

Articulation and the Small Room

by

Arthur M. Noxon

Acoustic Sciences Corporation

Eugene, Oregon USA

This paper first presented at the 85th AES Convention

Los Angeles, CA, November 6, 1988

Summary

This paper discusses the physics of how articulation relates to high end audio, and the steps involved in measuring it using the Modulation Transfer Function and Transmission Index. The research presented here led to the development of the industry standard ASC MATT Test.

INTRODUCTION

The Modulation Transfer Function (MTF) is used in room acoustics as a descriptor of the effectiveness of transmission down the signal path, between the speaker and listener. A major application for this has been speech intelligibility. Basis for MTF analysis is the signal to noise ratio. Noise can be any sound masking effect, steady state noise, transient noise of reverberation or apparent noise due to adjacent octave sound levels.

Narrow band MTF is used in the present work. This is in contrast to the octave band methods common to traditional speech intelligibility. Here, pure tone modulation used to develop spectral response detail. A rapidly gated, slow sine sweep is the test signal for the articulation response curve. This technique allows blurred transmission bands to be specifically identified. These narrow ranges of poor articulation are both audible to the listener and visible in hard copy data. Changes to the room acoustic are also easily documented. The responsiveness of this test to room acoustics in addition to the fine grain spectral information in the articulation response curve suggests that this system be used as a diagnostic tool. Although originally developed to demonstrate small room acoustics in the lower registers, it has found use in the full range of room sizes from the amphitheater and auditorium right through to recording studio vocal booth.

I ARTICULATION RESPONSE CURVE (ARC)

A. RATIONALE

The Modulation Transfer Function (MTF) is used in room acoustics as the descriptor of effective signal transmission between speaker and listener. A popular application of the MTF is for speech intelligibility. Here we look at an application of MTF developed for precision playback environments such as the hi-end, hi-fi listening room and the recording studio. The suitability of the standard Speech Transmission Index (STI) approach falls short on numerous points in these smaller spaces that have high musical articulation requirements.

The spectrum segment useful for STI prediction or measurement starts at the 125 Hz band and each octave band is weighted for significance in speech recognition. Music occupies two octaves lower than the range used for STI work, half the keyboard is below Middle C 2 5 2 . The weighting of these and other octaves in a calculation is not yet established. The musical spectrum and the relative significance of each octave band may well not be the same as for speech. The Music Transmission Index (MTI) may be convertible to STI, but the converse may not be possible. This would be due to the relative lack of full bandwidth information in the STI. Clearly, research remains to be done in this area.

The STI joins the group of single index acoustic descriptors, such as NRC, dB, A, IIC, RT60, et. al. Architectural specifications can be satisfied with a single index indicator. Acoustical engineers and consultants engaged in diagnosis and remedy have always required spectral detail and the subject of intelligibility is no different.

Measured STI only needs the signal to noise ratio to be detected. Tracking octave band decay rates is one method used and monitoring modulated octave band noise levels is another. Both use selected octave bandwidths and yield a single intelligibility rating. The approach contributes little to the diagnosis of room acoustics. The present technique provides narrow band spectral articulation information. This facilitates diagnostic efforts and evaluation of STC.

The predictive side of MTF analysis requires the ability to accurately estimate the signal to noise ratio. The noise level is due to the reflections in the room and due to its reverberation. Predictive methods that use room reverberation decay rates have the prerequisite imposed that the room sound field is instantaneously diffuse and has an exponential decay rate.

A non-linear method of predicting noise levels is to use ray tracing of the first 30 reflections. This method better correlates with measured STI. Complex room geometry limits this method. Neither linear acoustics nor ray tracing can be used for predicting in small rooms dominated by room resonant mode decays.

The musical line is characterized as a rapid staccato of complex tone bursts. Music then is a set of musical lines, overlaid and intertwining one another. The basic element of this woven fabric of music is the tone burst. The acoustic descriptor that relates to musical articulation may well be the tone burst, indeed a rapid staccato of bursts. Such a signal has been used for harmonic distortion analysis room acoustic transmission path. Here we only desire measurement of the signal envelope and the faithfulness of its modulated transmission. Wave form reproduction, although important, is not the issue addressed.

A synthesis of these constraints and requirements is embodied in the present approach to MTF. The Articulation Frequency Response Curve (AFC) is a relatively simple, direct physical measurement. Equally important is the subjective aspect. The auditor in a precision listening setting can play the test signal over headphones and hear the rapid, clean staccato of tone bursts whose frequency is slowly varied. The auditor expects the room acoustic to play this signal accurately. By removing the headphones and listening to the same signal in the playback room, defects in the transmission path become quite audible. In a small room, articulation dramatically varies with frequency. Typically, there are tenth-octave bands of totally garbled transmission adjoining similar sized bands of quite intelligible transmission. The Articulation Response Curve is a fine-grained quantification of the "fast tracking" ability of a listening room.

II COMPARISON WITH TRADITION

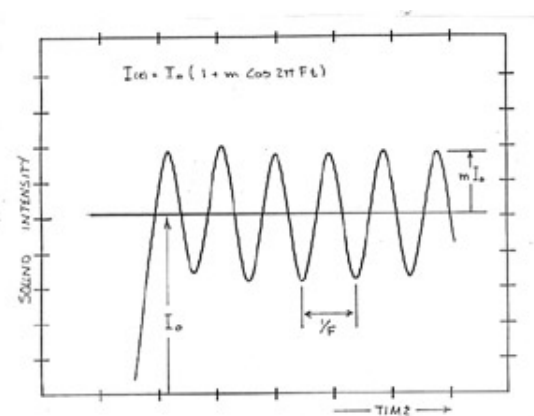
A. DEFINITION OF STANDARD TERMS

1. Signal Intensity (I)

Standard MTF format assumes the sound intensity envelope is a modulated cosine with a DC offset.

$$I(t) = I_0 \times (1 + m \cos 2\pi Ft)$$

The mean signal intensity (I_0) is modulated by the modulation amplitude (mI_0).

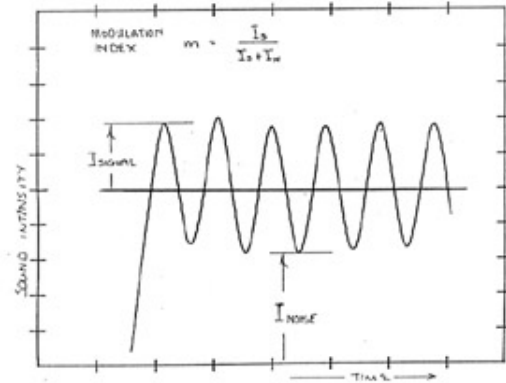


2. Modulation Index (m)

The modulation index is defined as the ratio of the intensity of the modulation to the mean intensity of the signal, modulation plus noise.

$$m = \frac{mI_0}{I_0} = \frac{I_s}{I_s + I_N}$$

It is also expressed in terms of the signal level $I_s = mI_0$ and the noise level $I_N = I_0 - mI_0$.



3. Modulation Transfer (MT)

This is the attenuation in dB of the modulated signal. It is a function of the modulation index.

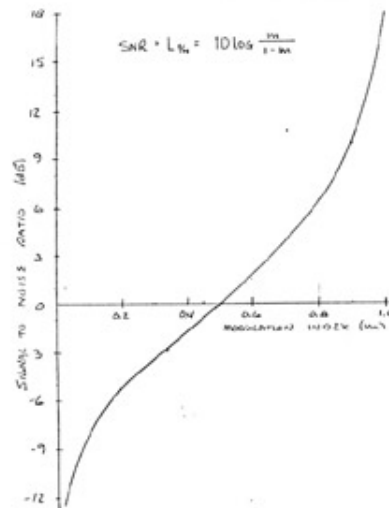
$$MT = 20 \log m$$

4. Signal to Noise Ratio (SNR)

The signal to noise ratio is the level of difference between the signal and the noise ($L_{S/N}$).

$$L_{S/N} = 10 \log \frac{I_s}{I_N} = 10 \log \frac{m}{1-m}$$

It can also be expressed in terms of the modulation index.



5. Transmission Index (TI)

The transmission index is the SNR measured in dB and expressed in percent. To do this the SNR is offset to a practical zero % level and then proportioned to the range of effective SNR. These are subjectively determined constants that relate the perceived threshold of modulation to the maximum value of modulation.

$$TI = \frac{L_{S/N} + 12}{30}$$

The offset is 12 dB and the range is 30 dB.

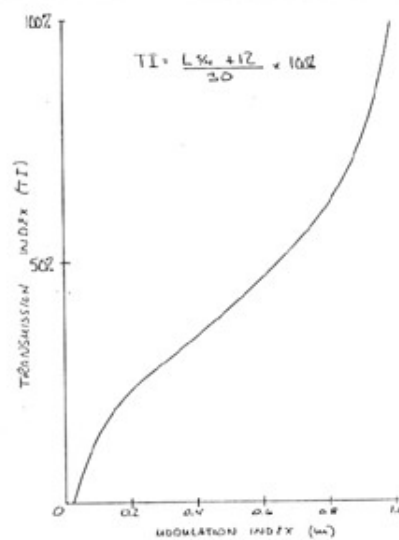
6. Speech Transmission Index (STI)

This is compiled as the sum of the weighted TI for each of the 7 octave bands and expressed in percent.

$$STI = \sum_k (W_k \times TI_k) \times 100\%$$

The weighting factors (W_k) normalize to 1.

$$\sum W_k = 1.0$$



7. Octave Masking Effect (m_o)

This occurs when the lower octave is louder than the measured one. 0.3% of the lower octave intensity is considered noise acting on the test signal.

$$m_o = \frac{I_K}{I_K + 3 \frac{I_{K-1}}{1000}}$$

The impact of simultaneous independent masking effects is carried by multiplying their independent modulation indices together.

$$m = m_1 \times m_2$$

B. MTF IN PRESENTLY MEASURED TERMS

1. Signal Modulation Level (L_a)

Separated signal and noise levels are not directly measured in the present test method. The articulation response curve is the timewise evolution of the received sound levels. This easily allows measurement of L_a , the fluctuation in dB of the test signal.

2. Modulation Index $m(L_a)$

The modulation level (L_a) can be expressed in terms of modulation index by rewriting its definition.

$$L_a = 10 \log \frac{I_2}{I_1} = 10 \log \frac{I + mI}{I - mI} = 10 \log \frac{1 + m}{1 - m}$$

Upon rearrangement, the modulation index is resolved solely in terms of measured level fluctuation (L_a).

$$m(L_a) = \frac{10^{L_a/10} - 1}{10^{L_a/10} + 1}$$

3. Modulation Transfer (MT)

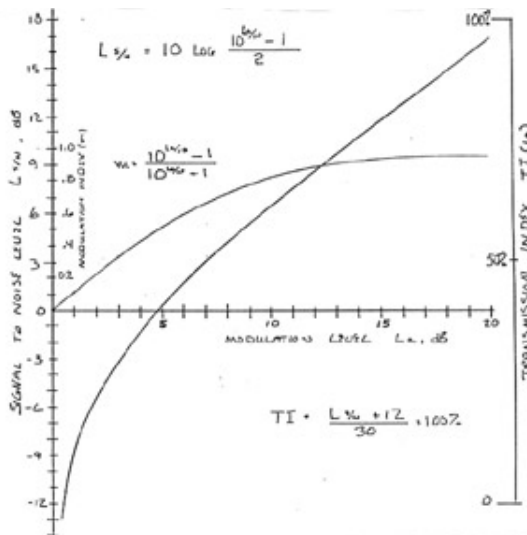
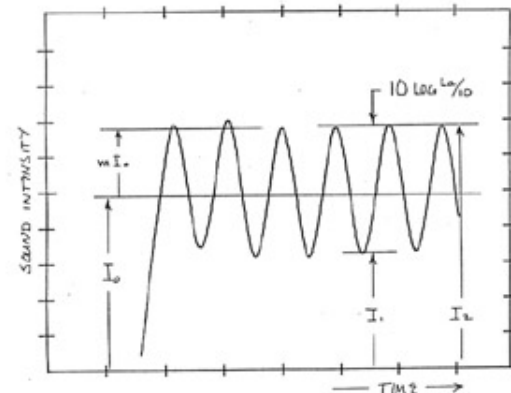
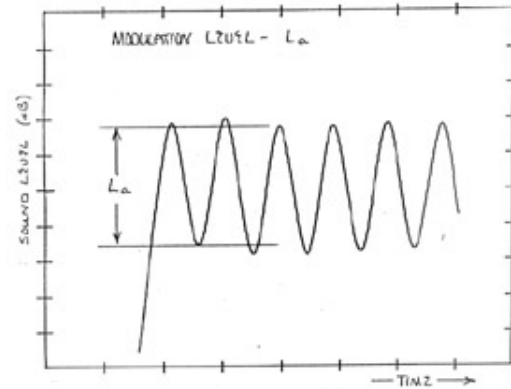
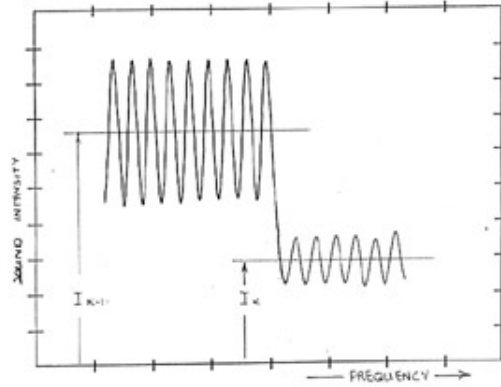
The reduction in modulation can be related to the modulation level at the receiver L_a .

$$MT = 20 \log \frac{10^{L_a/10} - 1}{10^{L_a/10} + 1}$$

4. Signal to Noise Ratio (SNR)

The signal to noise ratio is developed by using the new expression of the modulation index.

$$L_{S/N} = \frac{10 \log \frac{10^{L_a/10} - 1}{2}}{2}$$



5. Transmission Index (TI)

The transmission index remains except as the SNR term is above has been modified.

$$TI = \frac{10 \log \frac{10^{L_a/10} - 1}{2} + 12}{30}$$

6. Mean Transmission Index (TI)

The concept is to sum the various TI values similar to that as done with the STI. Data collected here is not from octave bands but from small bandwidths of tones having similar modulation levels.

The STI octave band weighting factor (W_k) here is undefined. It will be carried in the form of (W_i) to suggest that a listener based preference fit option still remains open.

The octave bandwidth weighting actor in STI appears here as a "log frequency" term in the averaged 5.

$$TI = \frac{\sum_{i=1}^N W_i x TI_i d\text{Log}f_i}{\sum_{i=1}^N d\text{Log}f_i} \times 100\%$$

$$\frac{\sum_{i=1}^N W_i 10 \log \frac{10^{L_i/10} - 1}{2} + 12 (\text{Log}f_i - \text{Log}f_{i-1})}{\sum_{i=1}^N (\text{Log}f_i - \text{Log}f_{i-1})} \times 100\%$$

7. Octave Masking Index (m_o)

This effect is left out of the current presentation. However, it should be thoroughly investigated and ultimately included. It clearly is an operative with small room acoustics. It is easy to find bandwidths with low level articulation and low mean sound level that are just upfrequency from a loud and strongly fluctuating signal.

A given mean intensity level is given by the mean sound level (L)

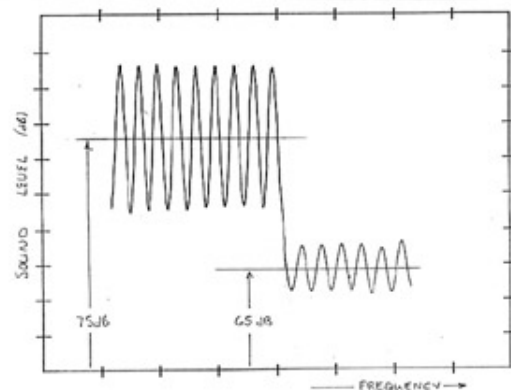
$$I = 10^{L/10}$$

Assume, for example, equal octave fraction band widths for a low frequency 75 dB level followed by a weaker 65 dB level.

$$m_o = \frac{10^{6.5/10}}{10^{6.5/10} + 3 \frac{10^{7.5/10}}{1000}} = 0.97$$

This single level shift is of small consequence but cumulative effects can occur due to a very rough response curve loaded with room resonances. Only 4 such 10 dB shifts would produce a 90% masking index.

$$m_o = m_{o1} \times m_{o2} \times m_{o3} \times m_{o4}$$

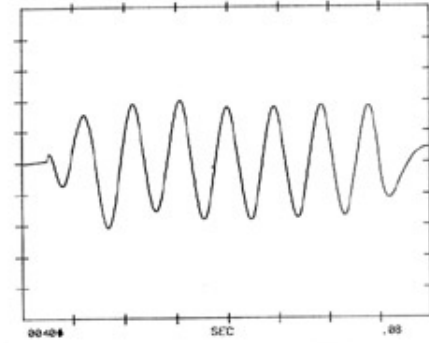


III APPLICATIONS

A. THE TEST SIGNAL

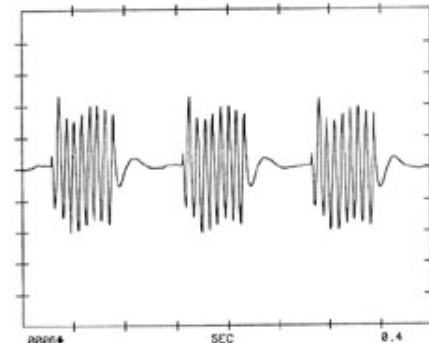
1. The Burst

The MTF (Modulation Transfer Function) method of testing for articulation uses a gated audio signal. For musical playback in listening rooms, a pure tone is gated 8 times per second. Shown in the figure is one burst, it lasts about 60 ms. The sweeping frequency changes about 1 Hz during each burst. This particular burst started at 183 Hz and over 60 ms has shifted to 184 Hz.



2. Duty Cycle

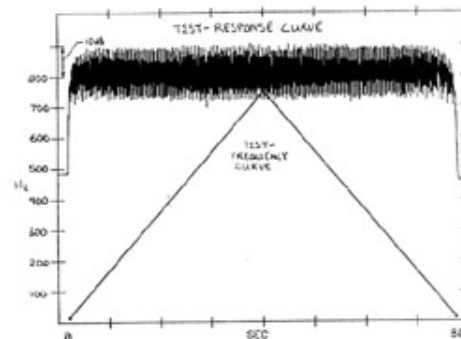
Each tone burst is separated by a dwell time. We use here a 50% duty cycle: 60 ms on and 60 ms off. During the silent period the signal generator continues to change frequency. The next burst after the 184 Hz signal would start at about 185 Hz and slide upwards to 186. Here we show three distinct tone bursts spread out over a 4/10 second time window. These bursts are clearly 60 ms long followed by a 60 ms dwell.



The burst has a square wave modulation. Typical MTF bursts are sine wave modulated, either amplitude or level. Here the square wave modulation has ringing in it, visible in both the on and the off parts of the duty cycle. A ramped attack and decay would reduce the ringing effect. Although the pure tone quality of each burst is degraded by the low level ringing, this coloration provided unique cues for the subjective perception of attack transients. At about 2 dB articulation level, the LF ringing loses audibility—this may suggest a method to evaluate perception thresholds of tonal transients.

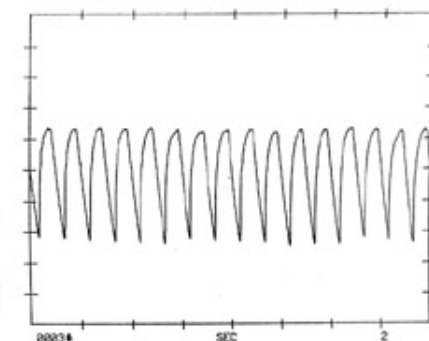
3. Frequency Range

The tone generator is set to be a linear ramp. The signal starts at about 20 Hz and sweeps with a rate of 20 Hz/sec up to 800 Hz and back down to 20 Hz. This symmetrical format has proven easy to read. The ramp up frequencies are not identical to those on the ramp down. This method also serves as a check on the repeatability and accuracy of the test.



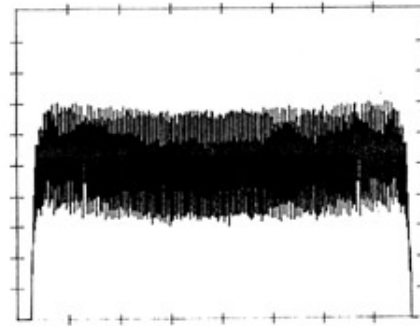
4. Signal Intensity

Next is shown the test signal as seen by a dB meter. If each burst is clean and each dwell period quiet, the dB meter output will alternate between loud and quiet levels. The signal rises in the presence of a burst and falls during the dwell time. There are 2 seconds shown and the 16 sound burst level peaks due to the 8 bursts/second test rate. The actual electric signal level shifts some 50 dB. The damping factors in the analyzer circuitry limit the level swing to only 17 dB. However, this seems to be more range than adequate for the analysis of most rooms.



5. The Complete Test

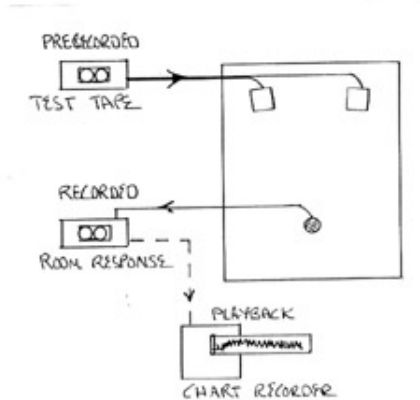
The entire test lasts about 75 seconds. The frequency from 20 Hz through 800 Hz and back down to 20 Hz again. The full test is shown. The level swing of each successive tone burst is clearly visible in this display. The burst's tone raises steadily to the 800 Hz peak frequency and then drops back down during the second half of the test. By listening to this signal on headphones, an articulate audition of the test tone is available.



B. THE RECEIVED SIGNAL

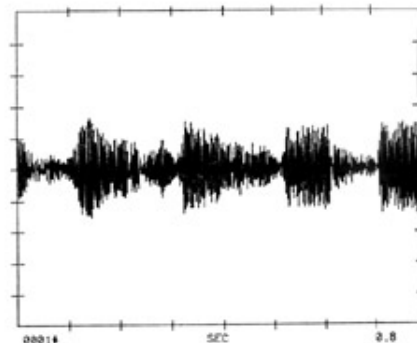
1. The Test Setup

The gated set of tone bursts is played into the room. This allows the distinct features of playback articulation to be observed. A good way to record the effect is with an omni mic and tape recorder without AGC (automatic gain control). Once the listener's signal is captured on tape, it can be played back through an analyzer circuit at a later date.



2. Articulation Response Printouts

The articulation response curve is developed by plotting the recorded sound level vs. time. This is most directly accomplished by running the signal into a chart level recorder. Another method uses the dB level output from a meter to feed the vertical sweep of a storage scope set at very slow horizontal sweep and a printout on an x-y plotter.



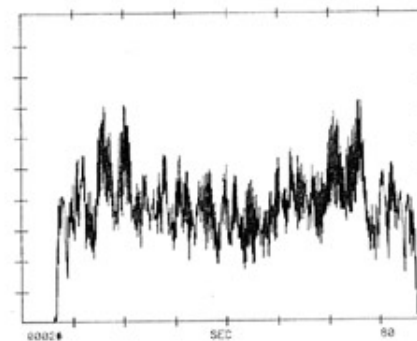
3. Burst Sequence

A closeup of consecutive tone bursts shows substantial acoustic energy can occupy the dwell period. There are 4 bursts in this 8/10 second display. Notice how the burst is deformed. What used to be a sharp attack, flat sustain and abrupt decay has been turned into a pulse that has lost distinctive features.

Ramps, both up and down take the place of the sharp attack and decay of the articulate signal. The sustain does not hold flat, it is foreshortened by the ramping transitions. In this inarticulate space, the room mumbles, slurs and often will "double-tongue" the rapidly gated signal.

4. Articulation Response Curve

Here is what a typical hi-fi demo room does to the fully articulate signal. The signal received by the listener will display some ranges of articulation but most of the test data looks very thin. When the vertical strokes are short, the articulation is weak. There will be little sound level difference between successive tone bursts and dwells. The only way to improve articulation is to "clear the air" between bursts by adding acoustic control.



IV ANALOG TRANSMISSION INDEX

A. APPROXIMATION TO TI

1. Fitted Curve

The STI or as generalized here the \bar{TI} is an equation based on clear definitions. The weighting factor feature (W_f) can be set and prorated to bandwidths used to convert the \bar{TI} into the STI. However the data taken must be converted into a computer and processed to calculate the STI. An analog electronic circuit approximation to this equation is desired.

The key is the \bar{TI} term. Within the range of desired values a simple expression has been found to closely match within a few percent. Also note the expression is in terms of L_a , the presently measured modulation variable.

$$\bar{TI} = 10 \text{Log} \frac{10^{L_a/10} - 1 + 12}{2} \quad \text{(exact)}$$

$$TI^1 \approx \frac{1}{3.2} \text{Log} L_a + \frac{1}{41} L_a + 0.08 \quad \text{(fitted)}$$

(Ref, $L_a = 1 \text{ dB}$)

2. Circuit Diagram for Measurement

The circuit diagram for the analog approximation equation is shown. The first stage develops the level of modulation (L_a). The second stage develops the dB level of the modulation ($\text{Log} L_a$). These two frequency dependant parts are properly ratioed and added to a DC offset then integrated against log frequency. Regardless of the reference level of either term, the DC offset can be scaled to fit.

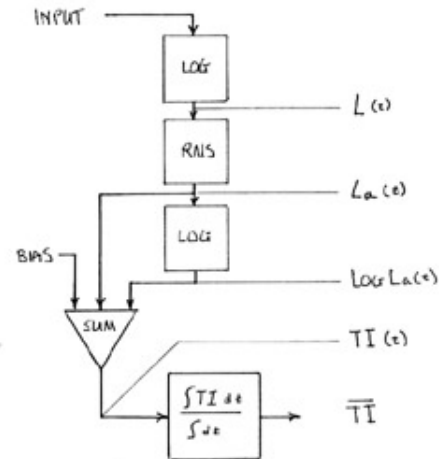
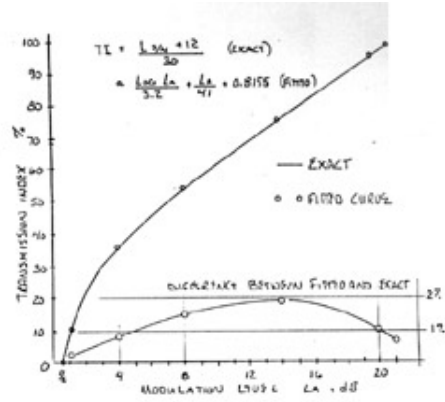
If the frequency sweep is a log sweep instead of linear, then log frequency weighting will be maintained by integrating over time. Substantial signal conditioning has been left out of this circuit to retain a sense of propriety integrity but the basic elements are presented.

B. DISCUSSION OF L_a AND $\text{Log} L_a$

1. Modulation Level (L_a , dB)

Articulation is measured here in terms of the modulation level in acoustic dB's. The weighting scale dB,A or dB,C doesn't affect articulation. Articulation is merely a difference in sound levels.

It is semantically possible to propose that an effect of negative articulation could exist and not be detected by the present circuit. This occurs whenever sound levels in the dwell period exceed levels, attained during the burst. This seems to be able to happen at a frequency for which sound cancellation occurs. The modulation transfer function is not defined in this situation of negative modulation level.



Negative modulation is physically improbable. It takes time for resonant conditions, strong enough to cancel a direct signal, to be developed inside the room. The direct signal will exceed reverb levels during this initial energy buildup period in the room. During this transition period, the direct signal will be heard. Energy is always split between the burst and dwell periods.

2. Articulation Level (10 Log L_a , dB)

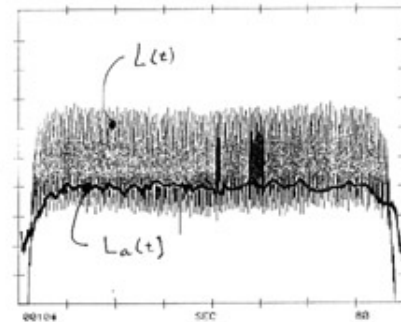
This is also measured in dB and the scale is adjusted so that 1.0 dB articulation is equal to zero articulation level (Ref, 1dB). This is really mathematically arbitrary but set here with considerations. The listener's minimum perceived level change is 0.4 to 0.5 dB for any tone. For the practical purpose of signal burst reproduction 1 dB level differences though audible have little to no perceived value for depicting quality music transition detail. Therefore, it was chosen as zero dB. Regardless, this is an empirical curve fitting arrangement and a different reference here would be reflected in a different DC offset constant than 0.08 above.

C. L , L_a AND Log L_a OF TEST SIGNAL

1. Constant Modulation Test Signal

The test signal has a constant signal to noise ratio of at least 45 dB or the full dynamic range of the test cassette tape. The corresponding articulation level shows as the solid, slightly fluctuating dark line. It is overlaid against the back drop of its gated sweep response curve (L_a).

Two curves are shown here. The sound $L(t)$ level vs. time articulation response curve is the wide fluctuating line. Overlaid on it is a solid, slowly changing and relatively flat line, the Modulation Level, $L_a(t)$.



2. Upper Limits to Sound Level

The sound level curve is not fully accurate because of the ballistics in the electronic detection circuits. For this data run the upper limit is about 20 dB. The real 40 dB signal modulation does not show. This is of no practical concern because 15 dB to 20 dB differences between peaks and valleys in the modulation envelope are subjectively quite adequate. Most of the data is often on the order of a 5 dB to 10 dB articulation level (L_a).

3. Calibration

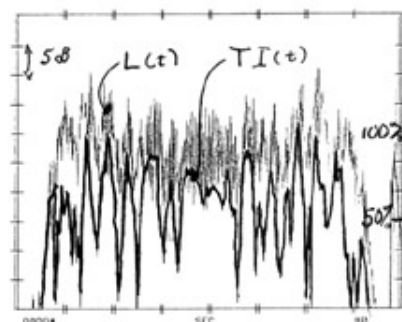
In future work a 1K test tone should be modulated at zero, 1, 5, 10, 15 and 20 dB modulation levels. This will allow calibration of testing circuits. An alternative to this is to step the 1K tone level (zero modulation) to develop calibration at the above -1, -5, -10, -15, -20 levels.

D. L , L_a AND Log L_a OF RECEIVED SIGNAL

1. Modulated Sweep Response Curve

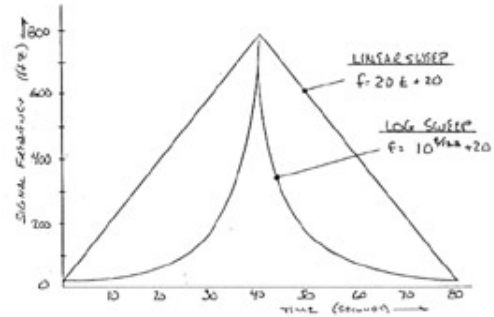
This shows the signal to noise ratio spectral response of the room to the rapidly gated tone sweep. The actual received signal level $L(t)$ is the wide, rapid fluctuating line.

The overlaid solid line is the transmission index vs. frequency at the 8 Hz gated modulation rate. The mean TI would be the averaged value of this curve.



2. This curve is a linear frequency sweep and the mean TI requires log frequency weighting. If a log frequency sweep was used instead of linear, then straight integration of the TI in time would produce the mean TI.

Linear sweep is often used in low frequency room measurements. It is said the ear hears quasi-linear frequency scale below 200 Hz. The log sweep spends $\frac{3}{4}$ of the time below 170 Hz about $\frac{1}{4}$ of the frequency range to be explored. The remainder $\frac{1}{4}$ test time packs the remaining $\frac{3}{4}$ frequency range (200 to 800 Hz). Although log frequency sweep accommodates a simple integration scheme for the mean TI, it most likely is not sampling sufficiently the room articulation. A more sophisticated integration must be used.



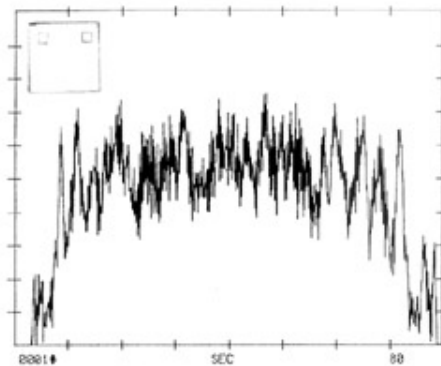
V SAMPLE TESTS

A. ROOM SEQUENCE

A listening room, 8' x 14' x 18' with double sheetrocked walls and concrete floor is tested at various stages of acoustic treatment. Fundamental, is the use of corner-loaded bass traps. The mic is placed at the hi-fi listener's position and two speakers, in phase are located at the opposite end of the room in a stereo setup.

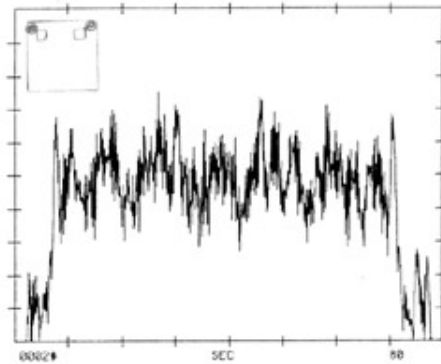
1. Bare Room Response

To read this type of printout, we focus on the percentage of the test frequency sweep that has a wide (10 dB) swing, peak to peak. Marginally acceptable is a medium swing (5 dB). Real garbling occurs with less than a 2 dB swing. Note also the irregular "median line." It is the average about which jitters the articulation signal. The terrain looks like a lot of steep hills and valleys covered with very little articulation.

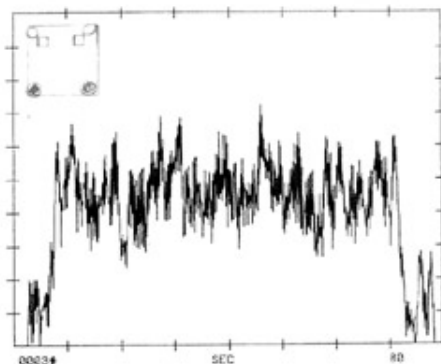


2. Absorption Added in Stages

a. Here, a simple Tx6 set has been added to the front of the room behind the speakers. Already a substantial pattern of low level articulation is established throughout the entire test. The hills and valleys have grown less severe and are covered better with a wider articulation band. Note also the overall flatness, the room is being acoustically EQ'd.



b. The next setup adds traps (16x3 plus 11x3 pair stacks) at the back of the room. Again, the frequency bands of improved articulation widen. The severity of the peaks and valleys is more reduced. A few peak/valley patterns have even disappeared.



The softening of the peak/valley profiles means the "Q" of the room, the sharpness of its resonance responses, have been lowered. As the room resonances are damped, the peaks drop, the valleys rise and there is an overall softening effect to the room response curve.

c. Next is added a side wall treatment, 4 sets of 9" x 5' ½ Rounds. This controls stage width and imaging, lateral flutter and cross talk. It develops overall a much deeper articulation. It produces wide bands of continuously full articulation, some 200 Hz wide between 400 and 600 Hz. Yet, curiously there seems to be some areas of thinning, reduced articulation around 150 Hz.

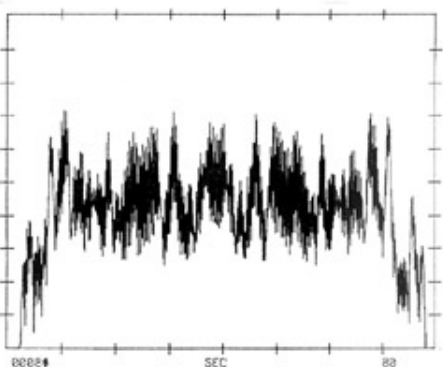
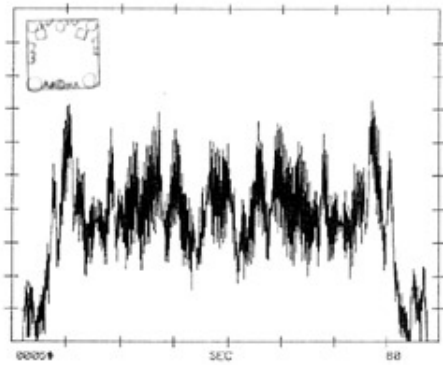
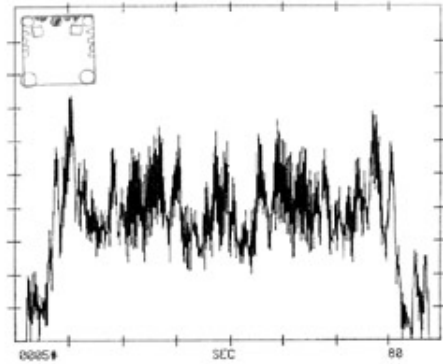
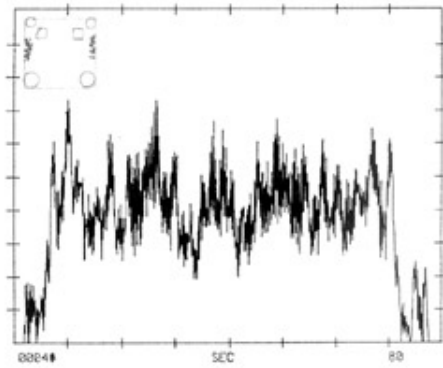
d. The head wall traps are the next to be set, 6-11x5 ½ Rounds plus a single column of 11x6 Full rounds in the center. This develops stage depth, clarity and imaging detail. Dramatic articulation improvement is seen broadband, the peak/valley terrain flattens substantially. The width of the articulation patterns have grown quite wide and improvement is seen in the mid-bass. The front/rear energy storage system of the room has been dampened to make this marked improvement.

e. Finally we have added the rear wall. A 16x3 + 11x5 center column and 4 sets of 11x5 ½ Rounds with one more pair on the front wall. The result is a very wide and steady articulation pattern that extends even into the deepest bass. Peaks and valleys now even more are soft, rounded. The room still retains a strong, comfortable ambience.

If you compare the overall before and after room articulation signatures, you will see that the sound levels below 100 Hz have not changed and those above 100 Hz are depressed by about 5 dB. In addition, we see that below 400 cycles the articulation signature increases from 2 to 8 dB and above 400 from 10 to 18 dB.

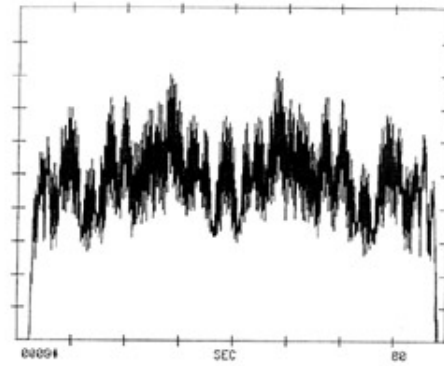
3. Equalizer Added

The effects of equalizing the signal were explored. The effort was made to get the trapped, articulate room to have an over flutter response. A 1/3 octave equalizer was set with pink noise and headphones. The following articulation test results. For better results, a parametric equalizer could be used. With this equalizer a noticeable ringing effect occurs, most likely not desirable in quality audio. Nonetheless, the response curve has been flattened, peaks lowered, valleys raised. Notice however, that there is "zero effect" on articulation. Electronic EQ only adjusts levels, not articulation.

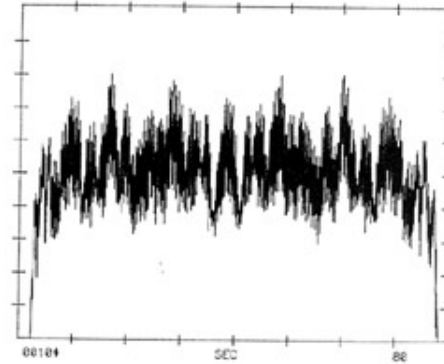


4. Full Acoustics Plus Equalizer

a. The “full on” room has also been tested. This is not too unlike the typical dedicated Hi-end reference listening room. Basically, a carpet has been added along with floor bounce traps. All the traps of the prior setup (#6) have been elongated from their 5-foot height to a full floor to ceiling length. A major articulation improvement is noted, especially in the 20 to 400 Hz range. The natural acoustic #Q is taking a strong control, the low-end boom below 100 is almost gone.



b. Finally, to this “ultra” system, we degrade its sonics but add equalizer effects. Again the EQ is set with pink noise, RTA and 1/3 octave equalizer. The result is pretty flat, and articulate response. There are a few small band widths with poor articulation remaining. Even these may well be cleaned up with additional tweaking. Again the ringing effect of the equalizer is clearly audible in this test, something undesired in precision audition.



B. 1/3 OCTAVE PINK NOISE, RTA

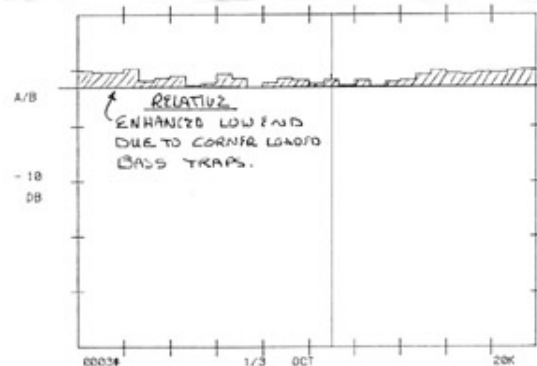
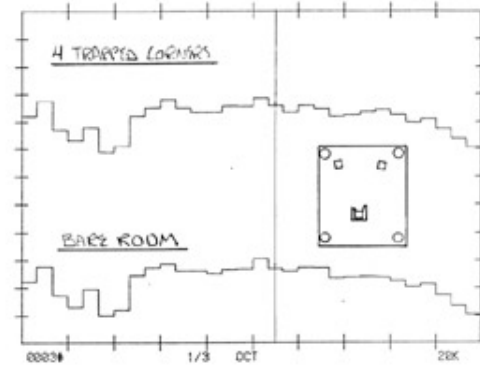
1. RTA and Room Treatment Sequence

For the entire series of test just described, 1/3 octave RTA was also taken. Above 40 Hz the overall levels are reduced by 2 dB. If we overlay and line up the mid-range levels, we see a relative increase in the lower octaves below 70 Hz by 2 dB. This is the acoustic EQ effect. This acoustic treatment brought the deep bass 2 dB closer to the mid bass levels.

Relatively minor corrections towards flattening the spectrum sound levels with no loss of deep bass sound power is how RTA sees the effects of the full on acoustics. Clearly RTA doesn't begin to suggest the fast tracking ability of the listening room.

2. RTA and Slow Sine Sweep

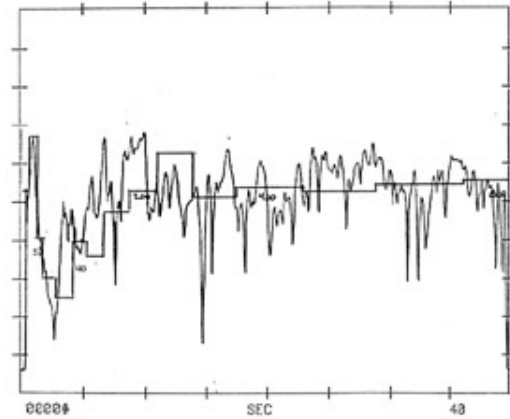
The narrow band frequency sweep room response curve is compared to the 1/3 octave pink noise levels. The frequency range is 20 to 800 Hz. The frequency scale is linear, this stretches the 1/3 octave bandwidths as the frequency goes higher.



The RTA levels are weighted higher with increasing frequency. This is due to wider bandwidths, more 1 Hz levels being added together. The equivalent narrow band spectrum can be had by subtracting the bandwidth weighting term from each bandwidth level.

$$L = 10\log f + 10\log 23\%$$

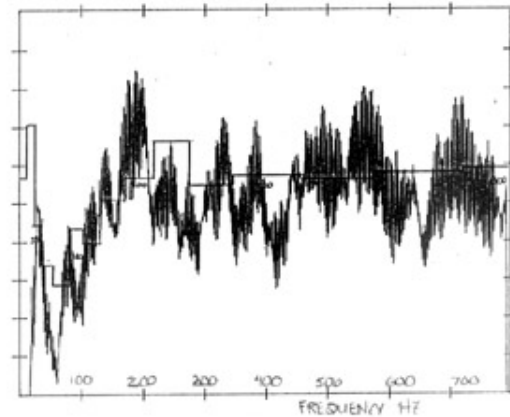
The 1/3 octave has 23% bandwidth. When the two curves are overlaid the general tendency is seen but the detailed narrow band sweep cannot be even inferred by the 1/3 octave measurement.



3. RTA and Articulated Sweep

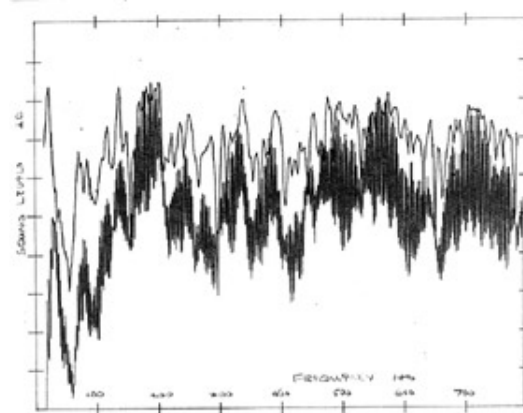
Not unlike the vague relationship between the frequency response of the room and the RTA, so it is with the articulated sweep. Overall trends do track, but the RTA gives no indications by which features in the articulated sweep response curve can be derived.

For example, 1/3 octave EQ suggests that the 250 Hz band should be cut some 5 dB. However, the articulated sweep response shows that the problem high sound level is a 1/3 octave band centered at 180 Hz.



C. SLOW SINE AND MODULATED SWEEPS

Here we compare the slow sine sweep to the modulated sweep. The sound levels at the listener's position are recorded in both cases between 20 and 800 Hz.

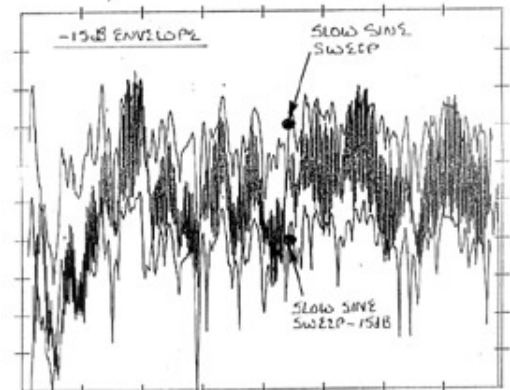


1. Observations and Tendencies

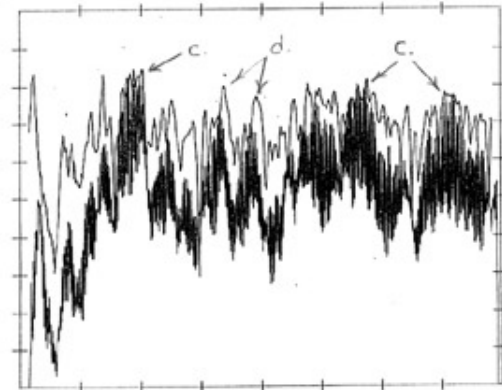
Tendencies are noticed. The trend of the slow sine sweep matches the trend of the articulated sweep.

a) Articulation levels L_a of 12 to 15 dB attain peak sound levels equal to that of the slow sine sweep levels.

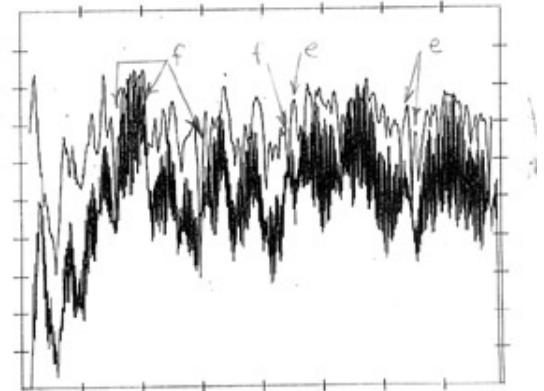
b) Articulation levels that are less than 12 dB fall short of the slow sine sweep level by an amount approximately equal to: $15 - L_a$.



c) Strong articulation is associated with wide bandwidths of relatively uniform sound level on the slow sine sweep response curve.



d) The lower the “Q” of sine sweep response curve the stronger the articulation signal.



e) Very low articulation levels are always accompanied by a very sharp, high “Q” room resonance section of the room response curve.

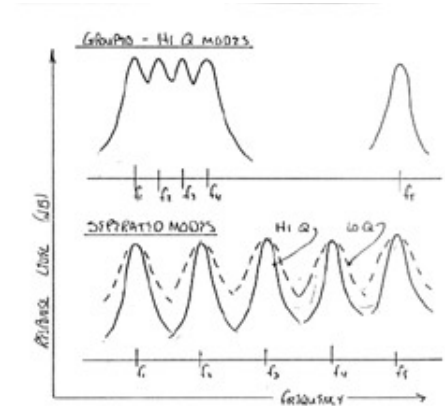
f) Rapid sound level changes in the slow sine sweep curve mark frequency bands with poor articulation response.

C. ROOM MODES AND “Q”

From the above it is clear that room mode spacing and the adjustment of room resonance “Q” are controlling variables in the development of articulation response in small rooms.

1. Mode Spacing

To illustrate by contrast, it can be safely concluded that a group of closely spaced high “Q” resonances will produce stronger articulation than if that given group was well separated having well isolated resonance peaks. The tight grouping of some modes leave more spaces between other modes. The real answer to an articulate room will be to have a set properly spaced and damped room resonances.

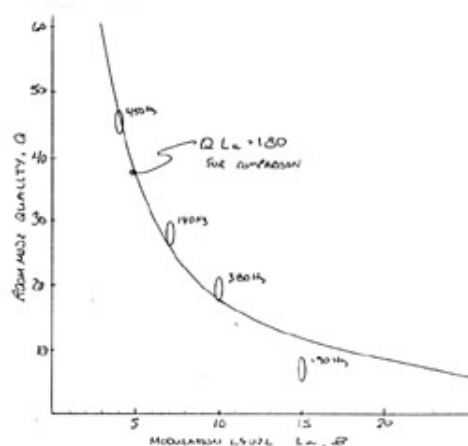


2. Modulation Level L_a and Room “Q”

It is straight forward to expect that the higher the “Q” for a particular room resonance, the lower the articulation levels would be. For the data presented above in Section V-B, an interesting curve “Q” vs. L_a is produced. The room “Q” has an almost exact inverse relationship with the articulation level L_a . The empirical data is found to lie on the curve of:

$$Q \times L_a = 180$$

Since the minimum L_a for acceptable listening is about 5 dB, the most probable maximum acceptable “Q” will be about 36. For the very desirable L_a of 10 dB we have room resonance “Q” of 18. The “Q” of a typical room is often 40 to 50 prior to specific acoustic conditioning.



D. LINEAR "Q" VS L_a

The classic sabine equation uses diffuse exponent sound fields. The "Q" vs. L_a relationship can be predicted, it is seen to not fit the measured relationship. This is expected because the sound field in small rooms and lower octaves does not exponentially decay.

1. "Q" and RT_{60}

There exists a group of "Q" relationships dependant on a variety of variables. The RT_{60} is no exception.

$$Q = \frac{1}{22} RT_{60} \times f$$

The frequency of the resonance (f) part of the dependant variables.

2. L_a and RT_{60}

For linear RT_{60} , the diffuse exponential decay sound field level drops proportional to time.

$$L_a = \frac{60}{RT_{60}} \times t$$

The gated tone burst has burst rate (F) and its dwell period is the time allowed for sound level decay.

$$L_a = \frac{60}{RT_{60}} \times \frac{1}{2F} = \frac{30}{RT_{60} \times F}$$

An 8 Hz gated frequency yields an equation relevant to the present test.

$$L_a = \frac{3.75}{RT_{60}} \quad (F = 8 \text{ Hz})$$

3. Linear "Q" and L_a

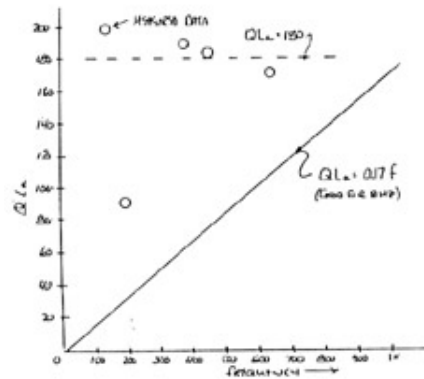
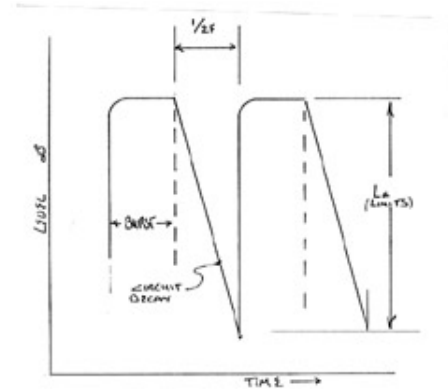
By combining the above equations the frequency dependant "Q" x L_a relation is developed.

$$Q \times L_a = \frac{1.36 f}{F} = 0.17 f (8 \text{ Hz})$$

For linear decay the Cis directly proportional to frequency. This is not what is measured, a constant. Since both definitions used, "Q" and L_a assume a linear acoustic relationship with RT_{60} , neither can be identified as the non linear term at this point.

DISCUSSION

The goal of this project has been to explore the Transmission Index of small rooms in the lower octaves. The rapidly gated slow sine sweep is an effective test signal. Although envelope shaping of the attack and decay should be explored, the existing coloration led to the observation that low level coloration becomes inaudible at a higher modulation level than does the modulation itself. This suggests that "quality" detection thresholds may well be much different from "quantity" detection. Research in perception along the lines of complex signal detection thresholds needs to be applied to the present work.



The difference between the linear and measured QLa term stands to illustrate that the prediction of TI in the lower octaves in small rooms has yet to be accomplished. More empirical work also needs to be done in this area. The observation presented here is only based on one data run.

A new, complex test signal and detection method may be considered to directly measure the masked partial signal level. A correlation between pure tone modulation levels at the partially frequency and the masking level of the partial wherein a complex tone burst ie. Linear, additive effects, may be fruitful.

The TI equation has been approximated here by a fitted curve using the same single variable. The only reason for this is to access the convenience of a relatively simple analog circuit.

Further work with the exact equation ought to be completed using analog or computational methods to develop the TI. There also may be additional terms added to reduce the error of the approximation curve.

There lies ahead a great opportunity to work on the theoretical side of the Transmission Index at lower frequencies in small rooms. The first step aside from large halls in linear acoustics was the ray tracing method, but this is not applicable to small room resonant modes.

The relative level effect needs to be factored into the present TI approach. A room with strong level changes in a slow sine sweep must be penalized when compared to a room with a relatively flat response. A method to isolate this effect needs to be developed and produce an independent modulation index.

In general, standards for speech in small rooms need to be applied to this work. The performance of STI analyzers needs to be compared to traditional listening tests in small classrooms where modes exist in the speech range. In large halls, little emphasis is given to the lower speech octave, 125 Hz. Small rooms, with their room modes and typical lack of low frequency absorption, may well require re-assessment of this weighting.

CONCLUSION

A method that develops spectral response curves for articulation has been demonstrated. The measure variables have been written into the equations that define the Modulation Transfer Function and the corresponding Transmission Index. The signal to reverberant noise level is directly measured and there is no conversion of data that requires the assumption of linear acoustics.

The equipment used to make this test is relatively common. The source is a pre-recorded cassette test signal. Analysis will use as little as a sound meter and strip chart recorder. By adding a circuit for signal processing, the Transmission Index response curve can be developed. With additional circuits even the STI can be stated.

The STI is fast becoming a standard specification. Engineers and consultants require a spectral version of the Transmission Index in order to remedy the acoustics. Now that this simple and low cost articulation test method has been shown to produce detailed spectral information, it is hoped that this technique will be the forerunner of a new class of sound system analysis.

BIBLIOGRAPHY

(The following were used in the preparation of the paper)

Houtgast, T. and Steeneken, H.J.M., Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function Parts I, and II. ACUSTICA VOL 46, 1980.

Houtgast, T. and Steeneken, H.J.M., A Review of the MTF Concept in Room Acoustics and Its Use for Estimating Speech Intelligibility in Auditoria. JASA 77 (3) March 1985.

Schroeder, M.R., Modulation Transfer Functions: Definition and Measurement. ACUSTICA Vol 49 (1981) 179.

Houtgast, T. and Steeneken, H.J.M., A Physical Method for Measuring Speech-transmission Quality. JASA 67 (1) Jan 1980.

Houtgast, T. and Steeneken, H.J.M., The Modulation Transfer Function in Room Acoustics as a Prediction of Speech Intelligibility. ACUSTICA Vol 28, 1973.

Kryter, Karl., Methods for the Calculation and Use of the Articulation Index. JASA 34 (II) Nov 1962.

Polack, J.D., Alrutz, H. and Schroeder, M.R., The Modulation Transfer Function of Music Signals and its Application to Reverberation Measurements. ACUSTICA 54 (1984).

Demany, L. and Semal, C., Amplitude and Frequency Modulation. ACUSTICA 61 (1986).

Fastl, H. and Hesse, A., Frequency Discrimination for Pure Tones at Short Durations. ACUSTICA 56 (1984).