Proceedings of the Institute of Acoustics

CONTROLLING THE LOUDNESS OF TELEVISION SOUND

From time to time members of the public on both sides of the Atlantic have complained about variable sound levels within and between television programmes, in particular, that background music and effects are sometimes too loud for speech and that some commercials are too loud.

It is well known that the impression of loudness is not due exclusively to peak sound pressure or intensity; it is determined by the energy in the sound wave which in turn is a function both of the amplitude of the wave and the shape of the envelope.

So where lies the problem, if any? It is true to say that there are some considerable difficulties to be overcome, and these difficulties arise from two fundamental causes; firstly the ear which after all is the final subjective measuring device is only capable of assessing absolute intensity in very general terms and the auditory system judges loudness largely by comparison. Secondly, with the advent of Digital Nicam there is a transparent path from the sound controller in a studio to the viewer that provides a very large dynamic range in the order of 70 - 80 dbs.

The conflicting aspects of the problem now become apparent; on the one hand it is necessary to compress the intensity range in such a way that the apparent values of loudness of the various forms of programme are acceptable to the vast majority of viewers without distracting from the programme producers artistic intentions, or the selling power of a commercial; and on the other hand it is a case of making the fullest use of the transmissions system.

BACKGROUND

In the 1938 C.G. Mayo at the BBC research department gave broadcasters the Peak Programme Meter (PPM), and to his credit must go the fact that its purpose and function have changed little apart from the tightening of specifications and its performance being defined for international standardisation.

The function of the PPM is to provide visible indications of

CONTROLLING THE LOUDNESS OF TELEVISION SOUND.

J White

Granada Television

variations in programme volume, and these variations must be confined to maximum and minimum values conditioned by the limits of intensity which the transmission system can handle.

Several fundamental principles were incorporated into the first PPM including the determination of the attack and decay time constants and the realisation that full wave rectification was necessary to display asymmetrical peaks which could differ by 8db or so in some cases. Furthermore, extended tests indicated that the ear required a period of time to recognise distortion; the shorter the peak the greater the percentage of overmodulation that can be tolerated. It was ascertained that a circuit registering 80 per cent of the peak value of a square wave in 4 milli - seconds provided the best compromise, and today there is an air of irrelevant mystery as to how this figure (now termed integration time) has crept up to 10 ms in the IEC standard, which incidentally runs to thirty pages.

Broadcasters are concerned for the majority of viewers who are usually described as having a pair of ears that are 'Normal' for their age, and both the BBC and IBA have drawn guidelines with the intention of defining as far as practicable the correct programme meter indications and limits for various types of television programmes, which are considered to be guidelines for the subjective factor of loudness.

GUIDELINES FOR LOUDNESS CONTROL

The final responsibility of choosing the correct sound level and loudness must be with the origination studio sound control engineer, and due to the subjective nature of the medium it is obviously not desirable or practicable to lay down rigid rules. However it must be said that too much latitude exercised in the past may account for some of the viewers complaints about variable sound levels.

Over the years a preferred list of relative levels for differing types of programme material has become commonplace to assist in the balancing of sound. Bearing in mind that the PPM has a well defined dynamic range of 30dbs, which is about the range of the human voice; the scale is divided into six equal 4db divisions by numbered lines 1 - 7, and the mid-scale '4' position is vertical; the following tables illustrate the principle.

(a) Drama

The level of significant dialogue should not fall below PPM 3; with good diction on close microphones, lower levels can be accepted with continued intelligibility.

The normal PPM limits within which the peaks should register most of the time (having regard to nature of the material and the normal aesthetic latitude required) for Drama are as follows:-

Average Dialogue	PPM 3 - 5	
Loudest Speech	PPM 5 - 6	
Quiet Speech	PPM 2 - 3	Full Range 2 - 6 (16dbs)

The level of Incidental Music and effects should be judged aurally.

(b) Variety

Speech	PPM 4 - 5 1/2
Musical Interlude	PPM 3 - 5
Laughter	PPM 3 - 5
Applause	PPM 2 - 4
	Full Range 2 - 6 (16db)

(c) Compression

Any material that is normally made with varying degrees of compression of the dynamic range should not be allowed to peak more than PPM4.

(d) News and station announcements (good diction on close microphones) PPM 4 - 4 1/2, and so on.

So much for the PPM. In actual fact the VU Meter is by far the most common form of programme meter with virtually 100% acceptance in the U.S., Australia, Japan, and practically every recording studio in every country; and they too have a loudness problem.

CONTROLLING THE LOUDNESS OF TELEVISION SOUND

RESEARCH 1959 - 1983

In the late 50's CBS Television Network carried out a study of audio volume levels for the express purpose of fully exploring the reasons why viewers were complaining that portions of television programmes are sometimes unpleasantly loud. (1)
CBS developed the idea of operating the VU meter in conjunction with the gain reduction meter of an automatic gain control amplifier in such a way that normal programme material was compressed up to a maximum of 6db. With an attack time of 25ms; a recovery time of 0.5 seconds and a new set of operational guidelines, this device was installed throughout the CBS network.

In 1965 CBS laboratories carried out a fundamental study of loudness from which several important results emerged:-
- a new set of equal loudness contours;
- a new summation function;
- a new duration versus loudness level function for octave bands
ranging in frequency from 125Hz to 8 kHz. (2)

In other words there are a great many competing methods for calculating loudness, and no one method seems to be entirely successful in dealing with the great variety of sounds which are encountered in everyday life.

Instead of converting sound pressures to phons and then to sones prior to summation C.B.S perform the summation on the sound pressure side of the equation.

Fundamental to the study of loudness is the measurement of levels of sounds of equal loudness as a funtion of frequency and intensity, and the result of such measurements is portrayed as a set of Equal Loudness Contours. (ELC's). Unfortunately there is a wide divergence of opinion with regard to ELC's even though there is an ISO standard (R226). It has been apparent to many workers in the field that the use of pure tones in an anechoic chamber and the use of earphones are not representative of the typical conditions of listening. The combination of white noise/octave bands/and listening in a diffused sound field produces results that are significantly different from pure tone contours, mainly by exhibiting less rise at low frequency and an absense of high frequency rise; such a specification describes the work of Stevens which resulted in a set of Equal Loudness Index Contours, which serve the role of ELC's in his method of loudness calculation. (ISO/R532-1966)

This 'hunk of granite' has been chipped by CBS, who in addition to realising the importance of determining a set of ELC's under conditions more applicable to the loudspeaker in the living room, realised that the quality of the sound system should be sufficiently high to ensure a uniform response throughout the frequency range of interest, and that the sounds presented should be in bands that correspond to the bandwidth of the filter sections to be finally employed for loudness analysis. Whilst one octave bands of noise is the obvious choice, unlike Stevens and others, CBS decided to use "pink" and not "white" noise because of the favourable spectral distribution of the former and as to be expected, there is a radical difference. The so called CBS practical contours unlike those of Stevens, indicate that there is a peak of maximum sensitivity at 3000Hz for males and 4500Hz for females which perhaps is why women are more bothered by the 'Hi-Fi'.

It was not until 1978 and after much encouragement from the Federal Communications Commission (FCC) that CBS instituted a new research and development programme which significantly advanced the state-of-the-art in loudness measurements and control.

As a result of this effort the FCC were able to evaluate the CBS loudness indicator and controller in 1983, and today it is the most sophisticated known technique for measuring and controlling the loudness of broadcast audio. Unfortunately even if it could be shown that the controller was 100% effective there are many reasons why some complaints would still be received; individual hearing response; viewers tastes and many psychological variables alter one's perception of loudness.
However there is one solid foundation to build on for example, Stevens' suggestion that loudness is a power

function of physical intensity is generally accepted, and a simple approximation to this is that two-fold change in loudness is produced by a 10db step in level; in point of fact CBS are using this traditional pyschoacoustic format for the secondary calibration of their Loudness Indicator, namely, where "one-half loudness" or "twice loudness" is equivalent to a 10db step in level.

Broadcasters too have built a solid foundation that takes the form of a videotape leader immediately prior to the start of all programmes and commercials that comprises visual identification, a countdown clock and test signals. The sound controller records a 1 kHz signal for 30 seconds (PPM 4) on this leader and this is used for setting the listening level; bearing in mind that it will be almost twice as loud with programme. (+8db)

LISTENING LEVELS

It is essential that the importance of loudspeaker listening level is appreciated because it is the monitoring level which affects the balance of sound and mono compatibility to the greatest extent, and it is noted that the Cinema Industry has a firm recommendation for a listening level of 85 db C per channel for 50% modulation.

However it is somewhat different in broadcasting; it is generally accepted that normal conversation level provides a firm basis for reference and sound controllers will adjust the listening level so that speech signals are comfortable in both the control room and the living room; it is a fact of life that sound controllers do follow their programmes into the living room and hope to do better next time.

The question now is how does a sound controller determine subjective loudness.

DETERMINATION OF THE SUBJECTIVE LOUDNESS OF A BROADCASTING PROGRAMME

The problem which arises in studying the loudness of sounds, particularly when these are complex or of a transient nature is of course that loudness is a subjective quantity, and as such cannot be measured directly. This problem has been tackled in a number of different ways; sometimes subjects are asked to match the loudness of a sound to that of some standard comparison stimulus (often 1000Hz tone); in other experiments subjects are asked to rate loudness on a numerical scale, a technique known as magnitude estimation; and suffice to say that there are problems associated with each of these methods.

In 1969 the Consultative Committee of International Radio (CCIR) considered that there was a need for a standard loudness reference recording with which programmes may be compared, and the BBC produced a provisional loudness tape with a comprehensive range of programme material having different subjective loudness characteristics. This was published by the CCIR Secretariat for study purposes leading to the possible production of a CCIR standard reference loudness tape. (3)

Two validation studies were conducted by CBS to determine the accuracy of their Loudness Indicator. The first employed 12 segments each of 1.5 to 2 second duration of programme and commercial material taken off the air (without further adjustments). These segments had been highly processed by compression and other means to achieve maximum loudness within prescribed modulation limits. The second study employed the use of eight 1.5 to 2 second segments of unprocessed material taken from the CCIR Tape Recording. (3)

Suffice to say that CBS were completely satisfied with the performance of the Loudness Indicator and this information was reported to the CCIR; this report was considered together with a report that described studies on level regulation of sound broadcasting signals. Various level regulation procedures were evaluated subjectively under studio and home reception conditions.

CONTROLLING THE LOUDNESS OF TELEVISION SOUND

The conclusion drawn by the CCIR was that it was impossible to satisfy all listeners with 'a single level' regulation of the programme, and that a great number of listeners can already be satisfied by a correct quasi-peak level regulation with a differentiation between music and speech sequences as a function of the type of programme. It was indicated earlier that the L.S listening level should be adjusted when listening to normal conversational speech signals and that the loudness factor could be controlled with a set of guidelines and a PPM.

It is not necessary to know how an engine works to drive a car, but a minority of people find it interesting to know precisely what makes one tick better than another; which raises the question what exactly is being measured by a PPM and how does the 'A' Weighting network fit into the scheme of things in view of its close associations with equal loudness contours and in particular, the fact that CBS developed a new set of contours.

THE PERFORMANCE OF THE PPM

One thing is for certain the PPM cannot indicate the loudness of a signal because loudness is not necessarily related to peaks in programme material; nor is it a function of the average or RMS signal level although it is probably closely related to the latter. The PPM in fact indicates quasi-peaks for signal envelopes that are disguised by a very slow decay time of the electronics, which holds on to the peak indication long after the signal may have returned to a lower average level; the PPM pointer does not start to come down off the scale until it's sure that no more speech information is forthcoming - and then it takes the pointer 3 seconds to move back 24 dbs.

Speech waveforms are characterised by abruptness of amplitude changeover accompanying a consonant and rapid changes in spectrum over a time period of 10-30ms following the release, and in fact listeners learn to identify sequences of sounds when the individual items are as short as 10ms which corresponds to a rate of 100 phonemes per second. Listeners do not perceive each successive item separately but rather learn the overall sound pattern and it is likely that for continuous speech something similar occurs. (5)

Consider plotting a speech waveform in terms of pressure variations as a function of time (ETC); such an energy time curve display indicates that the peaks are fairly regularly spaced which means that speech sounds are periodic and are 'in sync' with 'inner vibrations within the Larynx'.

The PPM tracks syllables with ease and registers transients/phonemes as well, for example, if one inspects a wideband spectrogram of the utterance "I can see you" (5) which incidentally has a duration of one second, one notes the concentration of energy at the frequencies corresponding to formants and in particular the energy of the sibilant sound 's', which lasts for a duration of 150ms at a frequency of 7 kHz. A sound controller must prevent this transient from being clipped too much.

The time interval between the application of the reference input voltage to a PPM and the moment when the pointer passes a point 1db below the reference indication is 100ms so therefore the PPM will track the phrase 'I can see you' and register the maximum peak of energy in the 's'.

The performance requirements for the PPM are very stringent, for example PPM indications are specified for burst durations of 0.5ms, 1.5ms, 5ms, 10ms and 100ms, and also the dynamic linearity within the 30db dynamic range is specified, which is just as well because the threshold for detecting a temporal gap in broadband noises and between clicks is 2-3 ms.

There is an idea that there is a process at levels of the auditory system higher than the auditory nerve which is 'slugged' in some way, thereby limiting temporal resolution. (5)

In 1938 it was found extremely difficult for two operators to compare the fluctuating readings of their instruments for the purpose of checking the programme volume between any two points on the BBC network. A so called 'slugged response' was devised that utilised a resistance of 10K in series with the meter itself and a capacitor of 100 micro-farad was connected across the combination, and this gave a satisfactory indication of the average programme volume which could be compared accurately with an adjacent station on the network. This requirement resides in the IEC history book.

The brain may slow the performance of the ear down and the PPM can be 'slugged'; but the precision sound level meter with its 'A' Weighting Network has been slugged with 500ms since day one, which usually takes precedence over the fast 200ms response.

CONSIDERATIONS OF dBA ('A' Weighting network)

Considering that the precision sound pressure level meter is married to the Phon, and the Phon has a close relationship with the Sone, it is of interest to note that there is only 6dBA difference between a voice speaking intimately and the same voice forcefully. Carrying this a stage further, dBA measurements were obtained for all of the 29 sources that are peaking to 100% modulation on the CCIR Loudness Study Tape, and it is noted that the average range from normal voice to Opera is not more than 7dBA. Furthermore, for many decades it has been established that there must be 6-8 dbs difference in level between commercials and unprocessed programme material to produce a smooth loudness effect for the viewer.

Similarly, it is evident that while the Energy/Frequency/Curves of speech signals emanating from BBC Radio 2 embrace varying degrees of bandwidth and hence produce varying degrees of loudness, once more the variation is no more than 6-8 dBA.

CBS use 8 filters in their loudness indicator and the loudness algorithm produces a summation response that could be said to be a modified 'A' Weighting Network which in actual fact is not far removed from the redundant 'B' Weighting Network that is based on the 70 Phon equal loudness contour, and this conveniently fits in with the average listening level in the home.

There is no doubt that the spectral content of Programme signals being transmitted today has changed considerably over the last two decades - more low and high frequencies. Bearing in mind that the CBS contours were developed in 1966, it would appear that the 'weighting curve' will be on the move again in the future.

After due consideration of the foregoing the following conclusion is drawn.

CONCLUSION

The PPM and appropriate 'guidelines' have obviously kept the loudness factor under satisfactory control for over fifty years on every conceivable type of sound signal; in partnership of course with a pair of so called 'golden ears' and manual dexterity. Sound controllers do not necessarily have to know what technology is built into the PPM and VU Meter anymore than they have to know the difference between a petrol and diesel engine to drive a car; similarly for the existence of the Phon/Sone and the idiosyncracies of the dBA; or indeed the magic that goes on in the 'auditory system'.

The PPM is obviously a quasi-peak reading instrument and its specification is not likely to change for another 50 years, but it is more than likely to be joined by a Peak Digital Coding Level Indicator, which is a device which measures the peak Levels of digital signals after analogue to digital conversion - but has it got a future? It has been shown that if an overload reserve of about 10db is maintained on a digital distribution system it is still possible to achieve an excellent signal to noise ratio which is a very distinct advantage over the traditional analogue technology, and just as important it presents peak signals to the transmitter that are not clipped.

Hopefully this paper has launched the idea that when the PPM is used in conjunction with a good set of "guidelines" it take on a new roll - it monitors loudness.

CONTROLLING THE LOUDNESS OF TELEVISION SOUND

References

1 Robert B. Monroe "A New Approach to Balanced Audio Levels in Television" SMPTE Journal August 1959.

2 Benjamin B. Bauer and Emil L. Torick "Researchers in Loudness Measurement" N.A.B. Convention March 1966

3 "Determination of the Subjective Loudness of a Broadcasting Programme" CCIR Report 465-3 (1970 - 1982)

4 Bronwyn L. Jones and Emil L. Torick "A New Loudness Indicator for Use in Broadcasting" SMPTE Journal Sept 1981

5 Brian C.J. Moore "An Introduction to the Psychology of Hearing" Published 1989

I would like to thank Richard Ellis the Chief Engineer of Granada Television for his encouragement and permission to publish this paper.