DIGITAL SOUND SIGNALS: THE PRESENT BBC DISTRIBUTION SYSTEM AND A PROPOSAL FOR BIT-RATE REDUCTION BY DIGITAL COMPANDING

M.G. Croll, D.W. Osborne and C.R. Spicer

1. INTRODUCTION

The introduction of stereophonic broadcasts, combined with the gradual withdrawal from service of the British Post Office (BPO) network of 'music lines', has caused the BBC to seek alternative means for distributing high-quality sound programme signals from the studio centres to the main transmitters. Investigations showed that digital systems could provide sound circuits of the particularly high quality and stability required for long distance transmission of stereophonic signals. Accordingly a pulse code modulation (p.c.m.) sound multiplex system was developed and first introduced into service in 1972, carrying signals for Radio programmes 2, 3 and 4 from Broadcasting House, London to the v.h.f./f.m. transmitters at Wrotham, Kent. In 1973 the system was extended from London to the v.h.f./f.m. transmitters at Sutton Coldfield, which serve the Midlands, and to the v.h.f./f.m. transmitters at Holme Moss which serve a large part of Northern England. Further extensions of the p.c.m. sound multiplex system to Scotland, Wales and to the North-Eastern region of England are planned to be installed during the current year.

In future it is likely that p.c.m. links will replace many of the present analogue links used by the BBC, not only for programme distribution to the transmitters but also for contributions to studio centres. The BPO is planning to develop a digit transmission network which, although primarily for telephony, could provide standard 2·048 Mb/s digit transmission circuits for other purposes. Assuming that such circuits were available to the BBC at a favourable cost, the more channels that can be accommodated in them the more economic they become. Hence the BBC has been investigating methods of bit-rate reduction for digital sound signals and, as a result, a system for multiplexing high-quality sound channels at a bit rate of 2·048 Mb/s is being developed. This uses digital companding techniques to enable six channels to be multiplexed at 2·048 Mb/s where otherwise only four channels could be provided.

The first part of this paper is concerned with the fundamentals of p.c.m. and describes the main features of the system already in service. The second part of the paper describes the investigations of bit-rate reduction techniques applied to high-quality sound channels and describes a proposal for transmitting six sound signals at a bit rate of 2·048 Mb/s.

2. FUNDAMENTALS OF P.C.M. APPLIED TO SOUND SIGNALS

2.1. Sampling

To convert an analogue signal into digital form it must first be sampled, thus converting it from a continuous function to a discontinuous function which is a measure of the signal at discrete time intervals. The rate at which an analogue signal must be sampled, to enable it to be accurately re-constructed, is at least twice its highest frequency, according to the well-known Nyquist theorem. In order to achieve an upper frequency limit of 15 kHz a sampling frequency of 32 kHz is used, being such that the required cut-off characteristics of the low-pass filters can be achieved without undue difficulty. Such filters are essential before the analogue-to-digital conversion (a.d.c.) and after the digital-to-analogue (d.a.c.) processes to remove components at half-sampling frequency and above. An advantage of using a sampling frequency of 32 kHz is that it is four times the sampling frequency used for digital telephony signals and synchronising problems would therefore be mitigated in possible future use of BPO digital circuits for high-quality sound signals.

The lowest frequency limit and any variations in response over the audio passband are set by the analogue sections of the a.d.c. and d.a.c. equipment.

2.2. Quantising

The signal-to-noise ratio of a p.c.m. system is dependent upon the number of bits used to describe each sample of the analogue signal. For high-quality sound distribution a signal-to-noise* ratio of at least 75dB

* Peak signal to r.m.s. equivalent white Gaussian, unweighted noise ratio

The authors are with the British Broadcasting Corporation.

is required[1] and, if allowance is made for up to four coding and decoding (codec) processes, the signal-to-noise ratio required from one codec must be at least 75 + 6 = 81 dB. This can be achieved by linearly quantising each sample to 13-bit accuracy which gives a theoretical signal-to-noise ratio, of 83 dB. However to remove 'granular' distortion on low-level highly critical programme material, a 'dither' signal is added to the audio signal prior to the coding process.[2] The dither signal consists of a component at half-sampling frequency (16 kHz) of amplitude equal to half a quantising step together with white noise at a level of 4 dB below the inherent quantising noise. This artifice removes granular distortion at the expense of worsening the signal-to-noise ratio by 1·5 dB, thus the final theoretical signal-to-noise ratio is 81·5 dB which closely corresponds to the target figure of 81 dB.

2.3. Effects of transmission errors

Although a p.c.m. signal is highly resistant to noise and other forms of interference, digit errors can occur when such unwanted signals become excessive. For example, if the p.c.m. signal is transmitted over a microwave link, then during abnormal propagation conditions deep fades may occur such that the random noise level at the receiver can give rise to unacceptably high error rates in the received digital signal. If an error occurs in the most significant bit (m.s.b.) of a sample word a loud click is produced whilst an error in the least significant bit (l.s.b.) would be virtually imperceptible. In linearly-coded signals a degree of protection can be provided against transmission errors by adding a single parity check digit to each 13-bit sample word. This can be used to check the more significant bits and, when an error is detected, the incorrect word is replaced by the previous correctly-received word. If the error rate becomes high enough to impair seriously the quality of the signal the audio output can be muted. The resistance to errors is thus improved, and it has been found that whereas without protection the error rate would need to be not greater than 1 in 10^7 for just perceptible impairment on critical programme items, with protection of the five most significant bits the error rate could rise to 1 in 10^5 for the same impairment.[3]

3. EXISTING 13-CHANNEL P.C.M. DISTRIBUTION SYSTEM

3.1. System description

Each audio input in the existing 13-channel distribution system is fed to a delay line limiter via a 50 µs pre-emphasis network if the channel is feeding a v.h.f./f.m. transmitter. Thus the single limiter can provide protection of both overloading of the p.c.m. system and over-deviation of all f.m. transmitters fed from the p.c.m. system. This is possible because of the excellent gain stability inherent in the p.c.m. distribution system. For stereo operation the left and right hand signals are coded separately but the limiters for each channel are interconnected so that if either limiter operates, both change gain by the same amount thus preventing shift of the stereo image.

Each channel is equipped with a 13-bit a.d.c. which uses dither signals and an extension of the counter-ramp technique in which the converter range of 8192 (2^{13}) uniformly-spaced quantising levels is split into 128 (2^7) 'groups' each group containing 64 (2^6) levels or 'units'. Conversion is achieved by first counting the groups and then 'changing gear' to count the remaining individual quantising levels. By this technique the clock rate for the sample conversion time of 15 µs is only about 13 MHz and the process is readily achieved using TTL logic integrated circuits. (A single-counter version of the converter would require a clock rate of about 550 MHz which would be very difficult with currently-available logic devices.) The coding accuracy achieved in practice is such that the measured signal-to-noise ratio is usually within 1 dB of the theoretical.

The digital signals from the a.d.c.s are applied to parity generator circuits and multiplexed, together with the framing and signalling information.

At the receiving terminal timing and framing information is extracted and used to control circuits which de-multiplex the signal. Each channel digit stream is effectively applied to a separate d.a.c. which uses the same groups-units technique as in the a.d.c. with the group and unit ramps arranged to have slopes in the ratio of 64 : 1. Initially parity is checked and if correct the conversion is allowed to proceed; if the parity is incorrect the previous sample is repeated by effectively inhibiting the conversion process.

Automatic monitoring of the coding, multiplexing, transmitting, de-multiplexing and decoding processes is provided. At the transmitting end the monitor is sequential and dwells on each audio channel in turn for 200 ms. During this time 256 checks are made by sampling the audio input and coding with a 5-bit a.d.c. The output of this a.d.c. is then compared with the 5 m.s.b.s of the appropriate identical word appearing in the bit stream at the output of the multiplexer. The monitor registers a fault if more than 50% of the checks for any channel violate the identity. At the receiving end a similar process checks for faults in the de-multiplexing and decoding equipment. Faults in the transmission circuits are detected by counting the rate of parity failures at the receiving terminal. The circuit is deemed to be unusable when parity failures equivalent to a sustained error rate of 1 in 10^4 or greater are present.

3.2. Spectral characteristics and digit format

The overall bit rate for the existing sound multiplex system was chosen to be 6·336 Mb/s, a rate that had been proposed as an International Standard. This enables the bit stream to be sent over a bearer circuit, usually a microwave link designed to carry 625-line monochrome television signals. For transmission purposes the individual pulses are shaped to be of cosine-squared form with a duration of 316 ns and with a spacing between pulses of 158 ns. There is therefore negligible energy in the transmitted spectrum above 6·336 MHz.

In practice the bearer circuit bandwidth can be reduced in exchange for a worsening of the system resistance to noise and distortion. With a circuit bandwidth of 6·336 MHz the system is satisfactory without error protection down to a peak pulse-signal/r.m.s. white noise ratio of about 20 dB; for a 4·5 MHz bandwidth this value rises to about 22 dB.

To render the spectral content of the transmitted signal to be zero at d.c. and very low frequencies, alternate digits in each channel sample word are inverted. This enables a.c. coupling to be used at the receiving terminal which then improves the reliability of pulse detection when d.c. 'bumping' effects occur during anomalous propagation conditions.

At the bit rate of 6·336 Mb/s and with a sampling frequency of 32 kHz and 13 channels there are 198 bits per frame. The 198 bits comprise 13 audio channels each of 14 bits (including the parity bit), 11 bits for synchronisation and 5 bits for auxiliary purposes. The latter are used for addressing and switching equipment simultaneously at all decoder sites or to send a message to an individual site. Typical uses of the data messages are for mono/stereo switching and transmitter remote control.

4. BIT RATE REDUCTION BY DIGITAL COMPANDING

4.1. General

As discussed in Section 2 of this paper 13 bits per sample with dither are required for a linearly-coded high-quality sound signal of 15kHz bandwidth, giving a basic bit rate of 416 kb/s per channel. In 1971 an investigation was commenced into various methods of reducing the bit rate without impairing the subjective quality, the ultimate target being that at least 6 high-quality channels should be accommodated in the European standard first-order multiplex rate of 2·048 Mb/s. Initially syllabic companding and other analogue forms of noise-reducing techniques were contemplated but were not favoured because, although they potentially offered a saving of about 2 bits per sample,[4] the elaborate analogue instrumentation was not attractive on grounds of cost and reliability. Attention was instead concentrated on instantaneous companding and a variant described as 'near-instantaneous', it being possible to achieve the required signal processing with these methods entirely in the digital domain. Experiments were first carried out using a simple simulation technique to assess the degree of programme-modulated noise likely to be introduced by this type of companding.[5] From the encouraging results obtained further work proceeded with the construction of a real digital compandor capable of operating with various companding laws both in the instantaneous and near-instantaneous modes. The remainder of this paper is concerned with a description of this experimental compandor and the results obtained with it, together with a method of multiplexing six high-quality sound channels at a bit rate of 2·048 Mb/s using near-instantaneous companding.

4.2. Near-instantaneous companding

The near-instantaneous companding technique has a similarity to the instantaneous discontinuous technique proposed by Bartlett and Greszczuk.[6] In the instantaneous case, a limited number of digits is transmitted per word, together with a scale factor indicating the significance or weight of the transmitted sample word. With the near-instantaneous technique, the scale factor is transmitted less frequently, say once every 30 signal samples, its value being determined by the peak signal level during that group of samples.

A practical implementation of a near-instantaneous digital compandor is shown in Fig. 1. The analogue signal is pre-emphasised before application to a 13-bit a.d.c. The peak value in a group of sample words is used to control the digital compressor. Delay is necessary to enable the compressor range to be set at the beginning of the group of sample words. At the decoder, the scale-factor word controls the digital expandor. After the 13-bit d.a.c. the analogue signal is de-emphasised.

The experimental compandor is illustrated in Fig. 1, and it includes manual switching facilities for altering both the rate at which scale-factor words can be transmitted and the number of segments in the companding law; this can be varied from 1 segment (i.e. linear, 13 bits) to seven segments in each quadrant.

The pre-emphasis characteristic used in the experimental compandor was adjusted to reduce the level of low frequencies rather than to enhance the level of high frequencies for reasons given below.

4.3. Performance

Level-dependent noise is inevitably introduced by compandors and althoughiit is possible to predict noise characteristics for any given system it is usually necessary to perform listening tests to assess the degree of impairment actually caused.

Fig. 2 gives theoretical quantising-noise characteristics for a 4-segment near-instantaneous compandor; perfect instrumentation is assumed and the effect of pre- and de-emphasis is neglected. For low signal levels, the r.m.s. quantising noise is similar to that from the existing BBC 13 bit linearly-coded p.c.m. system. At high signal levels the noise is almost constant as a proportion of the signal, the exact r.m.s. value of the noise depending on the analogue signal frequency. The broken line is for very low frequencies (when the action of the compandor is virtually instantaneous), whereas, for signals above about 530 Hz, the solid line applies. For comparison, the corresponding characteristic for an instantaneous A-law compandor (10 bits transmitted) is also shown in Fig. 2. The difference between the characteristics at low signal levels arises from the 14-bit initial coding that would be required for this A-law compandor. The advantage shown for the near-instantaneous compandor at high levels is consistent with the results of subjective tests.

Subjective tests were done using low-noise recordings of critical programme material with listeners experienced in assessing the performance of high-quality sound systems. An exhaustive search to find critical test material showed that programme-modulated quantising noise was particularly evident on the five-note piano scale C′, D′, E′, F′, G′; this was therefore used for most of the subjective tests. In the tests the listeners compared two presentations, one processed through the 13-bit linear system and the other through the digital compandor, each presentation consisting of 3 sequences of the five-note piano scale. An extract from the results obtained is given in Table 1 which shows that only 30% of the experienced listeners could detect the use of the near-instantaneous compandor when a 4-segment law was used; with more segments this proportion rises steeply. Moreover, for an approximately equal number of bits transmitted, programme-modulated noise was more audible with instantaneous than with near-instantaneous companding.

Subjective tests were also done to determine the relative merits of two different pre-emphasis characteristics, the standard 50 μs characteristic as used in the existing BBC p.c.m. system and a characteristic having the shape recommended by the CCITT[7] for carrier systems. These tests showed that the latter was superior to the 50 μs characteristic in reducing the audibility of programme-modulated noise. The results given in Table 1 apply for CCITT pre-emphasis with the gain at 15 kHz set to 4 dB.

TABLE 1

System under test	Approximate proportion of listeners who could distinguish between 13 bits linear and system under test
1. Near-instantaneous, 30 words delay, 5 segments (9·10 bits transmitted)	60%
2. Near-instantaneous, 30 words delay, 4 segments (10·07 bits transmitted)	30%
3. Instantaneous (i.e. no delay) 6 segments, (10 bits transmitted)	90%

The conclusion from the investigation using the experimental compandor was that the basic bit rate for digital transmission of a high-quality sound signal could be reduced to about 320 kb/s using the near-instantaneous technique (System 2 in Table 1).

4.4. Effect of transmission errors on companded p.c.m. signals

Further work was carried out to determine the degree of protection against transmission errors required by digitally companded signals. In general such signals were found to be more immune to errors than linearly-coded signals because the most significant digits are transmitted only when a high-frequency high-level signal is present. Compared to the linearly-coded signal, which uses 1 parity bit to check the 5 m.s.b.s of each 13-bit sample word it was concluded that for the System 2 near-instantaneous companded signal only the two m.s.b.s of the sample word, together with the scale-factor word needed protection.

Techniques for error concealment similar to those outlined in Section 2.3 for linearly-coded signals have been found effective for digitally companded signals. Errors in sample words are concealed by repeating the previous correctly-received sample and errors in the scale-factor words are concealed by repeating the previous correctly-received scale-factor.

4.5. Multiple-channel companded signals

As stated in Section 3.1. the objective in the bit-rate reduction investigation was to develop a system which would enable at least 6 high-quality sound channels to be transmitted in a standard 2·048 Mb/s multiplex. With the System 2 type of near-instantaneous companding, 10 bits per sample, plus a 2-bit scale-factor word for every 30 samples, are transmitted. At a sampling rate of 32 kHz, the gross bit rate per channel is 322·133 kb/s. Hence six such channels require 1·9328 Mb/s which, in a standard 2·048 Mb/s multiplex, leaves 115·2 kb/s for synchronisation, error protection and signalling. Frame synchronisation can be achieved using a 12-bit pattern distributed over 1920 bits. For a satisfactory performance in the presence of errors, only one parity check bit need be used for the two most significant digits of two sample words (taken from different channels) and one parity check bit for the scale-factor word.

Table 2 sets out one possible frame structure for transmitting 6 channels in 2·048 Mb/s. For simplicity, only one-sixth of the 1920 bit frame is shown. The complete frame would contain the full 12-bit framing pattern and six scale-factor words (one for each channel). No provision for signalling information is shown in Table 2. However, a signalling rate of about 1 kb/s could be derived either from the frame synchronisation bits or the sample-word error-protection parity check bits; exploiting these bits for signalling purposes would respectively either slightly increase the mean re-framing time or slightly degrade the performance of the system in the presence of transmission errors. On balance, the use of framing bits for signalling is preferred since the instrumentation required would be relatively simple.

TABLE 2

30 x 10-bit channel words (5 from each channel)	300 bits
15 error-protection bits for the channel words	15 bits
Scale-factor word	2 bits
Error protection for the scale-factor word	1 bit
Framing pattern (1/6 of total 12 bits)	2 bits
Total	320 bits

4.6. Instrumentation

Fig. 1 shows the basic functions required for a single-channel near-instantaneous compandor. A 6-channel compandor could consist of six such compandors with suitable arrangements for multiplexing the outputs into a single serial bit stream. However an alternative arrangement would be first to multiplex the six linearly-coded signals into a single serial bit stream and then apply this to one digital compressor circuit. Similarly, only one expandor circuit would be required at the receiving terminal thus reducing the cost per channel of companding. Moreover, because fewer integrated circuits would be required the reliability would be higher.

To retain the advantage described in Section 3.2 for the 13-channel system which has only one limiter per channel to protect both the p.c.m. system and the f.m. transmitters from large signal overloads, 50 μs pre-emphasis would still be inserted ahead of the limiter. The limiter would then feed a 50 μs de-emphasis network followed by the CCITT pre-emphasis characteristic required, as described in Section 4.3, to render programme-modulated noise inaudible. The gain adjustment of the CCITT pre-emphasis is such that overload protection is still afforded by the preceding limiter whilst the low-level signal-to-quantising noise ratio is maintained.

5. CONCLUSIONS

The lack of suitable forms of transmission circuits was a major technical barrier to the extension of stereo broadcasting in the U.K. The paper describes a 13-channel p.c.m. system which has been developed by the BBC to overcome this obstacle. The system is in service for distributing high-quality sound signals over monochrome 625-line television microwave links.

The paper also describes the results of an investigation by the BBC into methods of reducing the bit rate required for high-quality sound signals. A near-instantaneous digital companding technique has been evolved which will enable six channels to be multiplexed into a standard 2·048 Mb/s bit stream without noticeable impairment of sound quality. Moreover, it is considered that 4 such multiplexes could be transmitted over the existing 13-channel bearer circuits to provide 24 channels should this increase in channel capacity ever be required.

It is felt that digitally companded p.c.m. sound signals will be mainly of interest for transmission purposes where bit rate is at a premium. Such signals might be used in digital recording where again a reduction in the required bit rate could prove advantageous. However, within possible future 'all-digital' studio centres the signals are probably best left in linearly-coded form prior to transmission since such signals are more amenable to processing operations such as switching, fading and mixing.

6. ACKNOWLEDGEMENT

The authors are indebted to the Director of Engineering of the British Broadcasting Corporation for permission to publish this paper.

7. REFERENCES

1. SHORTER, D.E.L. *and* CHEW, J.R. 1972. Application of pulse-code modulation to sound-signal distribution in a broadcasting network. *Proc. IEE*, 1972, **119**, 10, pp. 1442 – 1448.

2. CROLL, M.G. 1970. Pulse-code modulation for high-quality sound signal distribution: quantising distortion at very low signal levels. BBC Research Department Report No. 1970/18.

3. MITCHELL, G.R. *and* MOFFAT, M.E.B. 1973. Pulse-code modulation for high-quality sound-signal distribution: subjective effect of digit errors. *BBC Engineering*, November 1973, No. 96, pp. 27 – 32.

4. DOLBY, R.M. 1967. An audio noise reduction system. *Journal of the Audio Engineering Society*, October 1967, Vol. 15, No. 4, pp. 383 – 388.

5. OSBORNE, D.W. 1972. Digital sound signals: further investigation of instantaneous and other rapid companding systems. BBC Research Department Report No. 1972/31.

6. BARTLETT, G.J.C. *and* GRESZCZUK, J. 1964. Companding in a p.c.m. system. Symposium on Transmission Aspects of Communications Networks. London, IEE, 1964, pp. 183 – 186.

7. CCITT Red Book, 1957, III, p. 273.