BBC RD 1973/41

RESEARCH DEPARTMENT REPORT

DIGITAL SOUND SIGNALS: bit-rate reduction using an experimental digital compandor

D.W. Osborne, C.Eng., M.I.E.E.
M.G. Croll, B.Sc., A.R.C.S.

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION December 1973

BBC RD 1973/41
UDC 534.86
621.37.6.56

DIGITAL SOUND SIGNALS: BIT-RATE REDUCTION USING AN EXPERIMENTAL DIGITAL COMPANDOR

D.W. Osborne, C.Eng., M.I.E.E.
M.G. Croll, B.Sc., A.R.C.S.

Summary

A previous investigation, using a simulation technique, indicated that considerable bit-rate reduction for the transmission of pulse-code modulated high-quality sound signals should be possible with a novel form of digital companding using a 'near-instantaneous' technique. A description is given of the design principles of an experimental model of a near-instantaneous digital compandor, which has been constructed. Subjective tests with this equipment confirm that the bit rate per channel can be reduced from 416 kb/s, required for the linearly-coded system now in service in the BBC, to about 320 kb/s, without significant impairment of programme quality.

issued under the authority of

P. Rainger

Head of Research Department

Research Department, Engineering Division
BRITISH BROADCASTING CORPORATION

December 1973

(EL-83)

DIGITAL SOUND SIGNALS: BIT-RATE REDUCTION USING AN EXPERIMENTAL DIGITAL COMPANDOR

DIGITAL SOUND SIGNALS: BIT-RATE REDUCTION USING AN EXPERIMENTAL DIGITAL COMPANDOR

D.W. Osborne, C.Eng., M.I.E.E.
M.G. Croll, B.Sc., A.R.C.S.

1. Introduction

The values of the essential parameters for a high-quality sound-signal pulse-code modulation (p.c.m.) multiplex transmission system have already been determined[1]. This system has a capacity of up to 13 channels and is now in service with the BBC for the distribution of programmes from the studios to the main v.h.f./f.m. broadcasting transmitters. The system uses linear coding with a quantising accuracy of 13 bits per sample at a sampling rate of 32 kHz, and thus it has a basic bit rate of 416 kb/s per channel. A parity check bit is included with each sample word for the purposes of transmission-error protection, bringing the overall bit-rate per channel up to 448 kb/s.

In the interests of economy, it is desirable to increase the programme-carrying capacity of p.c.m. communication networks, and, to this end, various techniques for reducing the number of bits per sample have been investigated. Instantaneous companding is one such technique, and it is potentially applicable to the digital transmission of high-quality sound-signals; it has been fully described in a previous report[2] on a preliminary investigation into the possible benefits to be obtained. The same report also covered a novel variant known as 'near-instantaneous' companding. This preliminary investigation used a simulation technique in which random white noise was used to represent the quantising noise which occurs in an actual digital system. The results suggested that a worthwhile bit-saving is achieved by companding and a decision was made to continue the work by constructing an experimental digital compandor. The present report outlines the principle of operation of this compandor, describes the tests carried out and gives the theoretical and practical results obtained for the systems of greatest interest.

2. Companding

2.1 General

Companding is a technique for improving the signal-to-noise ratio in sound-signal processing and transmission systems. In general there are two types of companding action — 'syllabic', a form often employed in analogue systems, and 'instantaneous', usually applied to digital systems. The main features of these two types of companding are discussed in detail in Reference 2 but for completeness an abbreviated description is given here.

In a syllabic compandor a compressor, which reduces signal gain with increasing signal magnitude in accordance with a predetermined characteristic or 'law', is used at the sending end of a system to allow smaller signals to be transmitted at a higher relative level. An expander, having a characteristic complementary to the compressor, is introduced at the receiving end to restore the overall gain to a constant value for all signals. The rate of change in gain of the compressor must be restricted so as to avoid a significant increase in the bandwidth of the transmitted signal, required by the additional products which are generated in the gain-modulation process, and a syllabic rate of change is suitable. Compandors of this type are used to reduce the effects of noise incurred in the intervening signal-processing or transmission path between the compressor and expander.

Instantaneous companding is based on a similar principle to that described for syllabic companding, except that no restriction is placed on the rate of compression and expansion. Instead the action is immediate and is equivalent to passing the analogue signal through a non-linear transfer characteristic for compression and a complementary non-linear characteristic for expansion. Because of the products generated by the compression process, and the consequent increase in bandwidth required to preserve the compressed signal waveform, instantaneous companding is not normally practicable in analogue transmission systems. However, in a p.c.m. system, instantaneous companding is practicable, because the required transmission channel bandwidth is independent of processes such as compression carried out before transmission of the digital signals.

The object of companding in a p.c.m. system is to make efficient use of the available number of quantising levels, i.e. to reduce the bit rate required for transmitting the signal, as distinct from companding in an analogue system which is used for minimising the effects of noise introduced in the transmission path. Instantaneous companding laws are such that the quantum steps, i.e. the intervals between quantising levels, are smaller for signals of low magnitude and larger for signals of high magnitude. For a given number of quantising levels, the amount of quantising noise imposed on small signals can thus be reduced at the price of imposing more noise on large signals.

An important problem with syllabic companding is that, in the expander, the level of the noise incurred during transmission is subjected to the same gain or attenuation as that of the signal level. The resultant subjective impairment which may be audible is programme-modulated noise, sometimes described as 'hush-hush' noise. A similar problem arises with instantaneous companding in p.c.m. systems, resulting in programme-modulated quantising noise.

Analogue syllabic compandors can be used for improving the performance of a p.c.m. system[3], but in general the complex circuits which are required, especially if the stringent matching requirements of the right- and left-hand channels of a stereophonic transmission have to be met, render them unattractive from the points of view of cost, reliability and stability. In contrast, digital instantaneous compandors do not suffer from these disadvantages.

2.2 Companding in digital systems

2.2.1 Companding laws

In theory the non-linear functions required for instantaneous companding in digital systems may be obtained either by an analogue compressor followed by a linear analogue-to-digital converter (a.d.c.), or by a linear a.d.c. followed by a digital compressor. In practice the use of the former technique has been found impracticable for high-quality sound p.c.m. systems because it is necessary to match the compression and expansion characteristics to a very high order of precision to reduce distortion to an acceptable level[4]. However, no such difficulty arises with the latter technique, which in this report will be termed 'digital companding', provided that a companding law having straight-line segments with slopes related by a factor of 2 is acceptable.

In describing a companding law, it is sufficient to specify the compression law since it follows that the associated expansion law must be precisely complementary in order that the overall transfer characteristic be linear. Moreover, since companding laws are skew-symmetrical through the origin, only one quadrant need be shown in each case.

One example of this form of law, shown in Fig. 1, is the 7-segment* law, known as the A-law[5], used for p.c.m. telephony systems to obtain a near-constant signal-to-quantising noise ratio over a wide range of input signal levels. In Fig. 1 the slopes of the segments are indicated by the number of bits per sample which would be required

*The number of segments given here is for one quadrant of the companding law. The well-known A-law is often described as having 13 segments which refers to the number of segments in the positive and negative quadrants taken together.

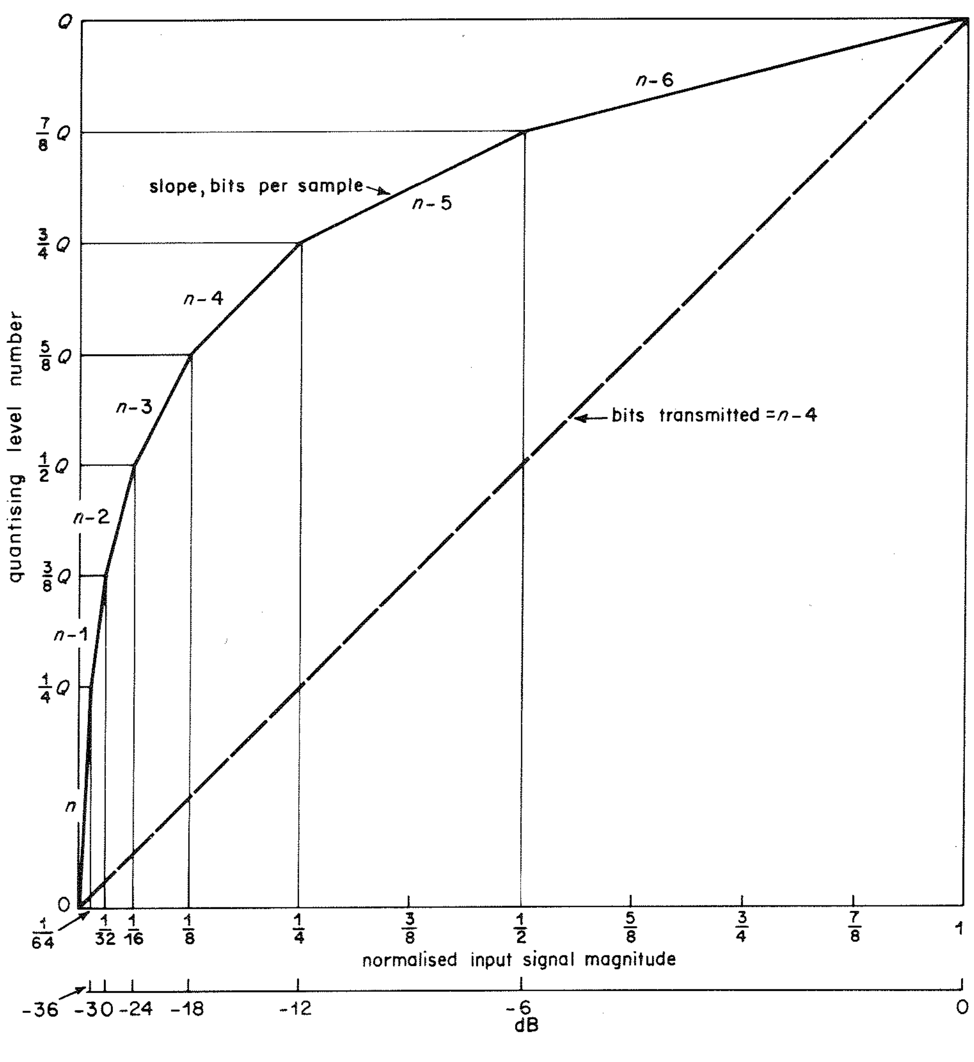

Fig. 1 - 7-segment A companding law: positive quadrant only shown
Q represents maximum number of digital codes available in positive quadrant; Q is a power of two

if that slope were used in a linear p.c.m. system to code the whole range of the signal. The reference value, n, corresponds to the initial slope of the A-law and hence to the bits per sample of a system having the same quantising noise level for very low-level signals. The number of levels, Q, required for transmission using the A-law is $2^{(n-4)}$, requiring n–4 bits per sample. A linear system with Q levels (see dashed line in Fig. 1) would be subject to 24 dB more quantising noise because the coarseness of quanta is increased by 2^4, i.e. by a factor of 16. The linear system would require 4 more bits per sample to achieve the same quantising noise background level. The subjective effect of noise modulated by programme level, however, becomes an important factor that reduces the net advantage of the A-law.

From the previous work[2], it appears that a better performance than that given by an instantaneous A-law compandor could be achieved in two ways:

(a) by the use of a less elaborate law which secured better balance between background noise and programme-modulated noise in terms of the subjective effect.

(b) by the use of the companding law in the near-instantaneous mode as explained in Reference 2.

We may consider a four-segment law in terms of transmission by means of a digital number corresponding to the discontinuous (and multi-valued) function as shown by characteristic (a) in Fig. 2[6], together with a further 2-bit word transmitted with each sample value to signify which segment applies.

However, the action may be modified with advantage in bit-saving by selecting one of the laws (c) in Fig. 2 to code a batch of, say, 24 successive samples, choosing the steepest slope that will just accommodate the largest sample value in a batch. The 2-bit scale-factor word now needs

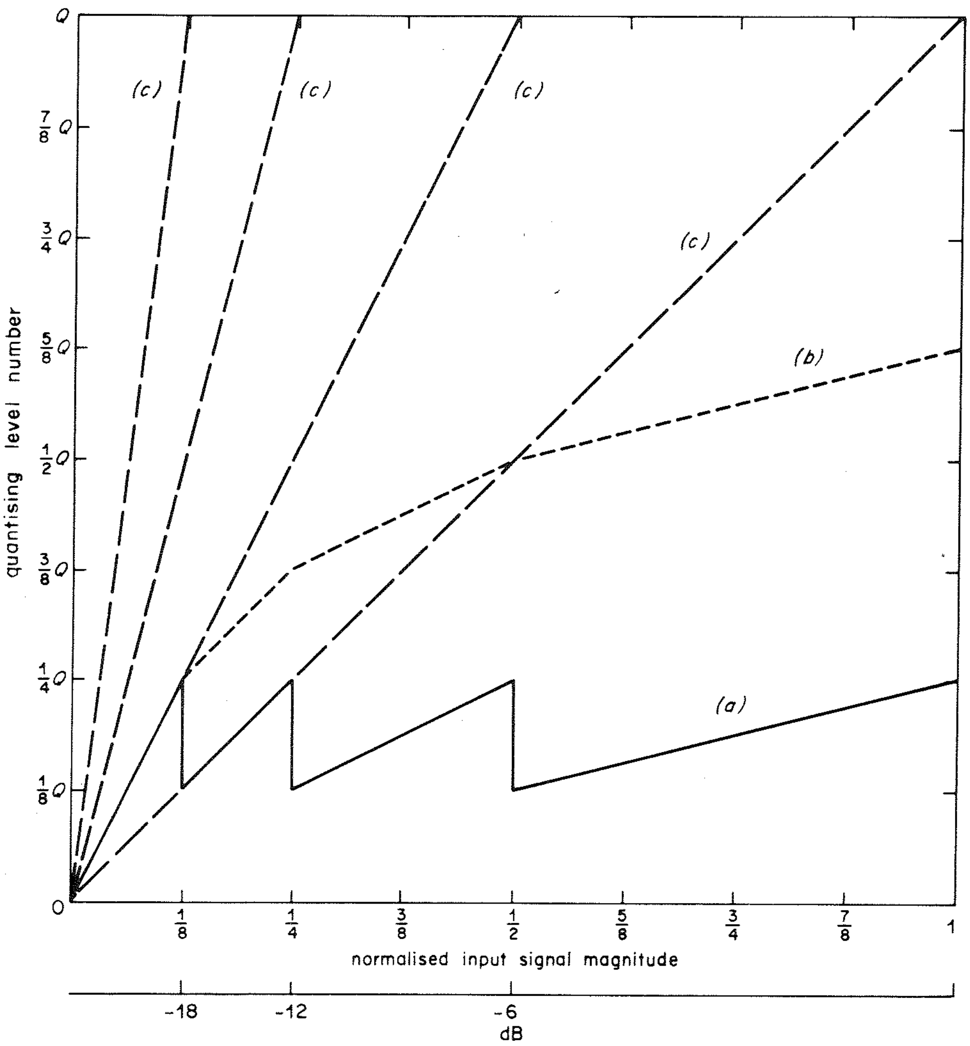

Fig. 2 - Comparison of instantaneous and near-instantaneous companding laws with a four-segment law: positive quadrant only shown

Q represents maximum number of digital codes available in positive quadrant

(a) instantaneous companding law actually transmitted (including 2-bit scale-factor word)

(b) instantaneous companding multi-segment law equivalent to (a)

(c) laws followed with near-instantaneous companding mode of operation of same law (ignoring occasional 2-bit scale-factor word).

transmitting only once per batch. This is the variant called 'near-instantaneous' in Reference 2, and offers the prospect of considerable bit-saving[7].

Fig. 2 illustrates the advantage to be gained from the near-instantaneous technique for a 4-segment law. When the companding action is changed from instantaneous to near-instantaneous, only a small bit rate is required for the scale-factor word, and almost the entire number of quantising levels is available for each segment of the law. Hence, the slopes (c) of the law radiating from the origin become four times the slopes for the corresponding segments of the instantaneous companding law. The resulting benefit is an improvement in quantising accuracy by a factor of four. This benefit is slightly offset since, except for low audio-frequencies, the companding action tends to follow the envelope of the analogue signal; the mean quantising noise output from the decoder must therefore of necessity be greater than for instantaneous companding (other things being equal) in which the companding action follows every detail of the analogue signal waveform. It is of interest to note that, neglecting the occasional scale-factor word, the number of bits required to be transmitted for the near-instantaneous companding law corresponds to the segment of least slope, i.e. the least number of bits per sample.

2.2.2 Pre- and de-emphasis

Pre- and de-emphasis is often used in signal transmission and processing systems to obtain an improvement in the steady-state received signal-to-noise ratio. An example is in f.m. broadcasting, in which a standard characteristic (defined by a network having a time constant of 50 μs for European transmissions[8]), is used to enhance the level of high-frequency components before modulation at the transmitter. A complementary de-emphasis network at the receiver ensures a uniform overall response. In this application, a reduction in perceived noise is effected because high-frequency noise components, to which the ear is particularly sensitive, are reduced by the de-emphasis characteristic in the receiver.

Pre- and de-emphasis is important in companded p.c.m. systems, not for reducing steady-state noise, but for reducing programme-modulated quantising noise. Previous simulation tests[2] had shown that the use of pre- and de-emphasis having the shape of the CCITT[9] characteristic reduced the audibility of programme-modulated noise to an extent that, on average, about one less bit would be required than in the case of no pre- and de-emphasis. In this application the pre-emphasis characteristic is set to enhance the level of high frequencies; Fig. 3 shows the CCITT pre-emphasis curve with the insertion loss at 15 kHz set to unity, as used for the simulation tests. The signal amplitude distribution after pre-emphasis is now such that signal excursions into the regions of larger quantum steps occur only for high-frequency components of above-average level, which are comparatively rare. Moreover, the increased noise arising from the coarser quantising at these times is effectively masked by the programme content.

It is possible that a pre-emphasis characteristic other than that recommended by the CCITT might be more effective in reducing programme-modulated quantising noise in companded p.c.m. systems. However, from a limited theoretical and practical investigation, it is thought unlikely that any substantial additional benefit would be obtained from another shape of curve and this is supported by results obtained by other workers[10].

2.2.3 Theoretical peak signal-to-quantising noise ratio for companded p.c.m. systems

In this Section the results of calculations of the quantising noise in companded p.c.m. systems are discussed; details of the method used for the calculations are given in Appendix 10.1.

The theoretical quantising noise characteristic for the 4-segment near-instantaneous companding law is shown in Fig. 4, curve (a). The characteristic was calculated assuming perfect instrumentation and, for simplicity, the effect of pre- and de-emphasis was neglected. The full-line curve applies for a frequency greater than about 1 kHz for which, because of the 0·5 ms intervals between transmissions of the scale-factor word, segment changes in the companding law cannot occur during a half-cycle of the analogue sinusoidal input waveform. The dashed curve

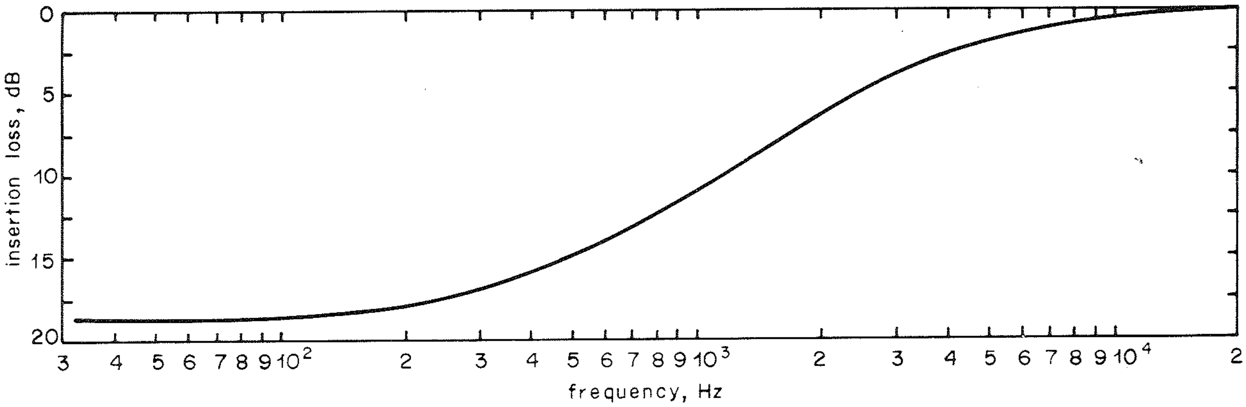

Fig. 3 - CCITT pre-emphasis characteristic as used in simulation tests

shows the characteristic for a low frequency sinusoidal input for which the compandor action can change segment during a period of the sinusoid and thus become effectively instantaneous.

For comparison Fig. 4 also shows the quantising noise characteristic for the instantaneous A-law, curve (b), calculated on the same basis as curve (a). As can be seen, the mean quantising noise for the near-instantaneous compandor is about 9 dB lower than that for the instantaneous A-law for the range of input signal above about −18 dB relative to the maximum. For input signals below about −30 dB relative to the maximum the former system has the higher noise level. The difference* between curves (a) and (b) for low-level signals arises because of the theoretical improvement between the signal-to-quantising noise ratio for the 14 bits-per-sample coding (without dither) assumed for the central segment of the A-law and that for the 13 bits per sample (with dither) assumed for the near-instantaneous law. However, for all signals lower than − 18 dB relative to maximum, the theoretical r.m.s. noise level for the near-instantaneous law is 81 dB below maximum peak signal, a figure considered to be sufficiently low.

*The difference is 7·5 dB, i.e. 6 dB due to 14-bit coding compared with 13-bit coding, plus a further 1·5 dB due to the use of dither[1] for the 13-bit coding.

The theoretical superiority of the near-instantaneous companding law is reflected in the results of the subjective tests described later in Section 4.3 of this report.

3. Instrumentation of experimental digital compandor

Fig. 5 shows the basic principle of the experimental single-channel digital compandor. Although primarily intended to verify the optimum near-instantaneous system, the equipment was arranged to be capable of operating with more than one companding law and, if required, in the instantaneous mode. After pre-emphasis the audio analogue signal is initially converted to 13-bit p.c.m. form using a linear coder and 'dither' to mask granular distortion at low signal levels. As discussed in Reference 1, 13 bits per sample with dither is considered sufficient for low-level signal-to-noise requirements for a linearly coded system with up to 4 coders and decoders (codecs) in tandem. It was therefore decided that 13 bits per sample should also be the maximum quantising accuracy necessary in any companded p.c.m. system for high-quality

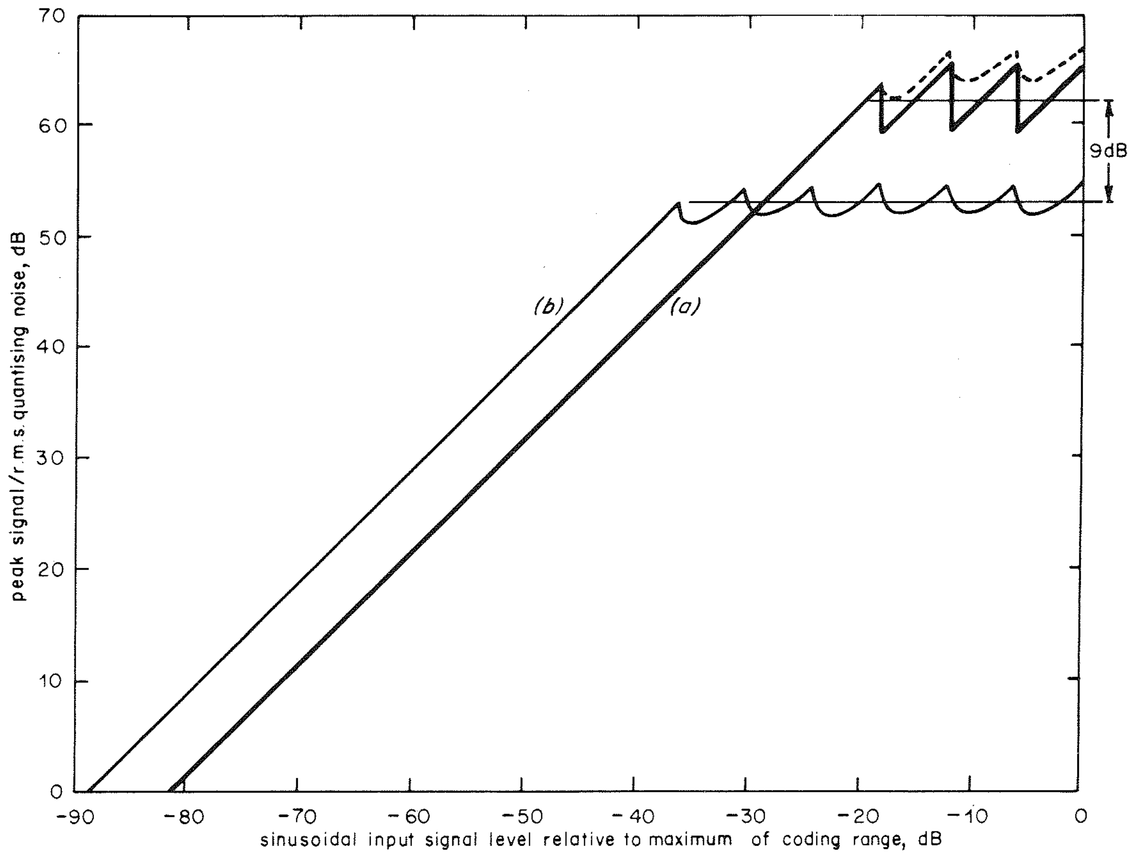

Fig. 4 - Calculated ratio of peak signal to r.m.s. quantising noise

(a) near-instantaneous four-segment discontinuous law, 10·125 bits transmitted

------ low frequencies ——— *frequencies above about 1 kHz*

(b) instantaneous A-law, seven segments, 10 bits transmitted

sound signals. (This decision was also partly determined by the known increased difficulty and costs of achieving and maintaining a quantising accuracy of greater than 13 bits per sample). A symmetrical binary code (i.e. a code in which one bit represents the sign, and the remaining bits the modulus of the analogue signal sample) is employed since this is convenient for the subsequent digital companding processes.

Companding is achieved using variable-length shift registers operating on the serial bit stream, and laws having from 1 (i.e. linear coding) to 7 segments in each quadrant can readily be set up by switches on the compressor. The laws are such that the changes in slope of the segments are related by a factor of two with 'breakpoints' occurring at 6 dB steps of input signal magnitude. Such laws are convenient to instrument in the digital domain; it is doubtful whether a worthwhile advantage would be gained by departing from this form of law.

When the compandor is operating in the instantaneous mode, the magnitude of each digitally-coded signal sample is measured, the shift-register compressor is then set according to the segment required for a particular companding law, and the corresponding scale-factor word (required for setting the digital expander) is multiplexed into the resultant serial bit stream for transmission to the decoder.

In the near-instantaneous mode, the peak magnitude in a group of digital samples is measured and the appropriate range set in the compressor as before. A digital delay preceding the compressor is necessary to enable the compressor range to be set at the beginning of the group of sample words. In this mode the compressor control and scale-factor words must be stored for the period corresponding to the group of samples. In the experimental compandor the action could be adjusted to cover a group extending up to 30 signal samples. The maximum number of samples in a group is governed by the subjective effect of prolonging the periods of increased quantising noise associated with programme-signal peaks; 30 samples, corresponding to a time interval of just under 1 ms, was considered to be the maximum likely to be of interest.

A 14-bit digital-to-analogue converter (d.a.c.) is shown in the decoder in Fig. 5. This is required to compensate ideally for 'rounding' error effects in the digital compression process, but tests showed that the alternative use of a 13-bit d.a.c. did not result in any perceptible degradation in performance. It is thought likely that the 'dither' applied to the original 13-bit coding of the signal nullified the slight theoretical advantage of 14-bit decoding. (See Appendix 10.2 for a more detailed discussion of 'rounding' compensation).

In the experimental compandor it was found convenient to use a 16-bit word format since this fitted in well with the many TTL integrated-circuits which are orientated towards a 16-bit configuration.

No severe practical problems were encountered with the equipment but in the analogue parts of the system, care was necessary to reduce hum and other unwanted signals to an acceptably low level. The use of pre- and de-emphasis as described in Section 2.2.2 tended to exacerbate this problem.

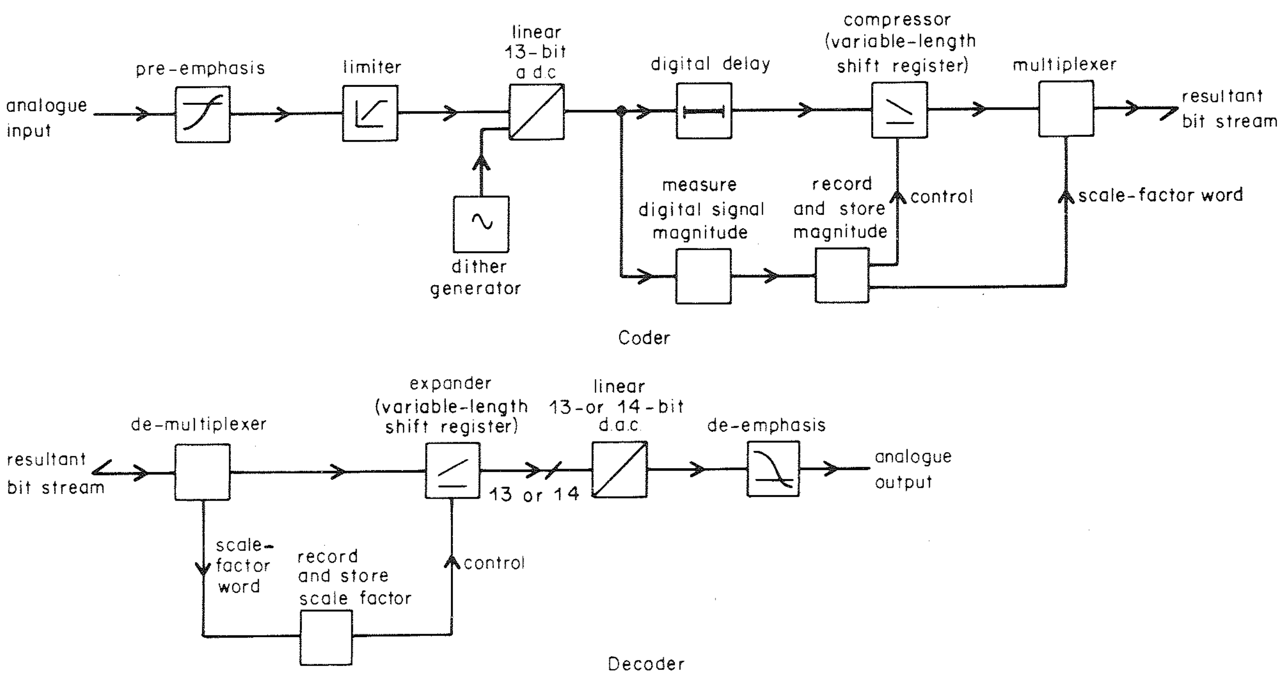

Fig. 5 - Experimental single-channel digital compandor

4. Subjective tests

4.1 Choice of test programme material

Preliminary tests were made with conditions corresponding as closely as possible to those used in the simulation tests described in Reference 2. In these, a short tape-recorded piano excerpt (from Fantasia in D minor, by Mozart) was re-recorded repetitively on a second tape, so that listeners could more easily assess programme-modulated quantising noise for a given condition by hearing the same excerpt many times. The results of these preliminary tests with the digital compandor were in reasonable agreement with those predicted from the simulation tests. However, it was realised that because of the nature of the impairment, and its very small degree which listeners would be asked to assess, the original test material was required to be as free from noise and distortion as possible. Moreover, it was felt that, before formal subjective tests were conducted, various types of programme should be tried to determine which of them provided the most stringent test, i.e. which of them was the most likely to reveal programme-modulated quantising noise. Accordingly, special low-noise recordings (using Dolby 'A' noise-reducing equipment) were prepared using solo instruments which were thought likely to be critical. The instruments were piano, glockenspiel, xylophone, tubular bells and trumpet. Recordings of speech were also made since Hessenmüller[11] had found that there had been a tendency in his tests with instantaneous A-law companding for listeners to prefer unprocessed (i.e. direct analogue) speech signals.

Initial pilot tests with the compandor, confined to a few critical listeners and with the various types of programme described above, showed that the piano was the most stringent test-programme item and that programme-modulated quantising noise was particularly evident with a five-note piano scale, C′, D′, E′, F′, G′. Surprisingly, programme-modulated noise was barely audible on the glockenspiel, an instrument having high-level, high-frequency components which caused the compandor to operate in the coarsest quantising range. Presumably in this case the quantising noise was masked by the programme content as discussed in Section 2.2.2. From the pilot tests it was therefore decided that only the very critical five-note piano scale should be used for formal subjective tests.

4.2 Subjective test conditions and procedures

Tests were conducted principally to determine the effect of three parameters on the overall objective, which was to arrive at an optimum bit-rate reduced p.c.m. system having a performance virtually indistinguishable from a linearly-coded 13 bit per sample system which uses dither signals. The latter, being the system now in service with the BBC for high-quality sound-signal distribution, was taken as a reference system. The three parameters were:

(a) The number of segments in the companding law (this related directly to the number of bits transmitted.)

(b) The time interval between transmissions of the scale-factor word in the near-instantaneous mode.

(c) The gain setting of the pre-emphasis curve.

A first series of tests was conducted with 20 listeners, ten of whom were technical staff experienced in the assessment of reproduced sound quality. A high-quality monitoring loudspeaker (BBC type LS5/5) was used in a room having acoustics designed to simulate good domestic listening conditions; the mid-band reverberation time was about 0·3s and the room volume 85 cubic metres. The background noise in the listening room was extremely low to ensure critical listening conditions, and the maximum sound-pressure level for listening ranged from 80 to 90 dBA.

Each of the 20 listeners was present individually in the listening room for identical sets of tests. In these tests a number of pairs of presentations, first A and then B, were made, and the listener asked to record any preference for either A or B according to the 7-grade comparison scale shown below.

Comparison	Grade
A much better than B	+3
A better than B	+2
A slightly better than B	+1
A = B	0
B slightly better than A	−1
B better than A	−2
B much better than A	−3

The pairs of presentations were made in a deliberately jumbled order so as to avoid a tendency for A to be consistently better than B or vice-versa and were intended to elicit information on the three parameters given above. The conditions presented covered:-

(d) Companding laws having 4,5,6 and 7 segments.

(e) Scale-factor transmission intervals of one word (i.e. corresponding to instantaneous companding), 16 words (0·5 ms) and 24 words (0·75ms).

(f) Pre-emphasis curve set such that the gain at 15 kHz was 3 dB, 6 dB, and 9 dB*.

It was found that some of the mean results of the first series of formal subjective tests tended to be inconsistent with expectations. For example, the 5-segment near-instantaneous companding law (corresponding to about 9 bits transmitted per sample) was assessed as somewhat better than the 4-segment version (corresponding to about 10 bits transmitted per sample). Several reasons were

*Tests were not made with the gain at 15 kHz set to 0 dB (as shown in Fig. 3) because with this setting the low-level signal-to-quantising noise ratio would be unacceptably degraded. In practice the gain at 15 kHz could be set to at least 3 dB and still leave sufficient allowance for the effect of pre-emphasis. This may be justified since an overload-protection limiter inserted after pre-emphasis would be caused to operate only rarely by above-average-level, high-frequency programme content.

considered to account for such anomalies; the impairments listeners were asked to assess were in general very small and therefore it was not surprising that the less experienced listeners were sometimes inconsistent in their assessment of programme-modulated quantising noise. Moreover, for comparisons between systems having only slight differences in performance, there was a tendency for presentation A to be preferred in relation to presentation B, even when the latter was the superior condition. It was also thought that the use of a recording of several playings of the same five-note piano scale test item, although a convenient technique for subjective test purposes, may have led to inconsistencies because of minor differences between each of the three playings of the five-note scale in each test item.

Accordingly, a second series of tests was made both to reduce and to check for inconsistencies. Twelve experienced listeners were used and pairs of given conditions were presented several times with the order of conditions reversed for some of the presentations. To avoid doubts about possible differences in recordings, a tape-loop system was used in which the same single recording of one playing of the five-note piano scale was replayed over and over for the tests.

Although the consistency of results was better with the second series of tests it was felt that room for improvement still existed. A third series was therefore made in which listeners were asked to select conditions A or B for themselves until certain of their assessment for each pair of conditions available. This technique yielded the most reliable results but the number of different conditions tested was limited to those of prime interest because of the practical difficulty of arranging for changes which could be controlled by the listeners' A/B switch, and also because of the considerably increased time required for this method of testing.

4.3 Results of subjective tests

Time did not allow the effect of parameters (b) and (c) (see preceding Section 4.2) to be evaluated in the third series of tests. The earlier less reliable tests showed that over the tested ranges of parameters (b) and (c), as given by conditions (e) and (f) in Section 4.2, and for any of the companding laws given by condition (d), no consistent significant differences in programme-modulated noise were discovered.

This result is interesting in the case of parameter (b) since programme-modulated quantising noise must increase as the intervals between transmissions of the scale-factor word are increased. A difference would in particular be expected between the one-word interval conditions (instantaneous companding) and the 16-word or 24-word interval condition (near-instantaneous companding).

The relative unimportance of the gain setting of the CCITT pre-emphasis curve is not unexpected since there would tend to be a compensatory action. As the pre-emphasis gain is increased the compandor is caused to operate more frequently in the coarser quantising ranges. However, the noise-reduction benefit derived from the corresponding de-emphasis following the d.a.c. is also increased, thus maintaining a *status quo.*

Parameter (a), the number of segments in the companding law, was evaluated as accurately as possible using the final test techniques described in Section 4.2. Tests were concentrated on two near-instantaneous laws having 4 and 5 segments respectively and on one instantaneous law having 6 segments. The latter is of special interest since it is a very close approximation* to the well-known A-law mentioned in Section 2.2 which has already been proposed as suitable for high-quality sound signals by other workers.[11,12] For all three laws, the pre-emphasis gain at 15 kHz was set to 3 dB; for the near-instantaneous laws, the interval between scale-factor transmissions corresponded to 16 signal sample words (0·5 ms). The results of the tests are given in Table 1 below, where the figures in the 'mean grade' column relate to the 7-grade comparison scale given in Section 4.2 and express the quality with reference to the linearly-coded 13 bits-per-sample (with dither) system. A negative mean grade corresponds to poorer quality than with the reference system.

These results show broadly that, when compared with the 13-bit linearly-coded reference system, the 6-segment law (A-law approximation), the 5-segment law, and the 4-segment law, are approximately graded 'worse', 'slightly worse', and 'same as' respectively.

*The difference is that the initial two segments of the proposed A-law, with slopes corresponding to 14 bits and 13 bits per sample, were replaced in the BBC experimental compandor by a single segment corresponding to 13 bits per sample with dither.

TABLE 1

Law	Bits per sample required to be transmitted	Mean Grade	Standard Error of the mean
4-segment, near-instantaneous	10·125	– 0·36	0·26
5-segment, near-instantaneous	9·187	– 1·09	0·3
6-segment, instantaneous	10·0	– 1·73	0·32

These conclusions were confirmed by later informal tests in which extremely critical listeners switched between linear coding and the 4-segment near-instantaneous law whilst listening to programme-signals from a piano being played 'live' in a studio.

5. Cascaded codecs using near-instantaneous companding

An attempt was made to simulate the effect of cascading up to four 4-segment near-instantaneous coders and decoders (codecs) by repetitive tape recording through the single experimental codec available. Dolby 'A' compandors were used to minimise the increase in steady noise from the successive tape recordings.

From comparative listening tests the inevitable increase in steady background noise due to cumulative contributions from both the tape recordings and the digital system were clearly audible and tended to mask the increase in programme–modulated quantising noise. Critical listeners were just able to discern an increase in the latter for a simulation of four cascaded codecs.

In general it was felt that the method of simulating cascaded codecs was not satisfactory, particularly since it necessarily included 'double' companding, i.e. a Dolby compandor for keeping the recording noise as low as possible, together with the digital compandor. The increase in programme-modulated quantising noise is, of course, predictable, (as is the steady background noise), i.e. it would be 6 dB for four cascaded codecs and, as such, might be expected to be comparable to that given by the 9-bit, near-instantaneous system assessed as 'slightly worse' than the 13-bit linearly-coded reference system. It is not therefore considered to be a serious shortcoming of the 10-bit near-instantaneous companding system, particularly since it is considered that only rarely would four such codecs be cascaded.

6. Effect of transmission errors in the near-instantaneous companded p.c.m. system

The effects of transmission errors in the near-instantaneous companded p.c.m. system were investigated both theoretically and experimentally and the results of this investigation are described fully in another Research Department Report[13]. Only a brief account of this aspect is therefore given in this Report.

For error protection of the reference 13-bit linearly-coded system, one bit per sample is used to protect the five most significant digits, erroneous words being replaced by the previous correct sample. In general, a companded signal is more immune to the effect of errors because, in effect, the most significant digits of the original linearly-coded signal are transmitted only rarely, i.e. when a high-frequency high-level signal is present. It is considered that the near-instantaneous companded signal requires protection against errors in the two most significant digits only of the sample word; in addition the scale-factor word needs protecting.

Summing up, fewer error-protection digits are required for a digitally companded signal than for a linearly-coded signal. In the near-instantaneous system it is considered that, for a satisfactory performance in the presence of transmission errors, only one protection bit need be used for the two most significant digits of two samples (taken from different channels in a multiplexed system) and one protection bit for the scale-factor word.

7. Application of near-instanteous companding to multiplex operation

The near-instantaneous digital companding technique is particularly suitable for multiple-channel operation in which, by time-division multiplexing techniques, a serial bit stream carries a number of digitally-coded audio signals. In such a system it is considered that the additional equipment complexity required for companding, although in any case not a great increase over that required for the basic a.d.c. and d.a.c. processes, could be reduced by using a single, time-shared compandor system instead of separate companding systems for each channel.

A detailed investigation, also the subject of a separate Research Department Report[14] and a contribution to *Electronics Letters*[15], has shown that it is feasible to fit six high-quality sound channels into a standard 2·048 Mb/s primary digital transmission circuit using the near-instantaneous companding technique. Such a system is of interest to the BBC since it is a potentially economically attractive solution to future requirements for the transmission of three stereophonic programmes.

8. Conclusions

An experimental digital compandor has been constructed and has substantially confirmed the predictions made from earlier work using a simulation technique. The experimental compandor has shown that it is possible to reduce the number of bits per sample to about 10 using a 'near-instantaneous' companding action, the programme quality being virtually indistinguishable by critical listeners from that of a linearly-coded system employing 13 bits per sample with dither. The A-law of instantaneous companding, which also requires 10 bits per sample and has been proposed by other organisations, results in inferior quality and is not considered to meet BBC requirements for high-quality sound transmissions.

Using a near-instantaneous digital compandor with a sampling frequency of 32 kHz, the gross bit rate per channel without error protection would be about 320 kb/s compared to the BBC linearly-coded system now in service, which requires 416 kb/s. Six high-quality channels could thus be transmitted in a standard 2·048 Mb/s bit stream leaving about 115 kb/s for synchronisation and transmission error protection.

9. 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. OSBORNE, D.W. 1972. DIGITAL SOUND SIGNALS: further investigation of instantaneous and other rapid companding systems. BBC Research Department Report No. 1972/31.

3. MANSON, W.I. *and* SHORTER, D.E.L. 1969. A p.c.m. sound-in-syncs distribution system: syllabic companding. BBC Research Department Report No. 1969/48.

4. CHEW, J.R. *and* INGLETON, J.G. 1968. Pulse-code modulation for high-quality sound signal distribution: instantaneous compandors using diodes as non-linear elements. BBC Research Department Report No. 1968/7.

5. CATTERMOLE, K.W. 1969. Principles of pulse code modulation. London, Iliffe, 1969, pp. 129–152.

6. BARTLETT, C.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. British Patent Application No. 10382/72. 6 March 1972.

8. CCIR Rec. 412, Vol. V, Part 1. New Delhi, 1970.

9. CCITT Red Book. 1957, III, P. 273.

10. SABATINO, D. 1972. Results of subjective tests on p.c.m. co-decoded high-quality sound signals. RAI Technical Reprot No. 72/11/E.

11. HESSENMÜLLER, H. 1973. The transmission of broadcasting programs in a digital integrated network. IEEE Trans. Audio Electro-acoustics, 1973, AU-21, 1, pp. 17–20.

12. CCITT, 1972. Doc. AP V, No. 131, pp. 94–106

13. REID, D.F. *and* CROLL, M.G. DIGITAL SOUND SIGNALS: effect of transmission errors in digitally companded systems. BBC Research Department Report in course of preparation.

14. CROLL, M.G., REID, D.F. *and* OSBORNE, D.W. DIGITAL SOUND SIGNALS: multiplexing six high quality sound channels for transmission at a bit rate of 2·048 Mb/s. BBC Research Department Report in course of preparation.

15. CROLL, M.G., MOFFAT, M.E.B. *and* OSBORNE, D.W. 1973. Nearly instantaneous digital compandor for transmitting six sound-programme signals in a 2·048 Mbit/s multiplex. *Electron. Letters.* 1973, **9**, 14 pp. 298–300.

10. Appendices

10.1 Calculation of quantising noise in companded p.c.m. systems

The signal-to-quantising noise ratio of a linearly-coded p.c.m. signal is discussed first before proceeding to the quantising noise characteristics of companded p.c.m. systems.

10.1.1 Linear coding

When a signal is quantised in a linearly-coded p.c.m. system it is rounded to the nearest quantising level. This results in quantising errors of up to ± one-half a quantum step which are randomly distributed for relatively high-level signals. At low signal levels the errors may not be random and the resulting subjective effect is described in Reference 1.

The simplest method of expressing the signal-to-quantising noise ratio of a linearly-coded p.c.m. system is in terms of the peak-to-peak signal to peak-to-peak noise. This is because the peak-to-peak signal is the sum of all the quantising levels available while the peak-to-peak noise is equal to the height of one such step. Hence for a binary coding system using n bits, giving 2^n quantising levels, all assumed to be equally spaced,

$$20 \log_{10} \frac{\text{peak-to-peak signal}}{\text{peak-to-peak quantising noise}} =$$

$$20 \log_{10} 2^n = 6n \text{ dB.}$$

However, in communications engineering it is more usual to express the signal-to-noise ratio in terms of the r.m.s. noise. The magnitudes of the random errors which form the quantising noise have a rectangular probability density function and, from relatively simple statistical methods, it can be shown that the ratio of the peak-to-peak to r.m.s. value is $2\sqrt{3}$ which is equivalent to 10·8 dB. It is also convenient to consider sinusoidal signals in terms of peak value (one half the peak-to-peak value). Hence,

$$20 \log_{10} \frac{\text{peak signal}}{\text{r.m.s. quantising noise}} = (6n + 4{\cdot}8) \text{ dB}$$

If dither is employed to mask granular distortion at low signal levels the signal-to-noise ratio is reduced by about 1·5 dB.[1]

10.1.2 Four-segment near-instantaneous companded system

For the reason given in Section 2.2.3 of this report the quantising noise characteristic for the near-instantaneous companded system depends upon the frequency of the input sinusoidal signal. For frequencies greater than about 1 kHz the peak signal to r.m.s. quantising noise characteristic is given by,

$$(6m + 4{\cdot}8 + S)\ \text{dB}$$

where S is the level of the input signal in dB relative to the maximum which can be transmitted and m is the number of bits of the equivalent linear system corresponding to the slope of each segment. For S in the range −18 dB and below, m = 13 and the dither correction factor of −1·5 dB is applied.

For S in the ranges −12 to −18 dB, m = 12
− 6 to −12 dB, m = 11
0 to − 6 dB, m = 10,

and in these ranges the effect of the 13-bit dither signal is appropriately reduced. Hence the full-line, saw-tooth characteristic, peaking to 65 dB in Fig. 4 may be derived.

For frequencies less than about 1 kHz the action of the near-instantaneous compandor becomes effectively instantaneous and the method outlined in Section 10.1.3 must be used to derive the dashed curve shown in Fig. 4.

10.1.3 Instantaneous companded system

With instantaneous companding the quantising characteristic is linear for signals which excurse only over the first segment of the companding law as shown in Fig. 4. For the four-segment law this part of the characteristic is already covered in Section 10.1.2. For the A-law the linear part of the characteristic is given by,

$$(6n + 4{\cdot}8 + S)\ \text{dB}$$

where n = 14 = the number of bits used for coding the first segment and, as before, S is the level of the input in dB relative to the maximum which can be transmitted.

When the input signal amplitude is such that more than the first segment of the companding law is required the quantising noise is modulated according to the waveform of the input signal. It can be shown that the r.m.s. value of the resultant modulated noise is given by,

r.m.s. value of the quantising noise for the first segment of the companding law x r.m.s. value of

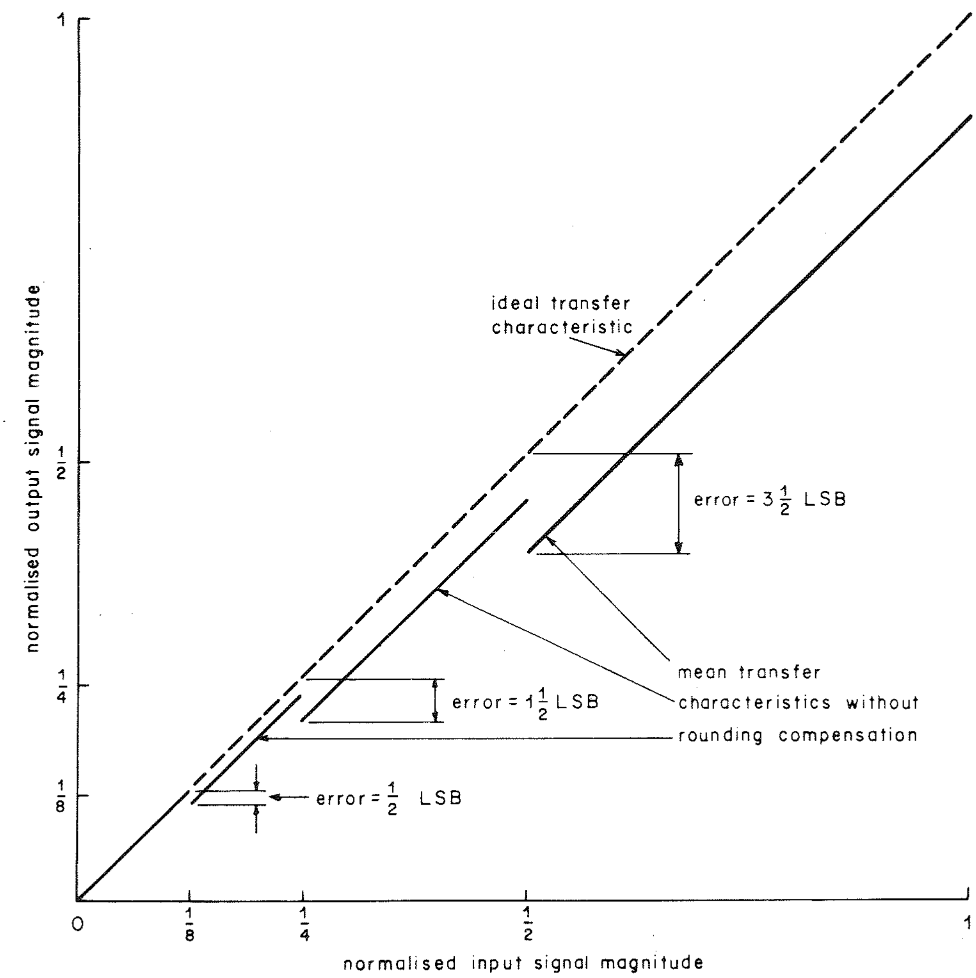

Fig. 6 - Rounding errors in a four-segment, instantaneous companding law: positive quadrant only shown

the modulating signal.

For a quarter-cycle* of a sinusoidal analogue input signal the modulating signal is a staircase waveform in which the number of steps is related to the number of segments of the companding law over which the input signal excurses and the heights of the successive risers of the staircase are related by a factor of two.

In order to derive the instantaneous companding characteristics shown in Fig. 4 calculations of the r.m.s. value of the modulating signal were made for 0·5 dB intervals of the input sine-wave signal amplitudes. In the case of the A-law, such calculations were made for input signals in the range −36 dB to 0 dB relative to the maximum of the coding range. For the four-segment law similar calculations were made over the range −18 dB to 0 dB.

10.2 'Rounding' error compensation in digital companding

Digital companding is done by removing the most significant bits (MSB's) or the least significant bits (LSB's) of the original linearly-coded signal, depending upon the

*Because of the symmetry of the input signal it is sufficient to calculate the r.m.s. value of the modulating waveform for only one-quarter of the input sine wave (i.e. from 0° to 90°).

instantaneous signal magnitude in the case of instantaneous companding, or upon the peak magnitude in a group of signal samples for near-instantaneous companding. When LSB's are removed, i.e. the quantised signal is rounded, the overall mean transfer characteristic of the compandor has discontinuities unless a compensating action is introduced in the decoder. Fig. 6 shows the effect for a 4-segment instantaneous companding law and Fig. 7 for near-instantaneous operation of the same law; for the purposes of illustration the quantising accuracy has been grossly reduced.

Fig. 6 shows that a discontinuity occurs each time the input signal waveform causes the compandor to change from one range to another. Fig. 7 shows that a discontinuity occurs at each zero crossing unless the signal magnitude is such that the compandor is in its most accurate quantising range. Compensation for these discontinuities is included in the digital expandor by adding the appropriate quantity to the digital signal according to which range of the companding law is in use. However, as explained in Section 3 of this report, no benefit was gained in practice by including compensation to ½ LSB accuracy (which requires 14-bit decoding accuracy), since this was masked by the 'dither' introduced in the original 13-bit coding of the signal*

* Investigation of the rounding error was carried out by D.F. Reid

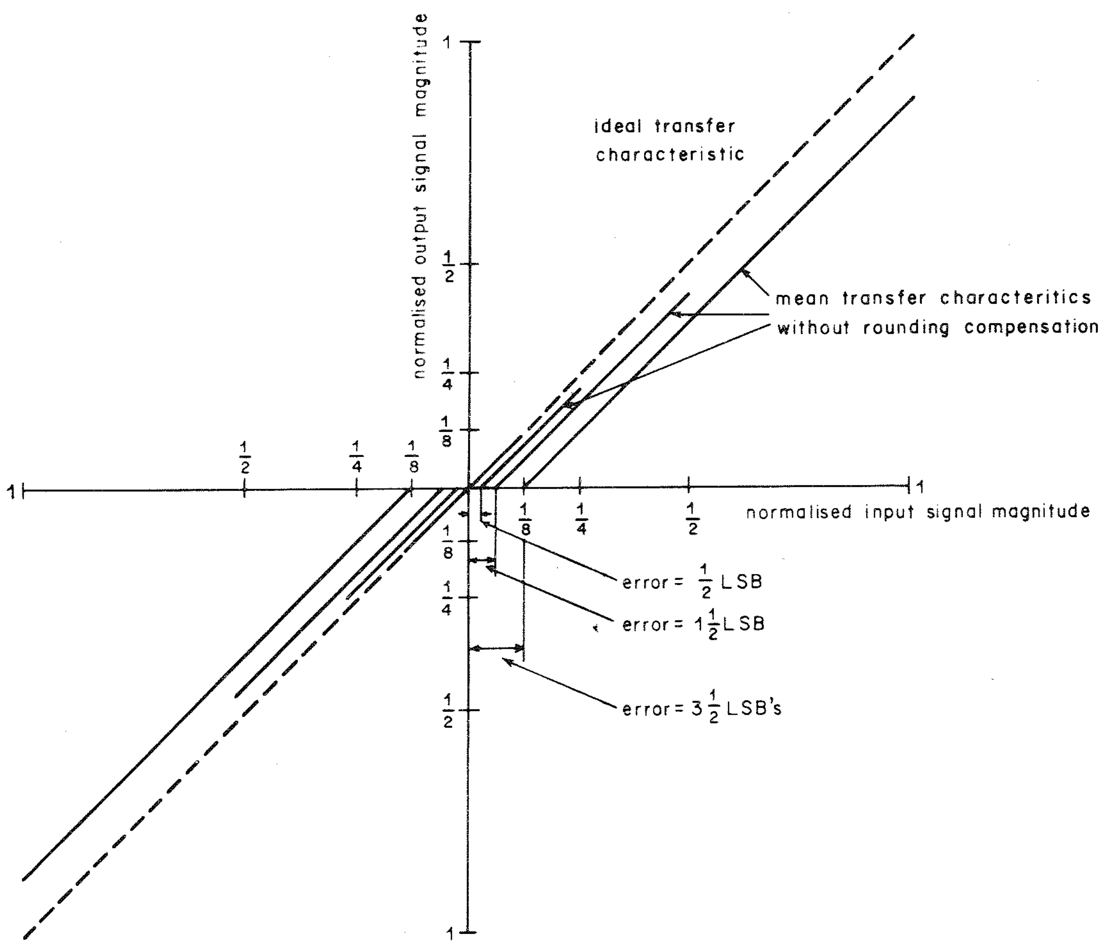

Fig. 7 - Rounding errors in a four-segment, near-instantaneous companding law

CLA/AM

(EL-83)

Printed by BBC Research Department, Kingswood Warren, Tadworth, Surrey