ANNOYANCE OF TONAL NOISE: A PARAMETRIC STUDY

D W Robinson

Institute of Sound and Vibration Research, University of Southampton, UK.

1. BACKGROUND

The four physical characteristics of environmental noise upon which depends its numerical rating are level, duration, intermittency (or variability) and quality. Quality is determined primarily by frequency spectrum, but also by impulsive nature and tonal character when present. This paper concerns only the last-mentioned.

Although, physically, tonal character belongs to the spectrum domain, it confers on environmental noise an annoying quality for which some adjustment of its rating is needed in comparison to a broadband noise of similar level. The origin of a "tone correction" can be traced to Rosenblith *et al.* (1953) in a prediction scheme for community reaction. It took a form equivalent to adding 5 dB to the measured level, and has persisted in this form to the present day.

The first edition of BS 4142 appeared in 1967 and specified a correction of 5 dB(A) when there was a "distinguishable discrete continuous note". Since then there have been numerous criticisms of this all-or-none rule. Replies to a questionnaire issued to users by the British Standards Institution in 1970 already indicated that the correction was felt in many cases to be too small to account for public complaints. However, no change was made since to substitute a larger value would have placed too great a premium on deciding subjectively what is "distinguishable" and "discrete". Fisher (1982) found that 60% of local authorities encountered difficulty with the Standard. In a later survey of 63 local authorities, reported by Williams and Robinson (1988), 40% of those with experience of tonal noise found no particular fault with the 5 dB correction, but of the remaining majority those who considered it insufficiently severe outnumbered those of the opposite opinion by 10 to 1. Several respondents would have preferred a graded system, dependent upon the individual circumstances. Dissatisfaction with the fixed 5 dB correction was most marked in the case of those local authorities possessing a capability for spectrum analysis.

The 5 dB correction was also specified on the international level in ISO R1996:1971, ratified as international standard ISO 1996 in 1975 and only amended comparatively recently. In the new ISO 1996-2:1987, there is advisory wording that 2-3 dB may be applied if a tone is *just* detectable by the observer and is demonstrable by narrow-band frequency analysis, or 5-6 dB is the tone is *clearly* audible and can be detected by 1/3-octave band analysis. However, these prescriptions still do not amount to a watertight specification. The new ISO standard has been implemented in UK as BS 7745:1991, but it stands alongside the 1990 revision of BS 4142 which continues to advocate the fixed 5 dB correction.

The importance of the matter is well attested. Among various estimates of the percentage of complaint cases involving tonal noise are 42-65% (Fisher, 1982), 40-50% (Haeberle and Schmid, 1980) and 50% (Williams and Robinson, 1988a) excluding domestic noises. Several studies have been reported (e.g. Wells and Blazier, 1962; Fidell and Teffeteller, 1978; Hellman, 1984) bearing on the magnitude of a graded tone correction and its dependence on the various physical parameters of the noise; there is also a well-established corpus of knowledge on the detection of tones in noise and the mutually inhibiting effect of tone and broadband noise on the total perceived loudness of the combination. However, the factor which has tended to inhibit adoption of these more complicated tone correction schemes, based on instrumental measurements rather than subjective observation, is the unavoidable complexity of resolving tone from complex. A graded system is, in fact, used in aircraft noise certification (ISO 3891:1978) but is of a highly empirical nature and its validity has been questioned (e.g. Ollerhead, 1971; Scharf *et al.*, 1977). With present-day techniques of on-line digital signal processing, this complexity has become less of a barrier. The investigation described in this paper was carried out with the aim of determining experimentally the requisite characteristics to be specified, and developing a quantitative formula, to underpin such a development.

2. BASIC CONCEPTS AND TERMINOLOGY

The way in which BS 4142 and ISO 1996 handle tonal character is by adding a correction to the A-weighted sound pressure level of the noise. This is simple in practice but is not the most revealing way of representing the underlying phenomena. Experiments, including those described here, have shown that the true "correction" defined in this way can sometimes be negative in sign, although more frequently it is positive.

Instead of focussing on the combined signal, it assists to consider the two components separately. If one adds a discrete tone to a broadband noise three things happen. First, the sound pressure level is itself raised (the increment will be denoted by ΔL); secondly the loudness also increases (but not in general by as much as ΔL); thirdly, the perceived quality change in the signal engenders annoyance over and above that represented by the increment in loudness. Whilst physical principles alone determine ΔL, and the theory of loudness permits calculation of the loudness increment in phons (ΔN), there is no equivalent theory to estimate annoyance (or specifically the increment in annoyance, since it is to be assumed that the broadband component of an environmental noise already evokes some annoyance). This increment can, however, be obtained experimentally by observing how much the broadband component, on its own, would have to be raised in level to match the combination for annoyance (Figure 1). This increment in level, symbol Δ, is referred to here as "overall tone penalty" (OTP). The difference.between this quantity and the physical increment, *viz.* $\Delta - \Delta L$, corresponds to the "tone correction" in the above-mentioned Standards; here it is referred to as "net tone penalty" (NTP).

It turns out that the parametric variation of tonal annoyance (with frequency, absolute level, etc.) is expressible neither as a direct function of NTP nor of OTP, but more nearly as a function of the intermediate variable ($\Delta - \Delta N$). The practical way in which this has been realized, without resorting to the considerable complication of loudness calculations, is described later.

3. PRELIMINARY EXPERIMENTS

3.1 *Detectability and tonality*

A "tonal" component, when present, is not always a simple pure tone. Experiments were first carried out on the detectability of pure tones and other signals added to a broadband noise. Subjects were tested under binaural free-field listening conditions, using the self-recording technique. Thresholds of the pure tones were in close agreement with classical critical band theory, but narrow-band signals exhibited small elevations of up to 2 dB for 1/9-octave bands of random noise (even though this bandwidth is well below the critical), and from 2 to 4 dB for 1/3-octave noise, depending on the centre frequency of the signal. These experiments were repeated under a different protocol with the subjects engaged in a loading task with high attention demand. Thresholds were determined by increasing the tonal component at a slow rate until the subject signalled an awareness of a *change* in the background. Under these conditions, thresholds were on average 10 dB higher than before if the added signal was continuous, but only 4 dB higher when pulsed at 2 Hz. When the broadband component was itself mildly fluctuant (using the recorded noise of dense road traffic), the threshold elevations were smaller (av. 6 dB continuous mode, 3.5 dB pulsed).

Real-life conditions lie arguably between full attention to the sound and full distraction from it, represented by the two test conditions, and the above results suggest that an appropriate allowance might be made. In practice, this would mean subtracting some 5 dB from the true sensation level (SL) of a tonal component extracted by frequency analysis.

In a further series of tests, "tonalness" was judged on a scale from 0 to 9, with stimuli consisting of broadband noise with a range of added signals at individually-determined sensation levels of 5 and 20 dB. Mean scores of 6-8 points (high tonality) resulted at all frequencies for pure tones, even at the low level of 5 dB SL. A 1/9-octave noise and a 4-tone complex of similar bandwidth elicited scores from 3 to 6 at 5 dB SL, whilst 1/3-octave noise gave scores below 3; on average, scores were 1.4 points higher at 20 dB SL. Whole-octave noise added to the broadband background elicited very low mean scores. It thus appears that a pure tone at 5 dB SL would qualify as clearly "distinguishable" and accordingly attract the standardized 5 dB correction. A component of finite, though narrow, bandwidth (up to 1/3-octave or so) would have to be several decibels more pronounced to qualify.

These preliminary tests have been described in detail by Williams and Robinson (1988b) and Robinson (1988). The results may be invoked to temper the rigour of the tone penalty requirements based upon the annoyance tests using pure tones in idealized listening conditions which form the principal part of this paper.

3.2 *Annoyance matching – pilot tests*

Two broadband noises were synthesized, the spectra being representative of the range recorded in a mixed industrial and residential area of central England. The two spectra differed by a mean slope of 4 dB/octave over the range 50 to 4000 Hz. These noises were reproduced in a listening room, with and without the addition of single tones at 250 or 2000 Hz. Immediately following measurements of the masked thresholds for the set of stimuli, the tone levels were set individually for 18 young listeners at 5 or 20 dB SL. Judgements of annoyance were taken on a scale from 0 to 9. For the mixed stimuli, the broadband component was maintained at 40 dB(A) in the first and 60 dB(A) in the second of two sessions of testing. The broadband noises were also presented at five levels from 40 to 60 and 60 to 80 dB(A) respectively, interlaced with the mixed stimuli in random order, and the mean judgements plotted against level provided the calibration curves (as in Figure 1), from which the level of the broadband noise matching each of the mixed stimuli for annoyance could be read off.

The OTP showed only a small variation with background spectrum slope and with tone frequency, but a significant drop between the 40 and 60 dB(A) conditions of the broadband component (from 8.9 dB to 5.5 dB); not surprisingly, the major determinant was sensation level (average 1.3 dB and 7.2 dB at 5 and 20 dB SL respectively). The relative importance of the four variables (sensation level; absolute level; tone frequency; spectrum shape) determined the protocol for the more extensive annoyance tests next descibed.

4. MAIN ANNOYANCE TESTS

Three series of subjective tests were carried out on the same lines as just described, in which the variation of OTP was explored systematically with respect to the variables:

Series I: SL range 5-26 dB; tone frequency 150-4800 Hz; absolute level 20-55 dB(A); one broadband noise spectrum designated "L" representative of the central range of environmental noise; three sessions (low, medium and high absolute level)

Series II: SL range 5-26 dB; tone frequency 300-2400 Hz; absolute level 30-55 dB(A); three noise spectra (as in Series I plus two others with further ±3 dB/octave slope imposed, designated "H" and "F" respectively); two sessions (medium and high absolute level)

Series III: Partial repeat of Series II with the stimuli presented in a different order; medium absolute level; three sessions ("F", "L", "R" spectra respectively).

As before, masked thresholds of the tones occurring in each session were first obtained for individual subjects, enabling the tones for annoyance judgements to be set at the (four) requisite sensation levels. Calibration curves were also obtained from judgements on the broadband noises alone over a range of 0-30 dB(A) relative to the lowest absolute level for the spectrum or spectra in use in a given session. Values of OTP were derived as before using the respective calibration curves. Direct measurements of the A-weighted SPL of every stimulus permitted determination of the physical level increments ΔL attributable to adding the tones. The results have been given in detail by Robinson and Wright (1991); the principal features are as follows.

Typical results are illustrated in Figure 2, indicating a roughly linear (but mildly accelerating) relation between OTP (Δ) and SL. The physical increments, ΔL, of course, follow the power summation rule and are all of the same shape but displaced horizontally depending on the spectrum shape of the broadband component (only "L" in the Figure) and tone frequency. Averaging the results for the three level ranges showed that OTP falls consistently at the rate of about 2% per decibel of absolute level.

The variation of OTP with tone frequency showed a generally rising trend, but this was complicated by an interaction with spectrum shape which had the effect of amplifying the upward trend in the case of spectrum "R" and reversing it over the middle frequency range in the case of "F". The analysis which follows shows how this interaction was resolved.

Some context-dependent effects were observed within the results. In Series I, the 2%/dB drop of OTP with level refers to the *intersession* effect; within each session, two levels of the broadband noise separated by 15 dB(A) had been tested, and the *intrasession* effect was substantially larger (Figure 3). Between Series II and III, the variety of stimuli presented in a given session was deliberately altered, those in Series III being considerably more restricted in variety and resulting in larger values of OTP.

The slopes of individual regression lines between annoyance score and SL, when converted to slopes of Δ versus SL, gave a grand mean of the order 0.5 dB/dB. The reliability of the results was expressed by 95% confidence limits on the coefficient of variation of slope, which were of the order ±30% for each combination of tone frequency, spectrum shape and absolute level. Much narrower limits applied to results averaged over frequency, level or spectrum shape, showing that each identified main effect was statistically significant.

5. ANALYSIS OF MAIN ANNOYANCE TEST RESULTS

As a first step, the results of Series I, II and III were amalgamated. Since no preference could be assigned to one set of test conditions over others, an average was deemed to be the best representation. Since the three series only overlapped in part, the process of averaging was complicated; the details have been described by Robinson (1992).

The principal stumbling block in finding an analytical expression for OTP was the afore-mentioned interaction between the variations with tone frequency and spectrum shape. Simple devices, such as correcting the values by a quantity dependent on both factors (e.g. the difference between the A-weighted SPL of the whole broadband noise and that of the 1/3-octave band embracing the tone frequency) proved ineffective.

Inspection of the results showed that the level dependence was very consistent across the other variables and could be extracted as a multiplying factor falling exponentially with level. Secondly, to a first approximation, the OTPs were roughly proportional to SL, permitting SL to be provisionally eliminated by representing each set of results by the slope of Δ on SL. With these simplifications, the interaction of tone frequency and spectrum shape was brought into sharper focus. Trials showed that use of the NTP in place of the OTP exacerbated the interaction, whereas an intermediate variable $(\Delta - k.\Delta L)$ replacing Δ would be effective for some value of k between 0 and 1. The procedure used to estimate k was to minimize the variance relative to a model in which the dependence on frequency and spectrum shape appear as independent (multiplicative) factors. The process was, of necessity, iterative since the slope of $(\Delta - k.\Delta L)$ on SL itself depends on k, and it was necessary to monitor that the relation between the modified variable and SL remained approximately linear; fortunately this was the case, and an optimum value of 0.38 for k was established. It is worth noting, and possibly of theoretical significance, that this factor is quite close to the typical ratio between ΔN (the calculated loudness increment) and ΔL. Once k was fixed, proportionality factors for the frequency and spectrum shape effects fell out automatically, leaving only the relation between these factors and the respective physical variables to be finalized.

In the case of frequency variation, the required factor was a rising function with a flatter portion in the middle range. This was expressed as a polynomial in log(frequency), normalized to 1 at 1000 Hz; it is illustrated in Figure 4. Spectrum dependence is, of its nature, more complicated in principle since not all spectra can be compared by means of a single-figure parameter. Nevertheless, the experimental results could be encompassed to a high accuracy in this way, by means of the factor $(17 + L_C - L_A)$, where L_C and L_A are respectively the C- and A-weighted SPLs of the broadband noise component.

6. GENERAL FORMULA AND DISCUSSION

Overall tone penalty is given by the formula:

$$\Delta = \{0.051 \times (17 + L_C - L_A) \times 10^{-L_A/92} \times \Phi(x) \times SL\} + 0.38\Delta L$$

where $x = \log(f/1000\text{ Hz})$, f is the tone frequency, and $\Phi(x)$ is the function shown in Figure 4. Net tone penalty, NTP, corresponding to the tone correction in standards documents, is given by the same expression with the second term changed to $-0.62\Delta L$.

Comparison of calculated and experimental values for the 162 data points of the experiments yielded a root-mean-square error of only 1.2 dB, most of which is attributable to experimental uncertainty within each datum.

To apply the formula, it is necessary to determine the SL of the tone, which can be done from the formula:

$$SL = L_{At} - L_{Af} + (5.44 + 3.10x - 5.26x^2)$$

where L_{At} is the A-weighted SPL of the tone,
L_{Af} is the A-weighted 1/3-octave band SPL of the broadband noise centred on f,
and x is as above.

In practice, it may be expedient to subtract, say, 5 dB from this calculated value of SL to simulate real-life conditions, as already described. The applicable ranges of the variables and indications for the signal processing paradigms required to find the quantities in the formulae are given in the report of the analysis (Robinson, 1992).

Comparisons were made, for a sample of 90 tonal noises, between the calculated NTP and the values that would apply according to present standards. Examples with calculated NTP from +0.8 to +3.6 dB would conventionally attract no correction (and this range could well be wider for a different set of examples). Conversely, noises that would attract the conventional 5 dB correction ranged in calculated NTP from −1.0 to as high as +21.6 dB.

Finally, it should be remembered that the sound levels in the experiments, and by implication in the formulae, are those at the ears of the listener. Since it is conventional to measure environmental noise outside rather than inside premises, certain adjustments ought to be applied to the formulae. The terms affected are the exponential factor for level, and the spectrum factor involving L_C and L_A. This could be done to an adequate accuracy by setting a standardized relation for outside-to-inside sound transmission.

7. ACKNOWLEDGEMENT

The work described was carried out in two phases under contract to the Department of the Environment and this summarizing paper is published by permission of the Department.

8. REFERENCES

Fidell, S. and Teffeteller, S., 1978. *The relationship between annoyance and detectability of low level sounds.* Bolt, Beranek and Newman Report 3699.

Fisher, P.S., 1982. *The assessment of industrial noise nuisances involving tonal character.* MPhil Dissertation, Open University.

Haeberle, M. and Schmid, D., 1980. Report – It is the information content of the noise that is disturbing. *Applied Acoustics* 13, 237-244.

Hellman, R.P., 1984. Growth rate of loudness, annoyance and noisiness as a function of tone location within the noise spectrum. *J. acoust. Soc. Am.* 75, 209-218.

Ollerhead, J.B., 1971. *An evaluation of methods for scaling aircraft noise perception.* NASA Contractor Report 1883, Washington, DC.

Robinson, D.W., 1988. *The assessment of noise containing tonal components.* ISVR TR 163, University of Southampton.

Robinson, D.W., 1992. *Annoyance due to discrete tones in broadband background noise: Part II – Analysis.* ISVR TR 202, University of Southampton.

Robinson, D.W. and Wright, P., 1991. *Annoyance due to discrete tones in broadband background noise: Part I – Experiments.* ISVR TR 201, University of Southampton.

Rosenblith, W.A. and Stevens, K.N., 1953. *Handbook of acoustic noise control: Vol. 2, Noise and man.* WADC TR-52-204, 181-200.

Scharf, B., Hellman, R. and Bauer, J., 1977. *Comparison of various methods for predicting the loudness and acceptability of noise.* US EPA Report 550/9-77-101, Washington, DC.

Wells, R.J. and Blazier, W.E., 1962. A procedure for computing the subjective reaction to complex noise. *Proc. 4th Int. Cong. Acoust., Copenhagen*, L24, 1-4.

Williams, P.R. and Robinson, D.W., 1988a. *A survey of practice and opinion of local authorities on the assessment of environmental noises containing tonal components.* ISVR Memorandum 686, University of Southampton.

Williams, P.R. and Robinson, D.W., 1988b. *The detectability and tonal character of tones and other narrow-band stimuli in noise.* ISVR TR 159, University of Southampton.

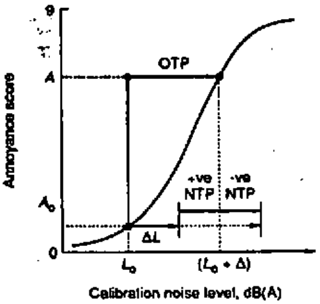

Fig. 1 Overall and net tone penalty. Broadband noise at level L_0 has annoyance A_0; addition of tone increases annoyance to A, which is also the annoyance of the broadband noise at level (L_0 + Δ). Tone also increases level by ΔL, so NTP is negative if ΔL exceeds Δ.

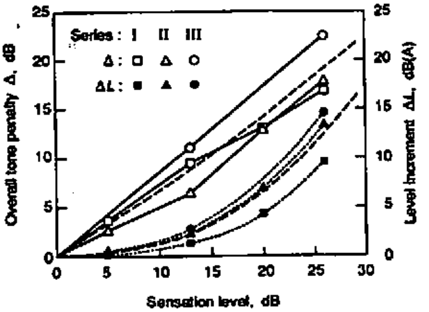

Fig. 2. Specimen results. OTP (open symbols) and ΔL (filled symbols) as functions of sensation level. Dashed curves are mean trends. Data are averaged over three frequencies and two levels.

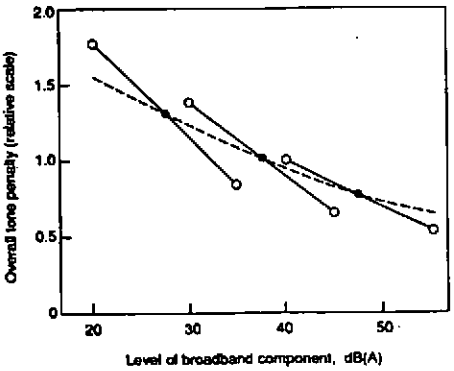

Fig. 3 Variation of OTP with absolute level. Solid lines: intrasession effect Broken line: mean intersession effect. Values are averaged over test conditions in Series I and normalized.

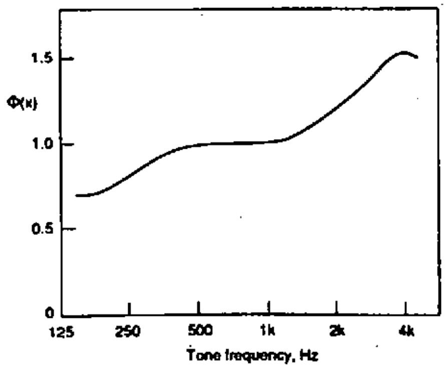

Fig. 4 Variation with tone frequency of the multiplicative term $\Phi(x)$ in the formula for Δ.