BURIED DATA IN NICAM TRANSMISSIONS

J R Emmett

Thames Television PLC, Broom Road, Teddington, Middlesex, UK.

1. INTRODUCTION

The Near Instantaneous Companding And Multiplex digital audio coding system was developed by the BBC in the late 1970's in order to improve the quality of their microwave feeder links to the then new stereo radio FM transmitters (1).

The actual bit rate reduction was performed by means of a block floating point conversion from 14 bit linear coding, down to 10 linear bits (Fig. 1) plus two range bits per block. The sample rate employed was 32kHz and so the block length was basically 320 samples plus the two range bits, which were well protected from errors. NICAM first came to television in the form of sound-in-sync. Here, a version known as NICAM III was used, which had an overall data rate of about 672 kbits per second. In the mid 1980's as a side project to the E-PAL vision system, a transmission digital sound carrier was developed which used a version of NICAM III coding, now known as NICAM 728 from the 728 kbits per second transmitted data rate (2). It was soon found that this sound carrier could be shoehorned into the conventional UK system I vision transmissions with a negligible adverse affect on existing receivers. Later, it was also found to fit experimentally into most other world terrestrial transmission standards, and is now being installed as the preferred dual channel Television Sound System in many countries.

There are two technical reasons why the NICAM carrier works so well. Firstly, only a low level of Rf is needed for good digital reception, since carrier to noise ratios can be tolerated down to around 10dB without errors becoming audibly noticeable. Secondly, the Rf carrier has a noise like amplitude character, and this ensures that any cross-modulation into the analogue vision or FM audio causes only the very minimum of impairment.

2. NICAM 728 DYNAMIC PROGRAMME RANGE

NICAM uses quite a severe amount of programme pre-emphasis. The characteristic is shown diagrammatically in Fig. 2, and the international reference CCITT J-17 (1972) shows it to stem originally from the telecommunications world. The purpose of pre-emphasis in this application is not so much to improve the high frequency signal to noise ratio, as to reduce the level of low frequency signals. Lower peak levels reduce block range changing during bass heavy passages of sound. Range changing manifests itself as low level programme modulated noise, which for plain bass programme would not in all probability be "masked" by higher frequency programme content.

Although this pre-emphasis undoubtedly improves both the measured and the perceived noise performance of NICAM, it does present a few engineering problems concerned with programme clipping levels. As a result there is a large area of the possible modulation "envelope" normally left unused. As an example of this the overall dynamic range of the system used by the BBC and ITV in the UK is illustrated in Fig. 3, with the FM sound clipping level shown as a reference.

3. THE CONCEPT OF MASKING

The first investigation of the audibility of one tone being masked by the presence of another tone, was probably made in 1924 by Wegel and Lane. Their use of pure tones caused complications in terms of beats and cross-modulation, both as a result of the non-linear nature of the inner ear, as well no doubt, as non-linearities in the electronic equipment of that period.

By the 1940's there was an intense interest around the Radio communications field in the detection of Audio signals in the presence of noise. During this period many important investigations were made concerning masking in the presence of variable bandwidth noise. From this work came the concept of critical bands (3), a rough definition in this context being the maximum bandwidth of noise which the ear treats as a single tone. Later, the masking effect was found to extend in time before and after a masker signal had been applied. Today, the work on masking has led to a number of perceptually based bit rate reduction system for digital audio (4), the best of which can subjectively attain a coding quality similar to that of NICAM with less than one quarter of the bit rate. The question might therefore be asked if programme material coded into NICAM form could effectively mask other signals carried at a low level within the digital system.

4. THE MASKING TEMPLATE FOR PROGRAMME MATERIAL

Although NICAM was very much the product of engineering development rather than auditory theory, there are two definite features which leaned towards the use of masking. Firstly, the choice of block size, 1 millisecond, fits in well with the ears perception of transients, and secondly the pre-emphasis applied gives a crude frequency dependent masking template of the background quantising noise.

If we were to filter the programme material into critical bandwidths (Fig. 4), masking experiments tell us that we could then probably accept background quantising noise *within that critical band* up to a level as high as 4dB below the signal. Even allowing a larger margin for indirect perception of the quantising noise, the allowable quantising noise still represents a huge digital data capacity.

In practice, the critical bands are continuous in nature and not conveniently arranged like a non-linear third-octave equaliser, so that the first function of any modern perceptual bit rate reduction system is to form a continuous frequency template of the signal which would be masked by programme material. The usual way this is performed is by Fast Fourier Transform and look up table, so that there is a degree of subjective human skill needed to choose a family of tables to suit all types of programme material. Fig. 5 illustrates this.

It is becoming apparent that for the optimum utilisation of this technique, it may be better to employ different tables for different programme types, the table employed being selected by a small fraction of the auxiliary data carried with each NICAM transmission. The accurate representation of stereo images is one example where there may be incomplete understanding of masking at present.

5. THE TREATMENT OF THE DATA

Once a masking template is selected according to the source programme material, the data to be carried needs to be processed in order to lie within the frequency/dynamic range space represented by that template. One way of doing this is to employ adaptive scramblers of the general form:-

$$x^{(m)} + x^{(m + n)} + x^{(m + n + o)} + \text{etc} + 1$$

(where m, n, o, etc are integer delays)

This algorithm can be represented by the circuit in fig. 6, and the total length will depend on the lowest frequency in use. One scrambler will be needed for each bit employed for data, and those employed for the bits of lesser audio significance will carry a higher data rate for a more continuous period, whilst those associated with higher audio bits will come into use less often. Another approach would be to use a code look up table associated with each masking template selected.

Whatever method is used to Channel code the data, the net effect within each 1 millisecond block is that the data signal will have a similar pattern and "sound" to that of the signal, and because of the lower levels occupied the signal will effectively mask the data as far as the final listener is concerned.

At the receiver, the audio can be decoded in the usual way, the performance of all recent D to A converters and television audio chains being such that the masked data will not become acoustically apparent. As far as the data is concerned, the template match can be performed in the same way as the transmitter, the nature of a digital signal ensuring the choice will be the same at both ends of the chain. The only difference at the receiver being a need to ensure that the minimum of error extension occurs to the data system. Intelligent choice of the scrambling algorithms or look up table hamming distance will ensure this.

6. POSSIBLE DATA RATES

During silence, no data can theoretically be carried if we wish to maintain the high decoded signal to noise ratio achievable with NICAM. In actual fact a data rate of up to 50 kbits could well be undetectable during many types of programme silence, just because of the masking effect of the inherent background noise levels.

During actual programme sound, an analysis of the possible data rate looks at first rather daunting, not only because the actual programme modulation level is so variable in the statistical sense, but because the available data rate is so dependent upon the programme frequency spectrum.

Fig. 3 can be taken as a starting point for the analysis of programme statistics and Fig. 7 shows a typical long term television peak and mean programme distribution with frequency for television sound.

As a point of incidental interest the experimental verification of this distribution made in 1991, was not significantly different to a similar measurement made for Radio in 1934(5). The expected higher levels of high audio frequencies today may be balanced out by better reproducing chains at high frequencies. If this mean programme level is mapped into the 10 bit linear code book for NICAM, allowing for range changes, a surprising peculiarity of NICAM becomes apparent. Since the data carrying capacity of any masking template is directly proportional to frequency, the activity diagram of Fig. 8 shows the result mapped into a linear frequency scale. With the surprisingly high activity spread well up the spectrum due to the pre-emphasis characteristics, a conservative rectangular mean masking template may be considered. The template shown on Fig. 8 represents no less than a mean raw data rate of 280 kbits for a stereo channel.

Another way of looking at this result is as follows:- As the peak signal level rises into the next NICAM range, the data rate will drop by some 20% initially, then rise again until the next range boundary is reached. This performance is repeated for each range in turn, leaving the curious results that *the available data rate will be relatively constant for peak programme levels down to 30dB below the clipping point*. Below that level the data rate will fall linearly with programme level.

7. CONCLUSIONS

Although this technique has been demonstrated in its most basic form, and the statistics of programme sound have not been fully studied, it still seems that at least 100 kbits per second of buried data could be sent using an unchanged NICAM transmission system, and without any audible effect on the programme sound. The obvious problem is utilisation of a data rate which is programme dependent, although as mentioned in section 6 the characteristics of NICAM keep the data rate more or less constant for a programme level change of 30dB.

One obvious application for a programme related data rate is data related to the programme! Here the candidates include the Audio Descriptive Service (ADS) for sight impaired viewers, or a programme sound channel suitable for the hearing impaired, mixed without high level effects, and with a reduced dynamic range. Surround sound channels and sub-woofer information could also be easily carried providing a useful and compatible future link between terrestrial television standards and those for HDTV and feature films.

7. REFERENCES

(1) C R CAINE, A R ENGLISH, J ROBINSON, "NICAM-3, A Companded PCM System for the Transmission of High Quality Sound Programme" - Proc. IBC 1980 (IEE).

(2) NICAM 728: "Specification for Two Additional Digital Sound Channels for Terrestrial Television". Joint IBA/BBC/BREMA document ISBN 0-563-20716-7. First published August 1988, also in the form of EBU and CCIR Recommendations.

(3) F ZWICKER, "Subdivision of the Audible Frequency Range into Critical Bands". J Acoust Soc of Am. 33 pp 248 (1961).

(4) G THEILE, G STOLL, M LINK, "Low Bit-Rate Coding of High Quality Audio Signals", EBU Tech Review 230 (August 1988).

(5) F E TERMAN, "Radio Engineering", McGraw Hill.

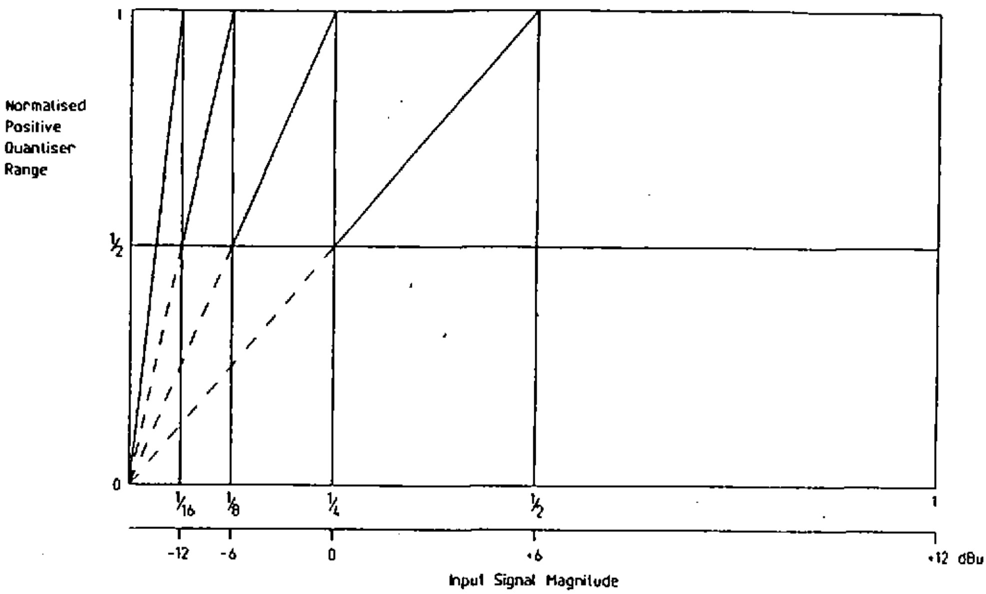

FIGURE 1.
Five-Segment Near Instantaneous Companding Characteristic. eg. BBC NICAM - 3

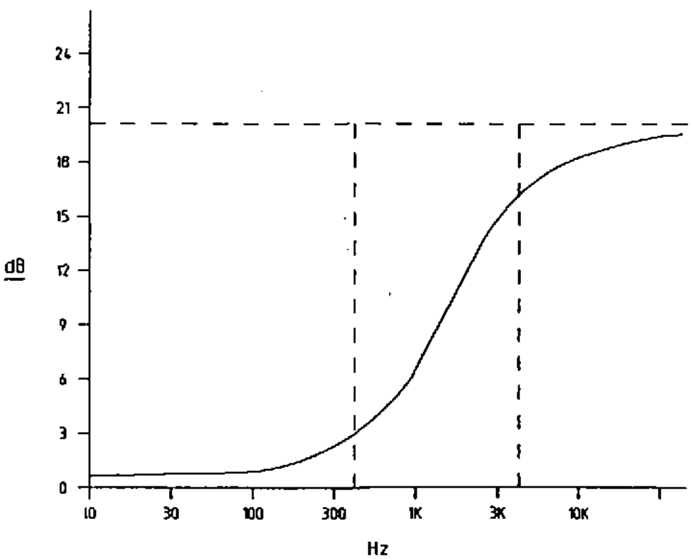

FIGURE 2.
J-17 Pre-emphasis Characteristic

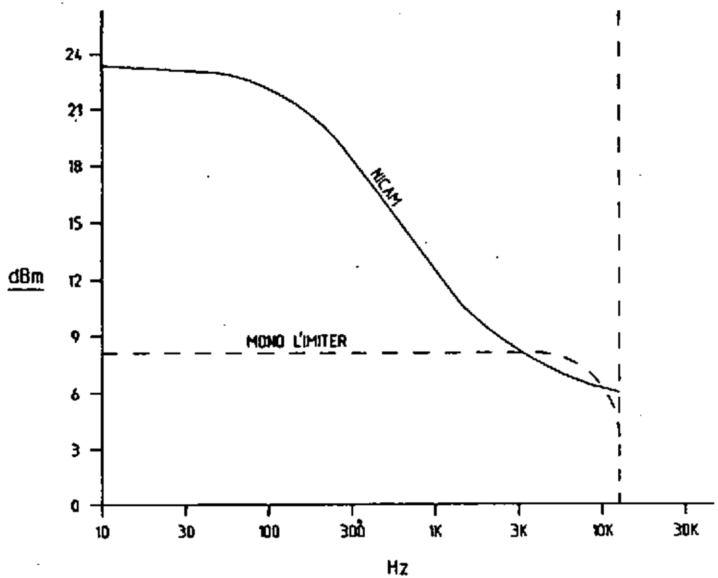

FIGURE 3.
U.K. Television Sound Clipping Levels

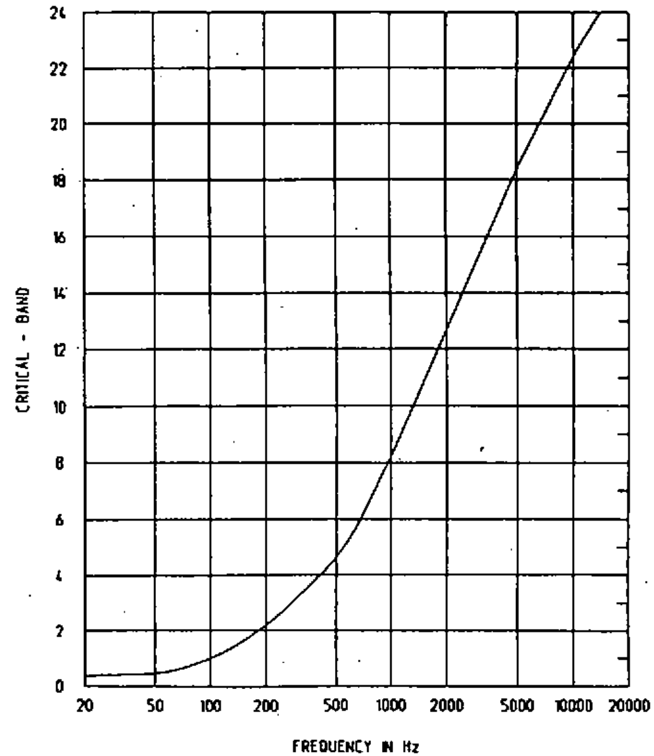

FIGURE 4.
Critical bands (after zwicker)

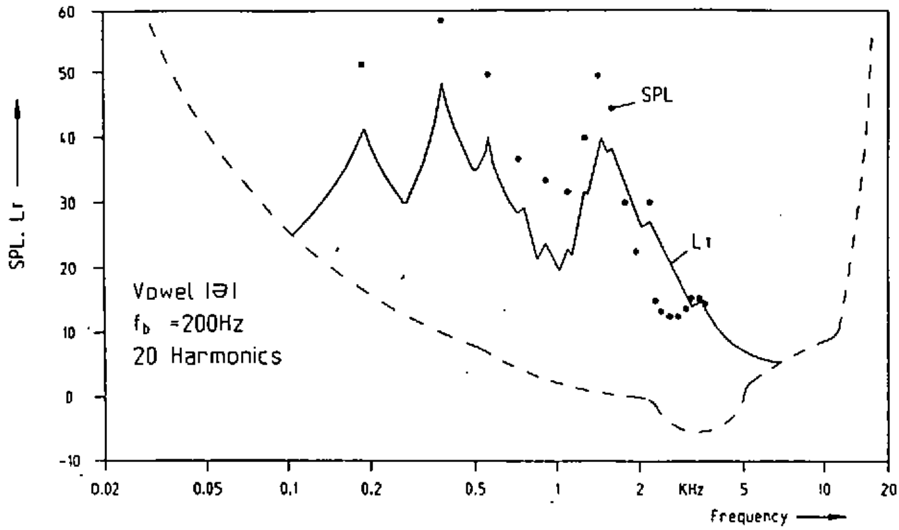

FIGURE 5.
Masking pattern Lт (full line) of the vowel |ə|

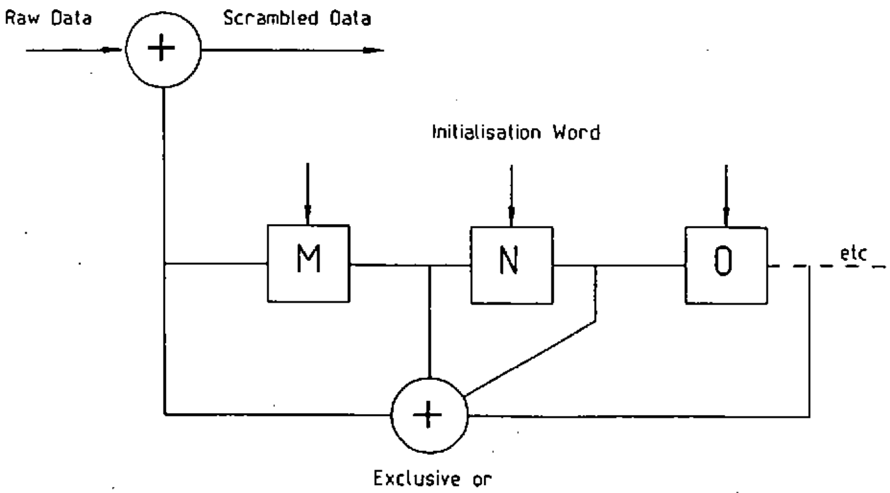

FIGURE 6.
General form of energy dispersal data scrambler

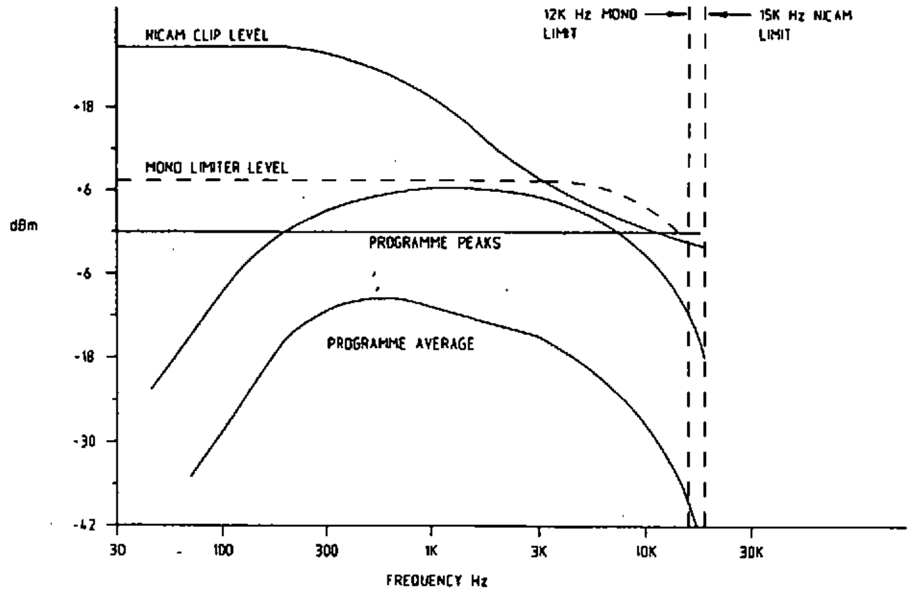

FIGURE 7.
Long term Programme statistics

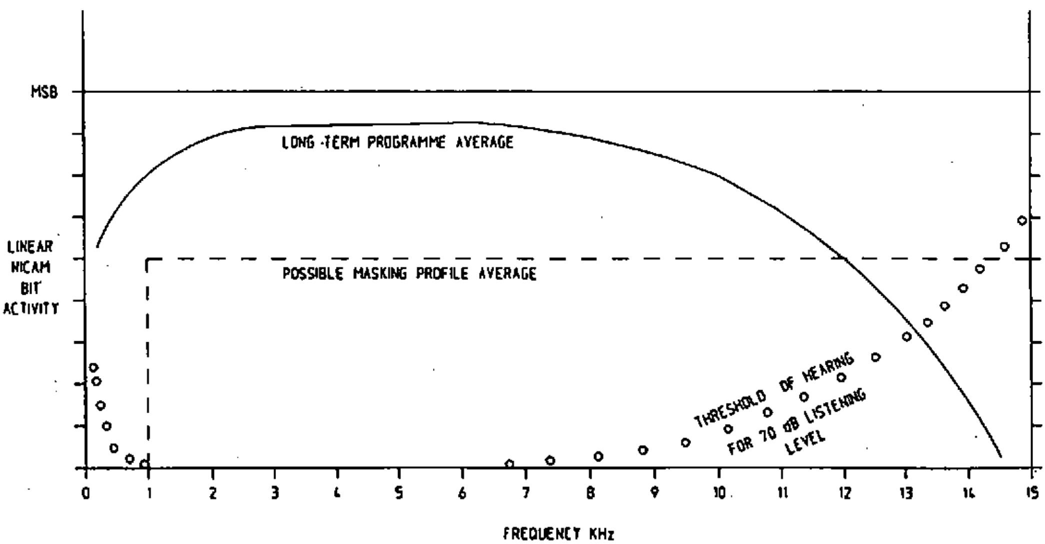

FIGURE 8.
Programme average remapped into Nicam bit activity