Estimation of Loudness by Zwicker's Method

Loudness is one category in the list of human perceptions of sound. There are many methods of estimating Loudness using objective measurements. No method is perfect. Simple methods seem to work reasonably well for simple sounds, but more complicated methods are needed when more complex sounds are included. Zwicker's method makes use of several psychoacoustical principles in a calculation process to give an estimate of the "average person's" impression of the Loudness of sound. Those principles are listed below and briefly discussed on the following pages along with an example of Zwicker's method.

For Steady-State Sounds

1. Non-linear Frequency Sensitivity e.g. Equal Loudness Contours

2. Masking & Critical Band Theory

3. Effect of Sound Fields on Loudness i.e. Frontal or Diffuse

Process Summary (Standard: ISO 532 method B)
> obtain one-third octave band spectrum
(the standard is very loose in the area of acquiring the spectrum)
> user identifies the kind of sound field
(the standard gives very little guidance concerning sound fields)
> apply principles from 1, 2 and 3 to re-scale / re-plot the spectrum
(i.e. Specific Loudness Pattern i.e. Sone/Bark vs. Critical Bands)
> the area under the Specific Loudness Pattern is proportional to the Overall Loudness
(when the overall loudness is expressed in "Sones" the loudness relationship is linear i.e. 4 sones is twice as loud as 2 sones.)
> for Identification a suffix "GF" or "GD" is added to the unit "Sones"
- "G" (from the German for Critical Band)
  identifies that Zwicker's method was used
- "F" identifies the sound field as Frontal e.g. SoneGF
- "D" identifies the sound field as Diffuse e.g. SoneGD

For Transient Sounds

In addition to 1, 2, and 3...

4. Temporal Effects - Growth of Loudness with time

5. Temporal Masking (vs. simultaneous masking i.e. steady-state)

Process Summary (The Process is not yet standardized)
Commercially available analyzers for transient sounds using Zwicker Loudness implement 4 and 5 using various approaches generally related to detector averaging times of the analyzer depending on the type of transient sound being measured. The uncertainty principle (for valid data the Bandwidth times the Averaging Time must be greater than or equal to one) is an important limitation on all current analyzers. Human hearing appears to be able to outperform the uncertainty principle, perhaps because human hearing is a non-linear process.

Zw_notes10/RJF/MY00 -MR03

1. Non-linear Frequency Sensitivity

Figures 1 and 2 show the Equal Loudness Contours that demonstrate the non-linear frequency sensitivity of human hearing. These contours were obtained by having human subjects compare the loudness of pure tones of various frequencies at various levels. The lowest curve is representative of the threshold of human hearing. At frequencies near 3000 to 4000 Hz the threshold is low and the human ear is the most sensitive. At low frequencies, especially below 100 Hz, the threshold of hearing is significantly higher than at 1000 Hz. Notice that at frequencies around 1000 Hz the contours are spaced at about 10 dB intervals. However, notice that at the low frequencies the spacing of the contours is much smaller than at 1000 Hz. This indicates that, above threshold, for a given increase in sound level the impression of loudness increases faster at low frequencies than at higher frequencies. Zwicker's method accounts for both the changes with frequency in sensitivity at threshold and the rates at which increasing levels contribute to the impression of loudness.

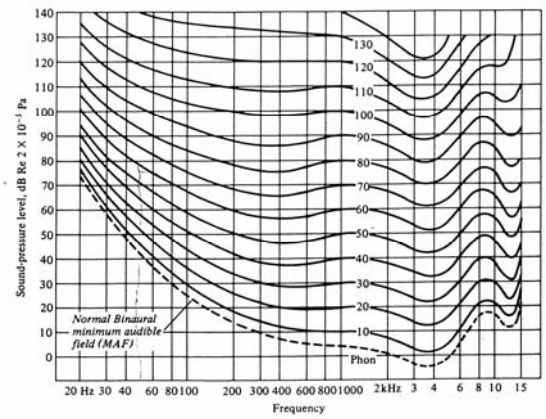

Figure 2.4 Internationally standardized set of equal-loudness-level contours. *(Source: J. T. Broch, Acoustic Noise Measurements, B & K Instruments, Inc., January 1971. With permission.)*

Lord, Gatley & Evensen, Noise Control for Engineers, Krieger Publishing, 1987.

Figure 1

Figure 2 shows an abbreviated version of the equal loudness contours with selected contours labeled both in phons and sones. Phon is a logarithmic unit like decibel while sone is a linear loudness unit. Remember that these contours are for pure tones only. Bands of noise and broadband sounds will be discussed later.

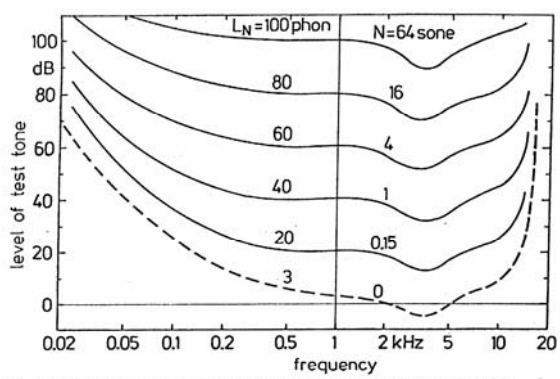

Fig. 8.1. Equal-loudness contours for pure tones in a free sound field. The parameter is expressed in loudness level, L_N, and loudness, N

Zwicker & Fastl, Psychoacoustics Facts and Models, Springer-Verlag, 1990.

Figure 2

2. Masking & Critical Band Theory

Masking is the reduction in the ability to hear one sound due to the presence of another sound. If two sounds are present at the same time, the lower frequency sound tends to mask the higher frequency sound (also known as simultaneous masking). Figure 3 demonstrates how a band of noise can reduce the audibility of a tone. At frequencies below the band of noise very little masking occurs. Above the band of noise a great deal of masking can occur depending on the level of the noise.

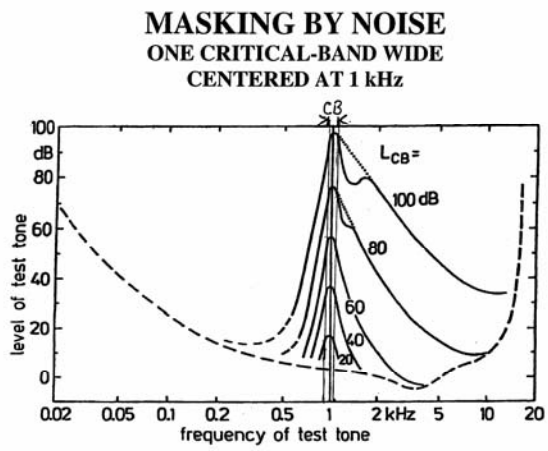

Fig. 4.4. Level of test tone just masked by critical-band wide noise with centre frequency of 1 kHz and different levels as a function of the frequency of the test tone

Example of Masking at
Higher Frequencies by Lower Frequency Noise

Figure 3

The critical band is a feature of human hearing that is important in both masking and the summation of loudness. Figures 4A and 4B demonstrate how human hearing adds the contributions of sound within the critical band differently than it does sound outside the critical band. Starting with a narrow band of noise and keeping the overall sound pressure level constant, the width of the band is gradually increased. As long as the band of noise is not wider than the critical band, the impression of loudness remains the same. Once the band of noise becomes wider than the critical band the impression of loudness begins to increase.

For audible frequencies, the critical band scale range is from 0 to 24. Critical band units are called “Bark” (named in honor of Barkhausen, an early researcher in the characteristics of human hearing).

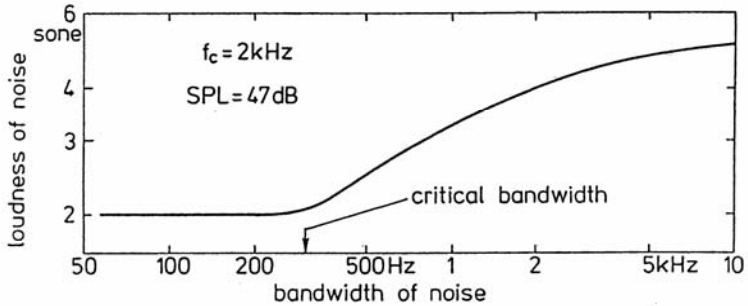

Fig. 6.7. Loudness of band-pass noise centred at 2 kHz with an overall SPL of 47 dB as a function of its bandwidth

Figure 4A

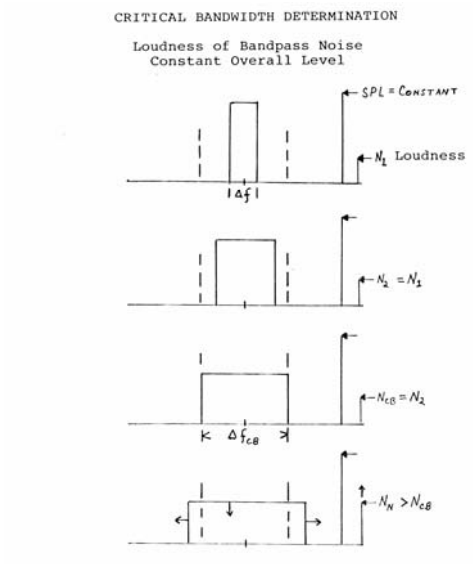

Figure 4B

The width of critical bands varies continuously with frequency in the manner shown in Figure 4C below. This relation was determined by psychoacoustical testing on humans. For practical loudness measurements Zwicker's method uses frequency analysis data obtained in parallel contiguous bands. When creating plots of specific loudness, the x-axis is linear in critical bands (units are Bark). However, as shown in Figure 4C, an axis that is linear in critical bands is neither linear in frequency nor logarithmic in frequency. This is currently a problem with commercial sound quality software packages when they attempt to plot the data in frequency. The creators of the software have used simple linear or logarithmic plotting packages instead of writing code to handle the more complicated psychoacoustical reality of critical bands. By plotting the data by frequency using either a linear or logarithmic scale, the resultant graph is distorted so that Zwicker's premise is no longer true (i.e. the area under the curve is proportional to the overall loudness). For an example of a specific loudness pattern scaled in both critical bands (Bark) and frequency (Hz) see Figure 8 below.

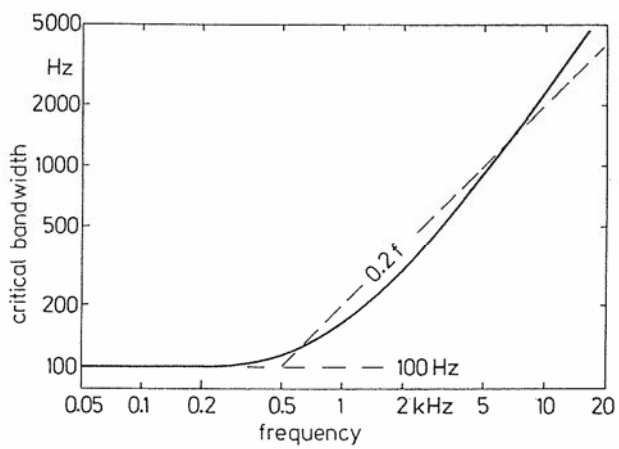

Fig. 6.8. Critical bandwidth as a function of frequency. Approximations for low and high frequency ranges are indicated by broken lines

141

Figure 4C

3. Effect of Sound Fields on Loudness i.e. Frontal or Diffuse

Figure 5 shows a representation of the two types of sound fields that are accounted for in Zwicker's method. In studies of how humans react to different sound fields, it was found that when a human subject is presented with sounds having the same objective level, a diffuse sound field was judged to be louder than a frontal sound field.

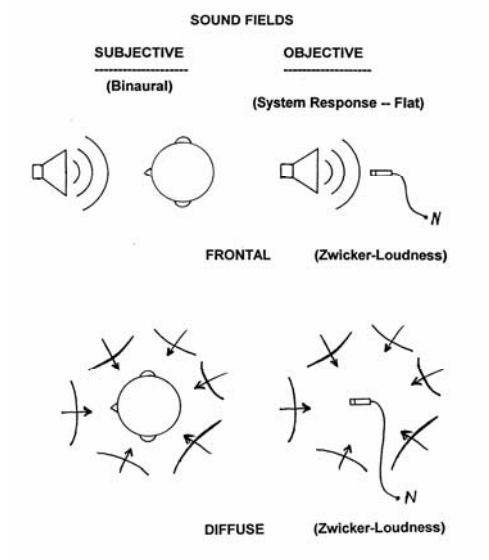

Figure 5

4. Temporal Effects - Growth of Loudness with Time

It takes time for our hearing to respond when a sound first turns on. Figure 6 demonstrates the change in response as the duration of a sound is lengthened. Even though the maximum objective level of the sound is the same (as shown in the top trace), the Loudness perception (shown in the bottom trace) is not the same. A sound must be on for about 100 ms for the perception of its Loudness to reach a maximum.

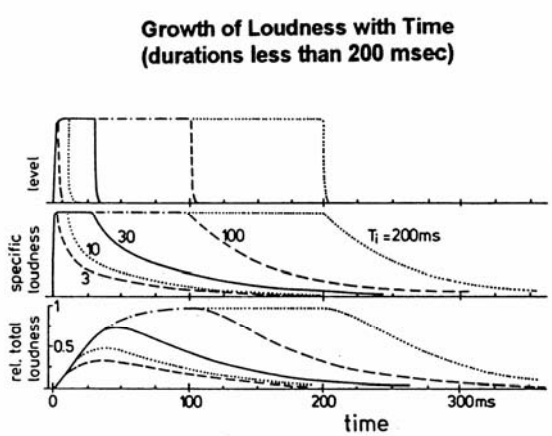

Fig. 8.20. Growth of loudness as a function of time: sound pressure level produced by tone bursts which are passed through a critical-band filter (*upper part*), specific loudness produced by the tone bursts (*middle part*), and relative loudness produced by these tone bursts as a function of time. The parameter is the duration T_i of the tone bursts as indicated

Figure 6

5. Temporal Masking (vs. simultaneous masking i.e. steady-state)

A sound can be masked by another sound even when that other sound is not physically present. Figure 7 illustrates the effects of temporal masking compared to simultaneous masking. In post-masking when the masking sound is turned off the masking effect does not abruptly stop but instead takes time to decay away. Pre-masking is like a reset function which interrupts the perception of a sound that has turned on just before the masking sound has turned on. Notice that its effect is much smaller than that of post-masking. It is not included in the current version of Zwicker's method.

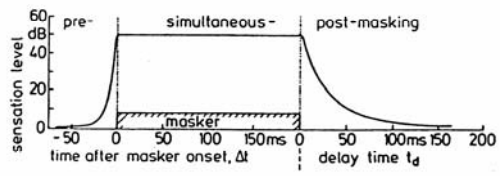

Fig. 4.17. Schematic drawing to illustrate and characterize the regions within which premasking, simultaneous masking and postmasking occur. Note that postmasking uses a different time origin than premasking and simultaneous masking

Figure 7

EXAMPLE

Figure 8 is an example of the input and output of a program that calculates Zwicker Loudness for steady-state sounds. A Specific Loudness pattern is calculated from a one-third octave band spectrum.
Part of the program combines certain one-third octave bands to simulate the first, second and third critical bands as shown in the figure. The overall loudness in sones is proportional to the area under the Specific Loudness pattern. For this Specific Loudness pattern the overall loudness value is 14.7 soneGF, shown in the upper right corner. The effects of masking are represented in the Specific Loudness pattern by the curved slopes on the on the high frequency side of bands that protrude above the rest of the pattern. Note that the threshold of hearing is the horizontal line with the value of 0 sone/bark on the Specific Loudness pattern. Compare this simple representation to the threshold of hearing curve in figure 1. Also note that the equal loudness curves in figure 1 would appear as horizontal lines on a Specific Loudness pattern.

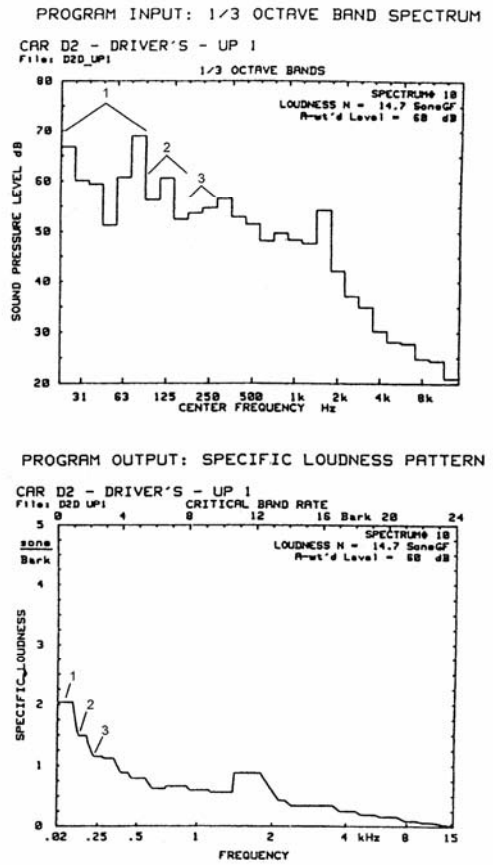

Figure 8

For impulsive sounds a single Specific Loudness pattern will not sufficiently describe the sound. Figure 9 is an example of a Multispectrum or Waterfall plot of an impulsive sound. The interval between successive patterns in this example is 1 ms. Figure 10 is an example of the Overall Loudness Time History that results from the Multispectrum in Figure 9.

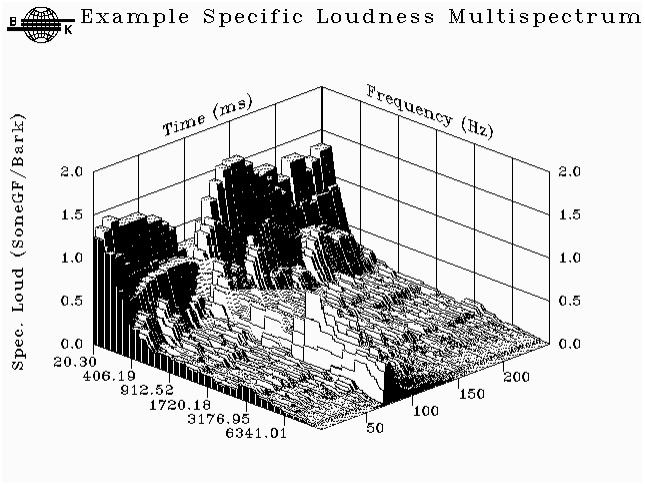

Figure 9

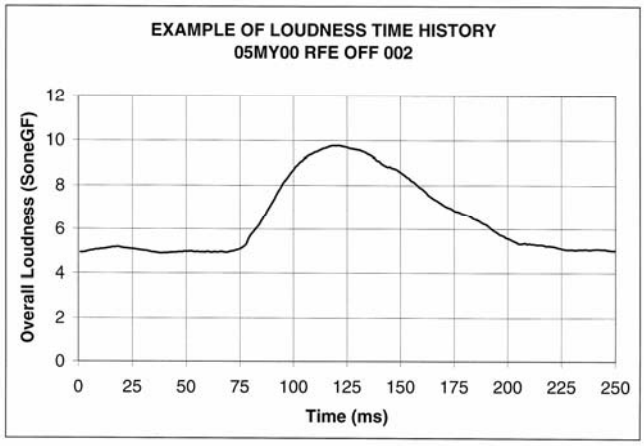

Figure 10

Figure 11 is an example of a Loudness Cumulative Distribution function, a useful way to statistically represent the Loudness Time History. The Loudness Time History shown in Figure 10 is a fairly simple shape but often a Time History can have a more complicated shape with multiple peaks and valleys. Simple shapes can be represented adequately with simple statistics such as the maximum value or the average value. However, no one can guarantee that a Loudness Time History will always have a simple shape, so more detailed statistics, such as the Cumulative Distribution function, are often needed. An important aspect of the Cumulative Distribution function is that it applies to a given time interval that is chosen by the user. In the case shown in Figure 11 the time interval is 250 ms. After selecting the time interval, the function is constructed by choosing a value of Loudness, determining the time that that value is exceeded in the Time History, and then determining, as a percentage, the ratio of the exceedance time to the user selected time interval. For example, for an impulsive sound, a value just below the maximum Loudness is exceeded a very small percent of the time. Similarly, a value just below the background Loudness is exceeded 100% of the time. Often only selected values of the function are quoted such as N5, N10 or N50, the value exceeded 5% or 10% or 50% of the time respectively. While these values may be useful for summary purposes, the shape of the function is more valuable in developing a well-defined specification. In any case, it is essential to state the time interval over which the statistics are calculated.

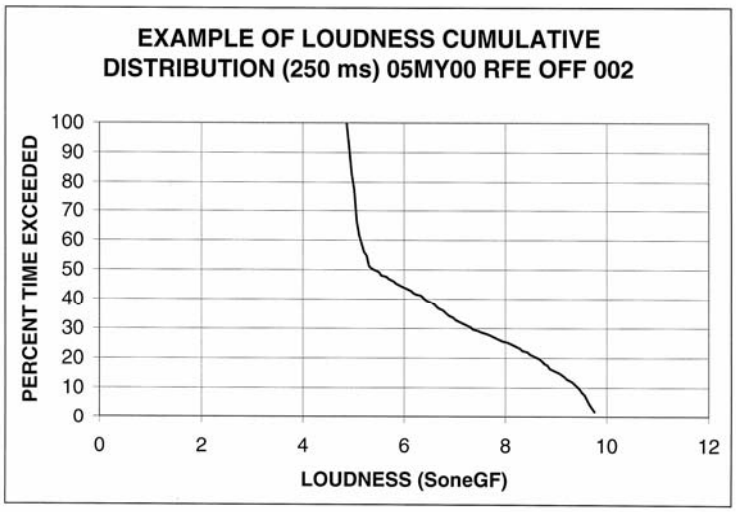

Figure 11

Illustration Credits:

Figure 1 - Lord, Gatley,& Evernsen, Noise Control for Engineers, Krieger Publishing, 1987.
Figure 2 - Zwicker & Fastl, Psychoacoustics Facts and Models, Springer-Verlag, 1990.
Figure 3 - Zwicker & Fastl, Psychoacoustics Facts and Models, Springer-Verlag, 1990.
Figure 4A - Zwicker & Fastl, Psychoacoustics Facts and Models, Springer-Verlag, 1990.
Figure 4C - Zwicker & Fastl, Psychoacoustics Facts and Models, Springer-Verlag, 1990.
Figure 6 - Zwicker & Fastl, Psychoacoustics Facts and Models, Springer-Verlag, 1990.
Figure 7 - Zwicker & Fastl, Psychoacoustics Facts and Models, Springer-Verlag, 1990.