LATERAL FRACTION OR CROSS-CORRELATION FOR SPATIAL IMPRESSION?

Mike Barron Combe Royal Cottage, Bathwick Hill, Bath BA2 6EQ m.barron@flemingbarron.co.uk

1 INTRODUCTION

In ISO3382-1:2009. 'Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces' [1], measures beyond reverberation time for auditoria were considered for the first time. In the case of Spatial Impression (or Apparent Source Width, ASW) two measures are proposed: the Early Lateral Energy Fraction (ELEF) in section A.2.4 and the InterAural Cross-correlation Coefficient (IACC) in Annex B. No advice is provided regarding when to use one or the other. This also poses the question of whether they are equivalent in some situations.

This author presented a paper in 1983 [2] which developed a theoretical relationship between ELEF and IACC measurements. In 1994 Bradley reported on results of simultaneous measurements of ELEF and IACC in 11 halls [3] and found evidence to support the proposed theoretical relationship, but some anomalies arose.

This paper revisits the above with the aim of advancing understanding towards a preferred measurement technique for spatial impression due to early lateral reflections. The degree of spatial impression or Apparent Source Width is in fact a function of two quantities: a spatial one (as measured by the ELEF or IACC) and the early sound level. (Because of this, large crescendos tend to create a significant spatial expansion, which can add to the excitement of an orchestral performance.) The sound level aspect will be omitted from this paper; a proposal for combining the two components is offered in [4].

2 DEFINITIONS

The Early lateral energy fraction (ELEF) was developed from results of simulation experiments [4, 5]. The subjective effect was found to be relatively unaffected by reflection delay; spatially it was found to be related to the cosine of the reflection angle to the axis through the listener's ears, θ. For measurement, the ELEF is usually measured by comparing the energy of a figure-of-eight microphone (with the null pointing towards the source) with the energy received by an omni-directional microphone:

$$ELEF = \frac{\int_{0.005}^{0.08} p_L^2\,(t)\,dt}{\int_{0}^{0.08} p^2\,(t)\,dt} \qquad (1)$$

where $p_L(t)$ is the lateral pressure response from the figure-of-eight microphone and p(t) is the response from an omni-directional microphone. From subjective experiments, $p_L^2(t) = p^2(t).\cos\theta$, whereas with a figure-of-eight microphone the relationship is $p_L^2(t) = p^2(t).\cos^2\theta$. In most cases, the $\cos^2\theta$ relationship is accepted.

The ELEF is a dimensionless quantity, which is linearly related to subjective spatial impression [5].

The IACC is defined as follows:

$$IACF_{0,0.08}(\tau) = \frac{\int_0^{0.08} p_l(t) \cdot p_r(t+\tau)\, dt}{\sqrt{\int_0^{0.08} p_l^2(t)\, dt}\sqrt{\int_0^{0.08} p_r^2(t)\, dt}}$$

$$IACC_{0,0.08} = \max\left|IACF_{0,0.08}\right| \text{ for -1ms} < \tau < \text{+1ms} \qquad (2)$$

The denominator for the IACF is for normalisation, giving values of IACF between 0 and 1. The IACC is a measure of similarity between two signals, in this case between the left and right ears and allows for a relative delay. For spatial impression, one is interested in dissimilarity, (1 – IACC). The variable τ is the delay for lateral sounds between the left and right ears. Note that reflection delay specifies that early reflections up to 80ms relative to the direct sound are being considered, hence the upper limit of 0.08s of the integrals.

2.1 Frequency

Marshall [6] was the first to suggest that spatial impression was principally a low frequency effect. Subjective experiments [5] indicate that low frequency lateral reflections create a vertical as well as a horizontal spread in the perceived source area. For lateral reflections containing only frequencies above 1500Hz, source broadening takes place very much in the horizontal plane of the source, with no vertical spread. Subjective evidence [5] suggested that for an objective measure of source broadening (ASW) the frequency range from 90 to 1500Hz is appropriate. In line with other proposed objective measures, octave filtering is generally employed, which in this case means the 4 octaves between 125 and 1000Hz. Potter *et al.* [7] among others support the view that low frequencies are important for spatial impression.

The seat-dip effect [8], with attenuations in excess of 15 dB in the octave between 100 and 200 Hz, is potentially significant for spatial impression. However mean measured values of the ELEF [9] do not indicate any dip in that frequency region. Indeed for measurements in British halls (17 auditoria, 189 positions) the mean ELEF is virtually independent of frequency at around 0.19; on average the seat-dip effect is influencing the frontal and lateral sound by equal amounts.

3 RELATIONSHIP BETWEEN LATERAL ENERGY FRACTION AND DEGREE OF INCOHERENCE

To derive a relationship between ELEF and IACC, a number of assumptions are necessary, as follows:

1. Given that low frequencies are the main interest, amplitude changes due to head shadowing are ignored. In other words, we treat the 'head' as if the ears were represented by two omni-directional microphones the correct distance apart to match the maximum interaural delay (for fully lateral sound).
2. We assume that the early sound field is symmetrical, that reflected energy is balanced between left and right.
3. We further assume that the test signal is of short duration.

As derived in the Appendix, these assumptions lead to the following relationship:

$$\mathbf{1 - IACC = k * ELEF}\text{ , where } \mathbf{k = 2 * (1 - \overline{\gamma(\tau)})} \qquad (3)$$

$\overline{\gamma(\tau)}$ is the mean of the autocorrelation function of the filtered source signal, with the average taken between τ = 0 to the maximum interaural time delay, τ_{max}. (Octave or similar filtering is assumed; this filtering dominates the autocorrelation function.) A crucial observation here is that $\overline{\gamma(\tau)}$ is a function of measurement frequency.

The dominant influence on the source signal is the octave filtering; the impulse response of an octave filter is shown in Figure 1. Figure 2 is the initial part of the autocorrelation function of this impulse response, namely $\gamma(t)$; the delay range shown being relevant for the maximum interaural delay. The shape of the function over this delay range closely matches the first quadrant of a cosine function.

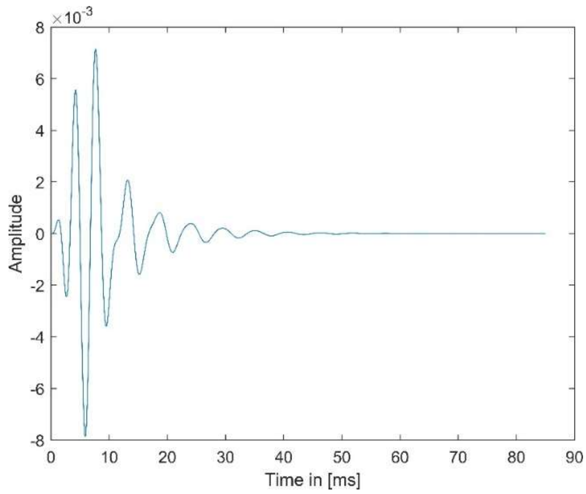

Figure 1. The impulse response of a 250Hz octave filter.

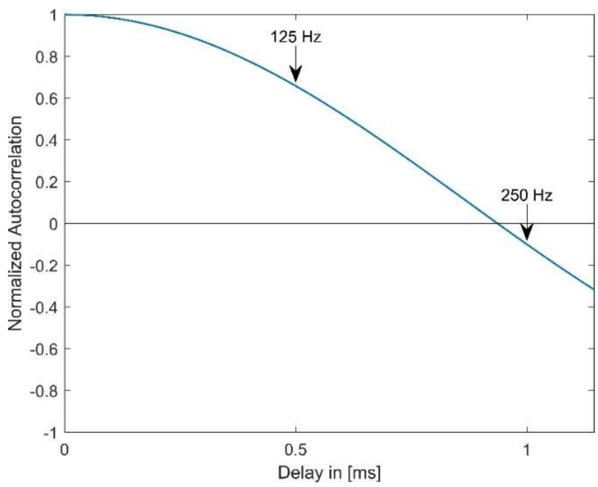

Figure 2 The initial portion of the normalised autocorrelation function of the 250Hz octave filter impulse response.

4 COMPARISON BETWEEN ELEF AND IACC ACCORDING TO HALL MEAN MEASUREMENTS

Bradley's comprehensive study involved measurements at 124 locations in 14 halls [3]. The IACC measurements were made with a Brüel & Kjaer Head and Torso Simulator. We start with the comparison between hall mean values of ELEF and (1 – IACC), Figure 3.

The first observation is that the two measures are linearly related, which is obviously encouraging. Secondly one notices that the measured ranges of the two quantities are different. The mean and range of the ELEF at the different frequencies is broadly the same. However the variation and values of (1 – IACC) at low as opposed to higher frequencies is much less, particularly at 125Hz. In spite of this, the correlation coefficients between measured ELEF and (1 - IACC) are virtually identical and highly significant ($p < 0.001$), as listed in Table 1. This suggests a relationship between these two quantities at these four frequencies. The ELEF was found to be a linear measure of subjective spatial impression [5], hence we can assume that IACC is also a linear measure.

The response of Hidaka *et al* [10] to the low values of (1 - IACC) at low frequencies was to ignore them to obtain a single value of (1 - IACC) and just assign a level contribution at 125 and 250Hz.

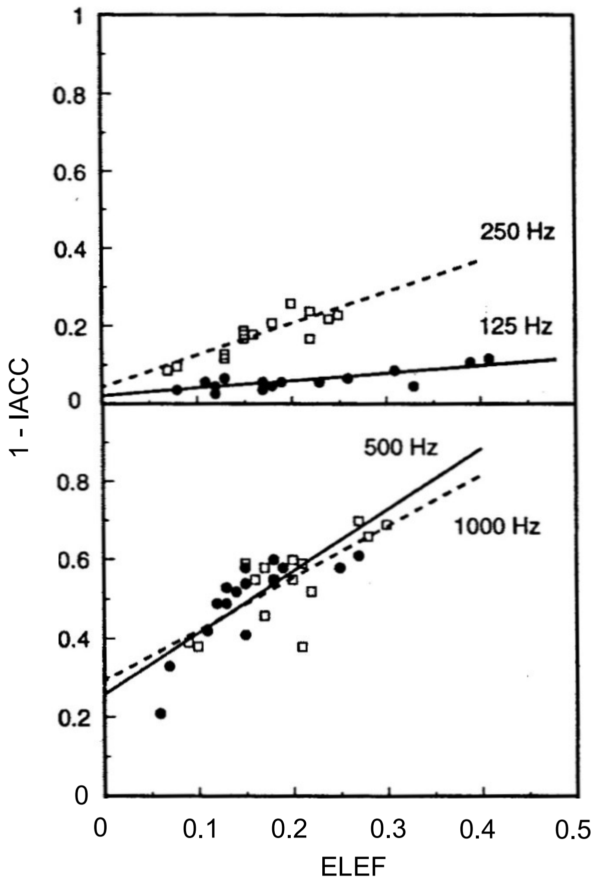

Figure 3. Comparison of hall mean values of ELEF and (1 - IACC) after Bradley [3, Fig. 3]

Table 1. Data relating to Figure 3, after Bradley

Octave frequency (Hz)	Correlation coefficient (r)	Slope (Figure 3)	k in equation (3)
125	0.80	0.20	0.20
250	0.85	0.81	0.81
500	0.80	1.61	2.00
1000	0.75	1.35	2.00

Table 1 and Figure 4 compare the slope of the hall mean comparisons with k from equation (3). There is identity at 125 and 250Hz but less agreement at higher frequencies. k is a function of $\overline{\gamma(\tau)}$, which depends on the value of the maximum interaural time delay, τ_{max}.

Bradley conducted an experiment in a reverberation chamber, which established that a separation of the omni-directional microphones assumed in the theory in section 3 of 0.3m is appropriate. This corresponds to a τ_{max} of 0.97ms. This is the time limit for the mean autocorrelation function to establish $\overline{\gamma(\tau)}$. Figure 2 shows the normalised autocorrelation function for a 250Hz octave filter. The arrow labelled "250Hz" shows the maximum interaural delay used for the mean at this frequency. The equivalent figure for 125Hz would be stretched between 0 and 2ms. For diagrammatic reasons in Figure 2, a 2nd arrow labelled "125Hz" is included. This shows the upper limit for the mean at 125Hz relative to the autocorrelation function. It is clear that the value of $\overline{\gamma(\tau)}$ is much higher at 125Hz than 250Hz and then smaller at higher frequencies. This gives values of k much smaller at 125Hz than 250Hz and above.

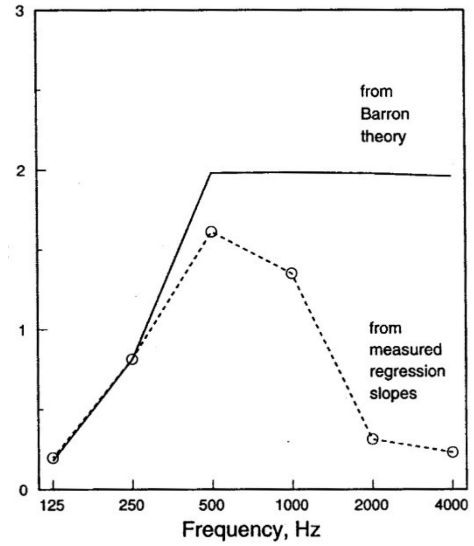

Figure 4. Comparison of the k parameter based on the autocorrelation function and slopes of regression lines from hall mean measurements in Figure 3 after Bradley [3, Fig.5].

Thus equation (3) explains the low values of (1 - IACC) at low frequencies and the reason for different slopes in Figure 3. Hence the equation provides a means for correcting this anomaly for measured values of (1 - IACC).

4.1 Measured (1 - IACC) for fully frontal sound

Returning to Figure 3, there is however a further anomaly: that the best-fit lines do not pass through the origin; zero ELEF corresponds to a finite value for (1 – IACC) at all octave frequencies. For fully frontal sound with no lateral components, the signals at the two ears are identical and hence (1 – IACC) is zero. For the measurement of ELEF, a point measurement of lateral sound is involved, as opposed to spaced microphones, so one would assume that a fully frontal sound produces an ELEF of zero.

There is no guarantee that the curves in Figure 3 remain linear for very low lateral sound. The plotted lines could curve to the origin, though the degree of non-linearity involved at higher frequencies would be extreme. For the moment, this behaviour is a mystery, which deserves further investigation.

The statement above regarding IACC measurement that "the signals at the two ears are identical" is not entirely true! The outer ears, the pinnae, of a dummy head are not mirror images of each other, hence there are slightly different signals at the eardrums for frontal sound. However for frequencies up to the 1kHz octave, the differences are unlikely to be significant.

4.2 Interaural level differences

At higher frequencies, level differences between the two ears become significant particularly for lateral sound. Differences of a decibel or so arise at 1kHz, which will therefore be relevant for measurement of the IACC. These level differences are omitted when a measurement is made with a figure-of-eight microphone. Nevertheless if the goal is a single number representation for spatial impression by averaging between the octaves 125 and 1000Hz, the omission of interaural level differences is numerically small. One can observe that the regression coefficient at 1000Hz compared with other frequencies in Table 1 is slightly less, though not markedly so.

5 COMPARISON BETWEEN ELEF AND IACC FOR INDIVIDUAL SEAT MEASUREMENTS

The above, based on hall mean values, provides encouraging evidence that equation (3) has validity but with two limitations: that finite cross-correlations exist for purely frontal sound and that assuming zero head shadow is only valid at lower frequencies. The comparison between ELEF and (1 - IACC) measured values for individual positions is less impressive with much greater scatter, Figure 5.

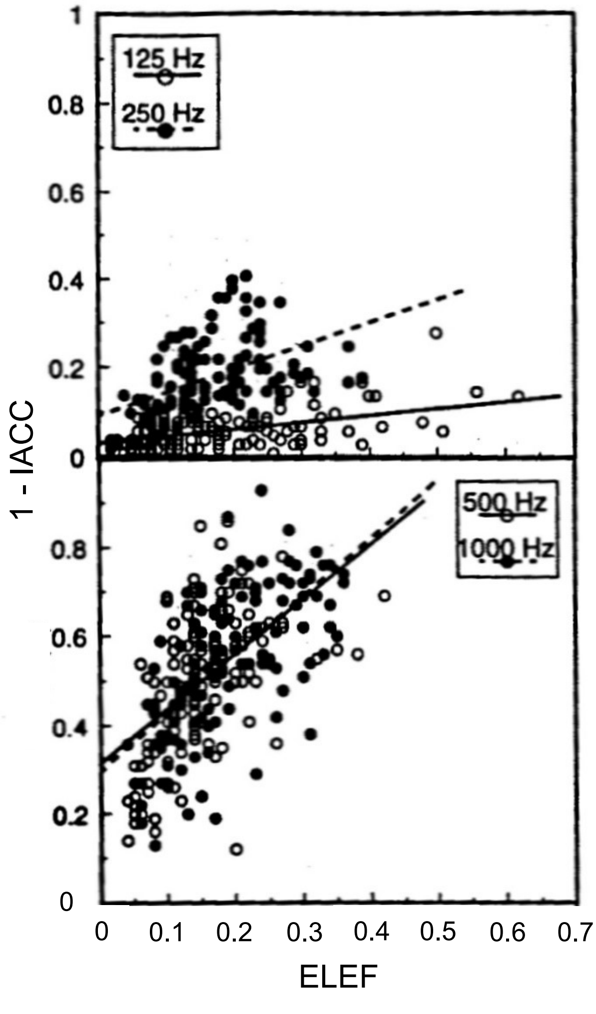

Figure 5. Comparison of individual position values of ELEF and (1 - IACC) after Bradley [3, Fig. 6]

In Figure 5, the correlation coefficients (125 – 1000Hz) are 0.62, 0.48, 0.53 and 0.62. Given the high number of points (124), these are all highly significant ($p<0.001$). However to quote Bradley: "knowing the value of a particular ELEF would not allow one to make an accurate estimate of the corresponding IACC value". By averaging values of ELEF and (1 – IACC) within a hall (Figure 3), the scatter is greatly reduced.

The principle cause of additional scatter for individual positions is most likely to be interference between different reflections. This was excluded from the theoretical derivation with assumption 3. In reality the effective length of the test signal includes octave filtering, so is no longer short. When a single value for ELEF is derived by averaging over the four octave values [5], the scatter observed in Figure 5 will almost certainly be reduced. Conducting a similar averaging process for IACC by correcting for frequency effects at low frequencies etc., there might be a reasonable relationship between single values of ELEF and (1 – IACC). This however is speculation. Sadly the numerical data from Bradley's measurements is no longer available so this frequency averaging cannot be conducted.

6 CONCLUSIONS

Based on certain assumptions, a theoretical relationship, equation (3), between ELEF and (1 – IACC) was developed. The theory shows that the constant relating the two quantities is a function of the normalised autocorrelation function of the effective source signal. Bradley [3] has conducted measurements of both quantities in concert halls and found that for hall mean values there is indeed reasonable evidence for the validity of the theoretical relationship .

The theory provides an explanation for the behaviour already observed elsewhere of low values of (1 – IACC) at low frequencies, particularly at the 125 and 250Hz octaves (Figure 3). However the relationship between ELEF and (1 – IACC) at these frequencies has a high regression coefficient so that simple correction for IACC values would be valid.

The comparison shown in Figure 3 also shows another 'anomaly' with measurements of (1 – IACC): that zero spatial impression does not correspond to a zero value of (1 – IACC). From the definitions of ELEF and (1 – IACC), a fully frontal sound field with no lateral components should result in zero values for both quantities. The reason for this behaviour with (1 – IACC) is not obvious, but should be capable of explanation following simple experimentation with signal analysis software etc. Nevertheless the linear relationships in Figure 3 indicate a possible correction process for (1 – IACC).

When one observes Bradley's plotted data relating measured ELEF and (1 – IACC) for individual positions in Figure 5, one finds that there is too much scatter for a reliable relationship between the two measured quantities. Unfortunately the data is no longer available to see if this remains the case for single number values of the two quantities by averaging over frequency.

ELEF looks like a reliable measure for spatial impression but the way our ears derive this sense involves comparison of the signals at the two ears, namely some form of cross-correlation. The simple form of correlation in the IACC, equation (2), has been shown to have some inherent problems. With suitable corrections these problems may be surmountable. Further investigations are needed.

7 ACKNOWLEDGEMENTS

Thanks are due to John Bradley for independently deciding 27 years ago that my theory was worth investigating. Thanks are also due to Evan Green and Otavio Colella Gomes, both of Kahle Acoustics, for help with Figures 1 and 2.

8 REFERENCES

1. ISO3382-1:2009. Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces (2009).

2. M. Barron, 'Objective measures of spatial impression in concert halls', *Proc. of the 11th ICA*, *Paris*, Vol.**7**, 105-108, Paper 7.2.22 (1983).

3. J. Bradley, 'Comparison of concert hall measurements of spatial impression', *J. Acoust Soc. Am*. **96** 3525-3535 (1994).

4. A.H. Marshall and M. Barron, 'Spatial responsiveness in concert halls and the origins of spatial impression', *Applied Acoustics* **62** 91-108 (2001).

5. M. Barron and A.H. Marshall, 'Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure', *J. Sound Vib.* **77** 211-232 (1981).

6. A. H. Marshall, 'Levels of reflection masking in concert halls', *J. Sound Vib.* **7** 116-118 (1968).

7. J.M. Potter, F.A. Bilsen and J. Raatgever, 'Frequency dependence of spaciousness', *Acta Acustica* **3** 417-427 (1995).

8. T.J. Schultz and B.G. Watters, 'Propagation of sound across audience seating', *J. Acoust Soc. Am*. **36** 885-896 (1964).

9. M. Barron, 'Measured early lateral energy fractions in concert halls and opera houses', *J. Sound Vib.* **232** 79-100 (2000).

10. T. Hidaka, L.L. Beranek and T. Okano, 'Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls', *J. Acoust Soc. Am*. **98** 988-1007 (1995).

9 Appendix: Derivation of the theoretical relationship

Lateral reflections arriving within 5ms of the direct sound are rare. If we ignore these, then

$$ELEF = \frac{\sum_{n=1}^{N} r_n \cos\theta_n}{\sum_{n=1}^{N} r_n} = \frac{\sum_{n=1}^{N} r_n \left(\frac{|\tau_n|}{\tau_{max}}\right)}{\sum_{n=1}^{N} r_n} \qquad \text{(A1)}$$

where r_n is the energy of the nth reflection and θ_n is its angle of incidence relative to the axis through the two ears. τ_n is the interaural time difference for the nth reflection, while τ_{max} is the maximum interaural time difference. The direct sound is represented by n = 1. (Note that the subjectively appropriate 'cos θ' definition is used here as mentioned just below equation (1), because it matches the cross-correlation measure.)

Turning now to the short term correlation coefficient, it is necessary to consider the cross-correlation function between signals at the two ears for the nth reflection, $\phi_{lr}^{n}(t)$. For a short duration test signal, two reflections are in general mutually incoherent and it can be shown that the cross-correlation function for reflections n and m is:

$$\phi_{lr}^{n+m}(t) = \phi_{lr}^{n}(t) + \phi_{lr}^{m}(t) \qquad \text{(A2)}$$

This can be generalised for N reflections as follows:

$$\Phi_{lr}^{N}(t) = \sum_{n=1}^{N} \Phi_{lr}^{n}(t) \qquad \text{(A3)}$$

The normalised cross-correlation function is then:

$$\Phi_{lr}^{N}(t) = \frac{\sum_{n=1}^{N} \phi_{lr}^{n}(t)}{\sqrt{\sum_{n=1}^{N} \phi_{ll}^{n}(0) \,. \sum_{n=1}^{N} \phi_{rr}^{n}(0)}} \qquad \text{(A4)}$$

Invoking the assumption of no head shadow, the ears are equivalent to two omni-directional receivers spaced a distance $c\tau_{max}$ apart (c = speed of sound).

$$\phi_{lr}^{n}(t) = r_n \,. \delta(t - \tau_n) \qquad \text{(A5)}$$

And since the cross-correlation function for an arbitrary signal is obtained by convolution with the autocorrelation of the signal, $\phi_{ss}(t)$:

$$\phi_{lr}^{n}(t) = r_n . \delta(t - \tau_n) * \phi_{ss}(t) = r_n \,. \phi_{ss}(t - \tau_n) \qquad \text{(A6)}$$

By similar arguments:

$$\sum_{n=1}^{N} \phi_{ll}^{n}(0) = \sum_{n=1}^{N} \phi_{rr}^{n}(0) = \sum_{n=1}^{N} r_n \,. \phi_{ss}(0) \qquad \text{(A7)}$$

If the normalised autocorrelation function of the test signal is $\gamma_s(t)$:

$$\gamma_s(t) = {\phi_{ss}(t)}/{\phi_{ss}(0)} \qquad \text{(A8)}$$

then the normalised cross-correlation function is:

$$\Phi_{lr}^{N}(t) = \frac{\sum_{n=1}^{N} r_n \,. \phi_{ss}(t - \tau_n)}{\sum_{n+1}^{N} r_n \;\; \phi_{ss}(0)} = \frac{\sum_{n=1}^{N} r_n . \gamma_s\,(t - \tau_n)}{\sum_{n=1}^{N} r_n} \qquad \text{(A9)}$$

For a symmetrical sound field, the maximum value of $\Phi_{lr}^{N}(t)$ occurs at t = 0. If we replace $\Phi_{lr}^{N}(t)$ by $\Phi(0)$ for simplicity,

$$1 - \Phi(0) = 1 - \frac{\sum r_n \, \gamma_s(\tau_n)}{\sum r_n} = \frac{\sum r_n (1 - \gamma_s(\tau_n))}{\sum r_n} \qquad \text{(A10)}$$

Comparing equations (A1) and (A10), the relationship between ELEF and IACC is seen to be dependent on $\gamma_s(t)$, the autocorrelation function of the test signal for cross-correlation. The form of relationship we are seeking is:

$$L_f = \frac{1}{k}(1 - \Phi(0)) \qquad \text{(A11),}$$

where k is a constant determined by the signal autocorrelation function. To determine the value of k, it is necessary to assume a directional distribution of reflected energy, a uniform distribution in the hemisphere above the listener is an obvious assumption. In this condition the mean value for $|\tau_n|/\tau_{max}$ is ½ and the mean value of $(1 - \gamma(\tau))$ is $(1 - \overline{\gamma(\tau)})$, where $\overline{\gamma(\tau)} = \frac{1}{\tau_{max}} \int_0^{\tau_{max}} \gamma(\tau). d\tau$.

$$\text{So } k = 2 * (1 - \overline{\gamma(\tau)}) \qquad \text{(A12)}$$