.3. Considerations concerning the ISO layer model

An essential consequence of introducing channel coding is thus a complete separation between the source-coding functions and the modulation. The source-coding system is limited to its primary function, that is to say, to exploit the redundancies of the signal and the physiological and psychosensorial properties so as to reduce the data rate to be transmitted without having to take account of the problem of error behaviour. For its part, the modulation system and the channel coding ensure error-free transmission of the information coming from the source coding down to a threshold as near as possible to the Shannon limit, without having to take account of the nature of the information transmitted.
This transparency permits us to completely respect the ISO layer model, thus avoiding specialization of the transmission channel.

8. Conclusions

The technique of OFDM multiplexing with a guard interval permits the problem of multipath reception in digital broadcasting to mobile receivers to be resolved. Moreover, the spectral efficiency thus obtained approaches asymptotically 2 bits/Hertz (ignoring channel coding and the guard interval) in the case of 4 PSK carriers.

In terms of complexity, the analysis shows that processing of the OFDM by a partial FFT demands only a relatively low computing power. In fact, the only really difficult question is that of the technical feasibility of a system of coherent demodulation offering a satisfactory complexity/performance compromise. From this point of view, the good performance obtained with differential demodulation, of which the implementation is extremely simple, provides a particularly useful reference element.

While OFDM provides a solution with respect to channel selectivity, it does not by any means overcome fading. This property of the transmission channel clearly indicates the need for a system of channel coding.

Associated with an appropriate interleaving system, the convolutional codes guarantee in this context excellent performance for a reasonable decoding complexity (64-state Viterbi decoder, in the examples chosen). Concatenation with a Reed-Solomon code improves the performance yet further. Nevertheless, the complexity of implementing decoding of Reed-Solomon codes leads us to reject this solution. We propose to use instead CSRS codes, which have comparable properties while offering a greater decoding simplicity.

The principles of modulation and channel coding explained above permit the definition of a digital sound broadcasting system having a performance with respect to $E_b/N_0$ and spectral efficiency close to the Shannon limit. Analysis of the complexity of the demodulation and associated decoding functions leads to the conclusion that the corresponding receivers in a domestic mass-production context are feasible.

---

ANNEX 1

The convolution codes used in the simulations are described in the table opposite. They correspond to the best possible choices among all the codes described in references [18, 19, 20, 21]. The corresponding concatenated codes are obtained by associating with each convolution code the Reed-Solomon or CSRS code of the same line.

<table>
<thead>
<tr>
<th>Type of code</th>
<th>Convolution code</th>
<th>RS code, j=8</th>
<th>CSRS code, j=12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Efficiency</td>
<td>Polynomial generators</td>
<td>(n,k)</td>
</tr>
<tr>
<td>A</td>
<td>1/4</td>
<td>111,135,167,173</td>
<td>(255,211)</td>
</tr>
<tr>
<td>B</td>
<td>1/3</td>
<td>107,135,161</td>
<td>(255,215)</td>
</tr>
<tr>
<td>C</td>
<td>1/2</td>
<td>133,171</td>
<td>(255,223)</td>
</tr>
<tr>
<td>D</td>
<td>2/3</td>
<td>135,163</td>
<td>(255,231)</td>
</tr>
<tr>
<td>F</td>
<td>4/5</td>
<td>133,171</td>
<td>(255,235)</td>
</tr>
<tr>
<td>G</td>
<td>5/6</td>
<td>133,171</td>
<td>(255,235)</td>
</tr>
<tr>
<td>H</td>
<td>6/7</td>
<td>133,171</td>
<td>(255,235)</td>
</tr>
<tr>
<td>I</td>
<td>7/8</td>
<td>133,171</td>
<td>(255,239)</td>
</tr>
<tr>
<td>J</td>
<td>8/9</td>
<td>133,171</td>
<td>(255,243)</td>
</tr>
<tr>
<td>O</td>
<td>no coding</td>
<td>no coding</td>
<td>no coding</td>
</tr>
</tbody>
</table>

* Code used by the NASA
Bibliographical references


