DYNAMIC RANGE REQUIREMENT FOR SUBJECTIVE NOISE FREE REPRODUCTION OF MUSIC

1772(B-2)

Louis D. Fielder
Ampex Corporation
Redwood City, CA

Presented at the 69th Convention 1981 May 12-15 Los Angeles

AES

AN AUDIO ENGINEERING SOCIETY PREPRINT

DYNAMIC RANGE REQUIREMENT
FOR SUBJECTIVE NOISE FREE REPRODUCTION OF MUSIC

Louis D. Fielder
Ampex Corporation
401 Broadway, Redwood City, CA 94063

Abstract

A dynamic range of 118 dB is determined necessary for subjective noise free reproduction of music in a dithered digital audio recorder. Maximum peak sound levels in music are compared to the minimum discernable level of white noise in a quiet listening situation. Microphone noise limitations, monitoring loudspeaker capabilities, and performance environment noise levels are also considered.

Introduction

The recent emergence of PCM recording techniques for music reproduction and the desire to standardize this format involves a re-examination of dynamic range requirements for natural music reproduction. Standardization of a 16 bit linear format would limit the dynamic range capability to 96 dB, and limit the quality of future PCM recorders if a wider range eventually became necessary. The most accurate of previous examinations of dynamic range requirements was done by Fletcher[1], who argued that 100 dB dynamic range was necessary. This was essentially a conjectural argument based on measurements of residential room noise by Hoff[2]. Unfortunately, these room noise measurements appear to be unrealistic; the spectra decreased too gradually compared with the frequency measurements presented here and thereby exhibit excessive noise in the most sensitive range of 3 - 7 kHz. In addition, Fletcher ignored the ear's ability to detect a noise source below that of the room noise by source localization.

In this work actual in situ noise threshold experiments and room noise spectra have been measured in ten home listening rooms and a number of studio environments. Next, a determination of any prerecording limitations that would affect the dynamic range requirement is made: effective microphone and recording environment dynamic range are investigated to ensure that under possible recording situations the dynamic range is limited only by the listener in the playback environment rather than shortcomings in recording equipment. Although the following method for determining necessary dynamic range is applicable to all recording techniques, the PCM recorder will be emphasized because it is the only technique that is capable of satisfying wide dynamic range criteria. It will be assumed that the noise spectra of the PCM recorder will be white and uncorrelated with the signal, as is the case of a dithered, unequalized PCM recording device.

The dynamic range criterion utilized in this paper is a comparison between the peak instantaneous sound levels occurring during a music performance and the just audible threshold for white noise when added to the program source. The minimum dynamic range requirement is determined by comparing a performance reproduced live at natural levels with the same signal entering through the recording channel: there should be no audible difference between the two due to background noise. Obviously this dynamic range requirement varies widely among performance situations; in some cases background noise is high while the performance itself produces relatively low peak sound levels and in other situations, such as recording in a quiet sound studio, the background noise may be extremely low while the peak sound levels are quite high. Because a recording system must be evaluated in terms of the most stringent situations, this analysis will concern itself with musical performances that have high peak sound levels in a quiet studio environment. To simplify the analysis and to provide for improvements in the electronic devices such as mixing consoles, equalizers and mike amplifiers, a very simple recording setup will be assumed.

Simplified Recording System

A schematic representation of a PCM music reproduction chain is shown in Figure 1 in which music reproduction is divided up into six basic sections. The first two sections are the recording side of the chain with the last three representing the playback process. The first section is the recording environment itself which consists of the musicians, audience, and the room in which the performance is being held. In this analysis, it is a quiet studio in which loud music is being performed. The next section consists of the microphones, the microphone amplifiers, and any mixing console before the recorder. It will be assumed for this analysis that microphones of the greatest dynamic range capability available today will be connected directly to the recording device to minimize dynamic range limitations.

The next element in the chain is the actual recording device itself. Any noise that this device produces on playback should either be below the threshold of the listener or be effectively masked by noise from the recording section. The playback section consists of the professional monitoring environment with the mixing engineer and producer who are assumed to have hearing acuity similar to the average 20-30 year's old audiophile. In this environment it is assumed that a normal complement of equipment is operating and the professional monitoring loudspeakers which are capable of reproducing natural sound levels are in use. In this typical monitoring room, the background noise of the environment is due almost entirely to acoustic noise generated by equipment fans and machinery. This makes the originally very quiet environment similar in noise level to the quiet home listening environment. The fifth step in the music reproduction chain is the processing and duplication that takes place from the master to produce the consumer copy. Although this step is not considered in this study, it is inevitable that some of the dynamic range present in the original master copy will be reduced during this duplication process. The final step in the chain is the playback environment of the consumer, represented by an average audiophile aged 20 - 30 with high quality audio equipment. Comparing the professional monitoring environment with that of the consumer, the major difference is that the consumer generally has speakers which are not capable of reproducing music at natural listening levels. Consumer equipment generally is capable of producing peak levels only between 100 - 110 dB SPL.

Maximum Level Determination

To determine this dynamic range requirement, it is first necessary to consider the peak instantaneous levels existing in a music performance. Unfortunately, very little information exists in the literature concerning absolute measurement of acoustic peak levels in music performances. Most studies investigate uncalibrated sound recordings or hearing damage rather than accurate music reproduction and thus have measured only average levels. This subject was studied by Sivian, Dunn and White[3] in their investigation of absolute peak and average acoustic pressures existing in classical performances. Additionally, several measurements were made during this study to obtain peak sound pressure information on several other types of performances. A literature survey on levels present in electronically amplified rock music yields only average levels. One such article by Cabot, Genter and Lucke[4] reports levels between 95 and 122 dB. Assuming even a very modest 6 dB ratio between the peak and average sound pressures in a rock performance, yields the prediction of peak levels up to 128 dB. These results for instantaneous sound levels at the listener position are summaried in Table 1:

Table 1 Peak Levels in Music

Source	Level	
Classical 75 piece orchestra	113	Sivian, White, Dunn
Classical 18 piece orchestra	112	Sivian, White, Dunn
Classical pipe organ	116	Sivian, White, Dunn
Country music electronically amplified (middle of crowd)	124	Author
Classical percussion music (front row center)	122	Author
Rock music extrapolated from average (electronically amplified)	128	Cabot, Genter, Lucke

This table reveals that classical music has peak levels up to at least 122 dB while electronically amplified music reaches levels of at least 128 decibels. While these measurements represent only a cursory look at the peak acoustic levels in music performances, it is unlikely that much greater peak levels than the ones reported exist. No special attempt was made to find examples of extremely high sound levels other than choosing performance situations incorporating music of a wide dynamic range. Thus, it is likely that these peak levels represent typical peak signals and may be used for a dynamic range determination.

In addition to the information on peak levels experienced by the audience, it is also useful to consider the peak levels available at short distances away from instruments. This applies to close miking situations in which the microphones are placed at various instruments and mixing is performed later. The paper by Sivian, Dunn and White[3] provides information for various musical instruments used in classical performances. The results are in Table 2:

Table 2 Peak Levels in "Close Miking" of Music

Source	Distance	Level
Bass drum	3 ft.	139 dB
Snare drum	4 ft.	124 dB
15" cymbals	3 ft.	129 dB
Trombone	3 ft.	121 dB
Trumpet	3 ft.	110 dB

From this table, it can be seen that extremely high amplitudes are present when certain musical instruments are recorded at close range. Because these levels are so high, microphone overload level probably limits the maximum sound pressure that the microphone can be exposed to rather than the ultimate loudness of the instruments. Examining the various recording microphones in use today it can be seen that few microphones have overload levels of 130 dB at their lowest noise setting. At least one microphone considered has this capability along with a low noise floor and it will be the one considered when examining recording limitations.

Noise Threshold Determination

The measurement of the minimum audible white noise sound level in the listening situation required a noise source which was acoustically flat between 1 - 10 kHz. One was built utilizing two 5 pole Chebyshev filters with adjustable upper and lower cutoff frequencies. A subject was exposed to this white noise source which was switched on and off at a half cycle rate. The subject then varied the acoustic level until he could just perceive the presence of the noise. Several measurements of this threshold were taken until consistent readings were obtained. In these experiments, the level equivalent to 20 kHz bandwidth white noise was obtained from a spot noise measurement at 4½ kHz. This was justified by further experimentation determining the frequency band responsible for the detection of the white noise. Two basic types of environments were examined. The first environment was the home listening room and was examined using ten listeners in their own rooms. This test revealed that the average white noise threshold 20 kHz bandwidth white noise threshold was equal to 4 dB SPL. The second environment investigated on a limited basis was the professional monitoring sutdio. These experiments involved three listeners in a typical good quality monitoring room. They were tested as above and a noise threshold of 4.7 dB SPL was obtained. The results for these noise threshold experiments are shown in Table 3:

Table 3 Threshold of Noise Detection for Various Listeners

	At Listener	
Subject	One Hertz Spot Noise 4.5 kHz	Level for 20 KHz Bandwidth
Home Environment		
1	45 dB	-2 dB
2	-41.5	1.5
3	-37	6
4	-40	3
5	-40	3
6	-40	3
7	-42	1
8	-35	8
9	-36.6	6.4
10	-34	9
Average	-39	4
Studio Environment		
1	-38	5 dB
2	-38	5
3	-39	4
Average	-38.3	4.7

0 dB = 2×10^{-4} dyne/cm^{-2}

These experiments on both environments were repeated with the noise bandpass filtered between 3 - 7 kHz. When listeners were tested in this way, a threshold shift less than 2 dB in the spot noise level at 4.5 kHz was obtained. This indicates that white noise perception is basically controlled by noise received in the 3 - 7 kHz band. This result supports the assumption that the 4.5 kHz region was most indicative for noise threshold criteria. This sensitivity at 4.5 kHz is also supported in the literature. In a paper concerned with hearing mechanisms governing loudness and masking effects, Fletcher[5] reports highest sensitivity to noise at 4 kHz. Another paper by Robinson and Whittle[6] on the loudness of octave noise indicates maximum noise sensitivity at 4 kHz for experiments using frontally incident noise. The results of these experiments were somewhat surprising in that threshold levels were much more dependent on the listener than on the environment even though the rooms had widely varying but moderately low noise levels. In fact, in the majority of the ten home situations and studio measurements the author's own threshold varied little from 4 dB. To get a significant threshold shift the noise had to be at least 15 decibels greater in the 100 - 10,000 Hz band than the average home listening room.

Additional measurements were taken in these environments to determine why such a low level white noise could be perceived in rooms which averaged broadband noise levels of 50 dB. It was discovered that noise levels in the critical 3 - 7 kHz region were lower than previously reported and that noise spectra decreased at 9 dB per octave rather than 6 dB per octave. The results of such measurements can be seen in Figure 2. This figure displays the one hertz noise level between 100 - 10,000 Hz for various playback environments: a monitoring room with and without mechanical equipment operating, the averaged results of the ten home listening room measurements, and a representation of the level for just audible white noise. In this figure, it can be seen that the total noise level in the listening environments is much greater than the threshold white noise signal. This indicates that there is no masking of the critical 3 - 7 kHz region by the high level low frequency noise. In addition, the room noise level present in the 3 - 7 kHz region was greater than the threshold white noise perceived. This led to the hypothesis that the ear was using direction clues to perceive the threshold noise. An experiment using headphones was performed in which simulated room noise (white noise filtered by a 9 dB per octave filter) was added to a threshold white noise signal. It was discovered in this case that the white noise threshold increased only when the simulated room noise exceeded the noise level everywhere in the region between 3 - 7 kHz. This demonstrated that even without directional clues, the ear was still unaffected by low frequency energy rising at 9 dB per octave. This explains why most quiet rooms will allow noise perception at threshold.

Support for the second conclusion was derived from comparison between the room and headphone threshold experiments. In the case of headphone measurements, which do not allow for directional clues, the subject was unable to detect the white noise addition when it was lower than the simulated room noise in the 3 - 7 kHz region. Support for these two conjectures was found in the literature. First, experiments by Young and Wenner[7] in which attempts to mask a white noise signal by narrow band noise at low frequencies on a headphone listener indicated that masking effects did not occur. Their results indicate that even in the case of the noisiest of the three listening environments shown in Figure 2, masking effects should not affect perception in the 3 - 7 kHz band. Second, the listener's ability to utilize spatial clues to perceive noise from a source was mentioned in a paper on necessary dynamic range by Stuart[8]. In this paper, it is claimed that the ear is able to use spatial clues to hear noise levels from a single source 10 dB below room noise which originates from all directions. This result is in basic agreement with the threshold experiment results presented here. It is therefore reasonable to hypothesize that the 4 dB threshold of hearing for white noise in music listening environments would not be lowered in an otherwise totally quiet environment. Assuming that the white noise perception in listening situations is identical to noise-free threshold values, it is permissible to compare previously existing work on quiet environment white noise thresholds. In a work by Thurlow and Bowman[9] considering the effect of repetition rate and duration on the threshold of noise, a threshold level of 4 dB SPL was obtained for 7 kHz bandwidth noise produced by headphones. Extrapolating this result to a 20 kHz bandwidth would raise this noise threshold to 9.5 dB SPL. Comparing this result with the threshold levels obtained in listening rooms results in a 5.5 dB discrepancy. A possible explanation is that in a room listening situation the ear is able to use the diffraction properties of the human head and ear to obtain an effective amplification of the noise impinging on the head. This explanation is supported by the comparison between the head phone and room hearing measurements by the author and in a paper by M. Killon[10], who considered the relationship between various types of hearing threshold experiments. In the critical region of 3 - 7 kHz, he reported a 4 dB difference while a difference of 6 - 9 dB was observed by this author. The 9.5 dB figure from Thurlow and Bowman should then have a correction factor of 6 dB applied to it, producing a 3.5 dB result which is in good agreement with the 4 - 5 dB result obtained here. In summary, it is found that a single noise source producing a 4 dB sound level over a 20 kHz bandwidth is just inaudible for listeners in a home listening environment or a professional monitoring environment. This value is found to be consistent with the previously mentioned psychoacoustic literature involving noise perception and masking.

Recording Environment Limitations

The final area to be investigated is the effect limitations on the recording side of the music reproduction chain have on the playback derived dynamic range requirement. Limitations resulting from microphone or recording environment properties are now considered. To this end equivalent acoustic noise spectra for recording environments and microphones are measured or obtained from the literature (Meares[11] in a BBC research report dealt with recording studio noise standards and provided recording studio noise spectra). These measurements are considered in light of the previous conclusions: first, that a listener is most sensitive to noise intrusion in the 3 - 7 kHz region; second, that masking effects on the 3 - 7 kHz region will be negligible as long as the disturbing environmental noise is less than that of the average home listening room or studio; and finally that in cases where spatial clues are not present for the listener, noise perception will only take place when the intruding noise is greater or equal to the existing noise in 3 - 7 kHz region. Figure 3 represents noise spectrums of various recording environments, a white noise threshold level, and a wide dynamic range recording microphone. With the exception of the spectrum of audience and orchestra noise during a classical music performance, the other recording environments have noise levels significantly below that of the quiet recording microphone in the critical region. Thus, in these latter cases the limitation will be that of the recording microphone noise in the critical frequency region with the threshold level. If the microphone is positioned at typical audience position, the total system dynamic range would be reduced 5 dB since the microphone noise averages 5 dB greater than the 4 dB room threshold in the 3 - 7 kHz region. To remove this limitation it is simply necessary to use a close miking technique in which a higher acoustic level is available to the microphone. This technique is useful as long as the microphone overload point is higher than the maximum sound levels at the audience position. For this particular microphone, the overload point is 130 dB and thus would allow the capturing of an equivalent dynamic range of 121 dB if peak levels of 130 dB exist in a performance. From the tabulation on peak sound levels close to musical instruments in Table 3, it is seen that musical instruments are capable of producing these high sound levels especially at distances less than 3 feet. In an example of an actual classical performance with audience present, it is interesting to note that the microphone noise was less than the environment noise. This noise spectrum was acquired by recording the quiet sections in a traditional classical performance and the splicing out the music passage to determine the "quiet" noise spectrum when recorder background noise would be most likely heard. Examining the resulting spectrum in the 3 - 7 kHz region yielded an average noise level equivalent to 16 dB white noise. This then reduces the dynamic range requirement by only 12 dB from that of a quiet studio, and demonstrates that even with audience present music reproduction requires wide dynamic range. Several other microphones were also measured to ensure that the performance represented by the microphone in Figure 3 was comparable to other existing microphones. The results are shown in Figure 4. Four different microphones were measured which had overload levels between 120 to 140 decibels. They were all condenser microphones and as the graph shows the noise levels in the 3 - 7 kHz region were witin 5 dB of each other. In summary, it is shown that close miking techniques and the proper selection of a microphone produces no limitation or reduction on the dynamic range requirement as determined by the playback experiments. Even a natural miking technique results in only a 9 dB white noise threshold. In the case of a large auditorium with orchestra and audience present only a 16 dB noise threshold level for white noise perception is obtained.

Limitations of A-Weighted Measurements

An additional conclusion coming from these threshold tests is that A-weighted noise measurements have limited usefulness in determining the perceptual effect of the addition of a number of noise sources with different spectral characteristics. In the case of recording and playback environments, noise levels as measured by an A-weighted technique do not allow the determination of whether the listener will be sensitive to background noise introducted by the recording device. Figure 5 demonstrates this very clearly. Figure 5 is a representation of the contribution versus frequency after A-weighting of the noise of three recording environments; a professional monitoring room, the average home listening room, and a recording studio. The sound levels at each particular frequency were corrected by the A-weighting squared, and then normalized for ease of graphical comparison. When this is done, contributions to the final value are represented by the area under the curves and in all three cases the major contribution is due to energy below 800 cycles. This indicates that the A-weighted measurement value has little or no bearing on the levels present in the critical 3 - 7 kHz region. This lack of correlation between 3 - 7 kHz noise levels and A-weighted noise measurements is also true for microphones but to a lesser extent. Two microphones tested differed by 5 dB in the 3 - 7 kHz region but had identical A-weighted noise values because the quieter of two had more noise below 500 Hz. In summary, it is shown that the A-weighted measurement has very little use in determining the effective noise introduced by various parts of the music reproduction chain. A more useful technique would be the comparison of the various spectra of noise produced by the separate parts of the reproduction chain with special emphasis on the 3 - 7 kHz region.

Dynamic Range Requirement

Comparing the previous data on the peak sound levels to the white noise thresholds it is possible to derive a realistic dynamic range requirement for a noise-free recording system. Non-electronically amplified music with a peak level of 122 dB has been measured and by comparison with a white noise threshold of 4 dB yields a 118 dB range requirement. Electronically amplified music may create an even larger requirement because of the higher sound levels possible. This requirement for natural level music reproduction will apply in musical situations where the performers are closed miked and in a very quiet environment such as a recording studio. Typical classical music performances in which orchestra and audience are present can reduce the dynamic range requirement to 106 dB. Microphone limitations can reduce this range if natural miking techniques are used but in any event reduce it no more than 5 dB resulting in 113 dB. Considering now the dynamic range for consumer reproduction it is seen that the only difference between consumer and professional listening situations is that the consumer's loudspeakers produce significantly lower sound levels than professional monitoring speakers. Today, quality home audio systems have peak sound pressure capabilities of up to 110 dB, yielding a present requirement of 106 dB for the consumer.

Conclusion

In conclusion, several experiments were made to determine the dynamic range requirement for a recording system to produce no audible hiss when used to play back music at natural listening levels. These experiments resulted in a dynamic range requirement of 118 dB (non-amplified music), 124 dB (amplified music) for the professional, and 106 dB for the high quality consumer playback system. Both the results of peak sound levels and noise threshold were reconciled with the existing literature and it is believed that these dynamic range requirements are realistic. Even though the analysis considered the more extreme situations, the need for an additional 10 dB head room margin in the professional recording situation makes these requirements even more plausible. It was also discovered that A-wieghted noise measurements were not suitable in the evaluation of the effect the measured item had an overall system background noise as perceived by the listener.

Acknowledgement

I would like to thank Paul Stubblebine and The Automatt recording studio for assistance in obtaining the studio data presented here. I am also grateful to Bob Meuse and the San Jose Symphony Orchestra for assisting me in the measurement of the noise present during quiet periods in a classical music performance.

References

1. H. Fletcher, "Hearing the Determining Factor for High Fidelity Transmission", Proc. IRE, Vol. 30, pp. 265-277, June 1942.

2. D. F. Hoth, "Room Noise Spectra at Subscribers Telephone Location", JASA, Vol. 12, pp. 499-504, 1941.

3. L. J. Sivian, H. K,. Dunn, and S. D. White, "Absolute Amplitudes and Spectra of Certain Musical Instruments and Orchestras". JASA, Vol. 2, pp. 330-371, 1931.

4. R. C. Cabot, C. R. Genter, and T. Lucke, "Sound Levels and Spectra of Rock Music", AES Preprint, No. 1358 (G-6), 1978.

5. H. Fletcher, "Loudness, Masking and Their Relation to the Hearing Process and the Problem of Noise Measurement", JASA, Vol. 9, pp. 275-293, 1938.

6. D. W. Robinson, and L. S. Whittle, "The Loudness of Octave Bands of Noise", Acoustica, Vol. 14, pp. 24 - 35, 1964.

7. I. M. Young and C. H. Wenner, "Masking of White Noise by Pure Tone, Frequency-Modulated Tone, and Narrow Band Noise", JASA, Vol. 41, pp. 700-706, 1967.

8. J. R. Stuart, "Tape Noise Reduction", Wireless World, Vol. 78, pp. 104-110, 1972.

9. W. R. Thurlow and R. Bowman, "Threshold for Thermal Noise as a Function of Duration and Interruption Rate", JASA, Vol. 24, pp. 281-283, 1952.

10. M. Killon, "Revised Estimate of Minimum Audible Pressure", JASA, Vol. 63, pp. 1501-1508, 1978.

11. D. J. Mearies, "Revised Background Noise Criteria for Broadcasting Studios", BBC Research Report, RP1980/8, 1980.

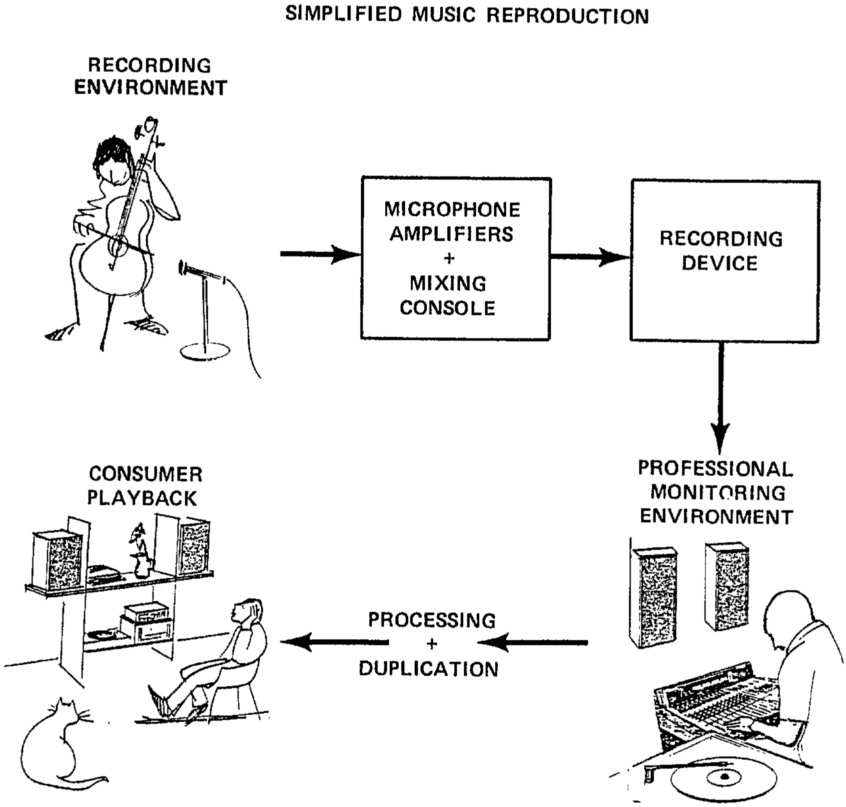

Figure 1 Schematic Representation of a Music Reproduction System

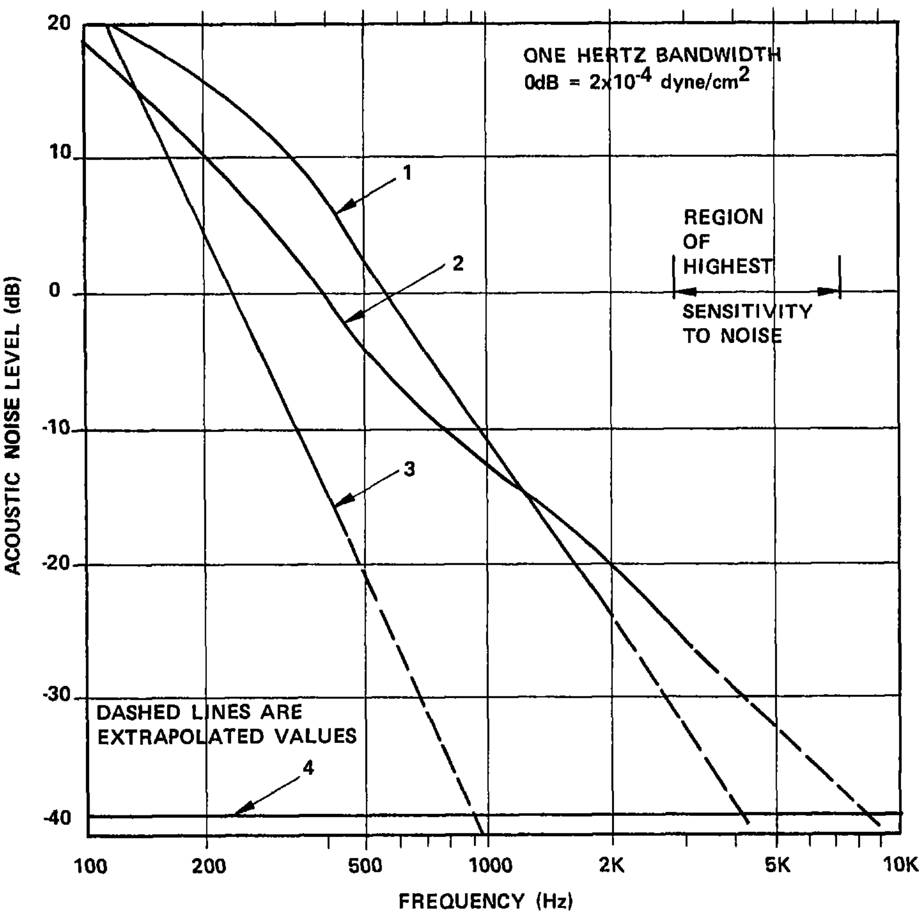

Figure 2 Noise Spectra for the Playback Environment

1. Professional Monitoring Room with Equipment On. (Automatt, S.F.)
2. Average Result of Ten Home Listening Rooms
3. Professional Monitoring Room with Equipment Off. (Automatt, S.F.)
4. Threshold Level for White Noise

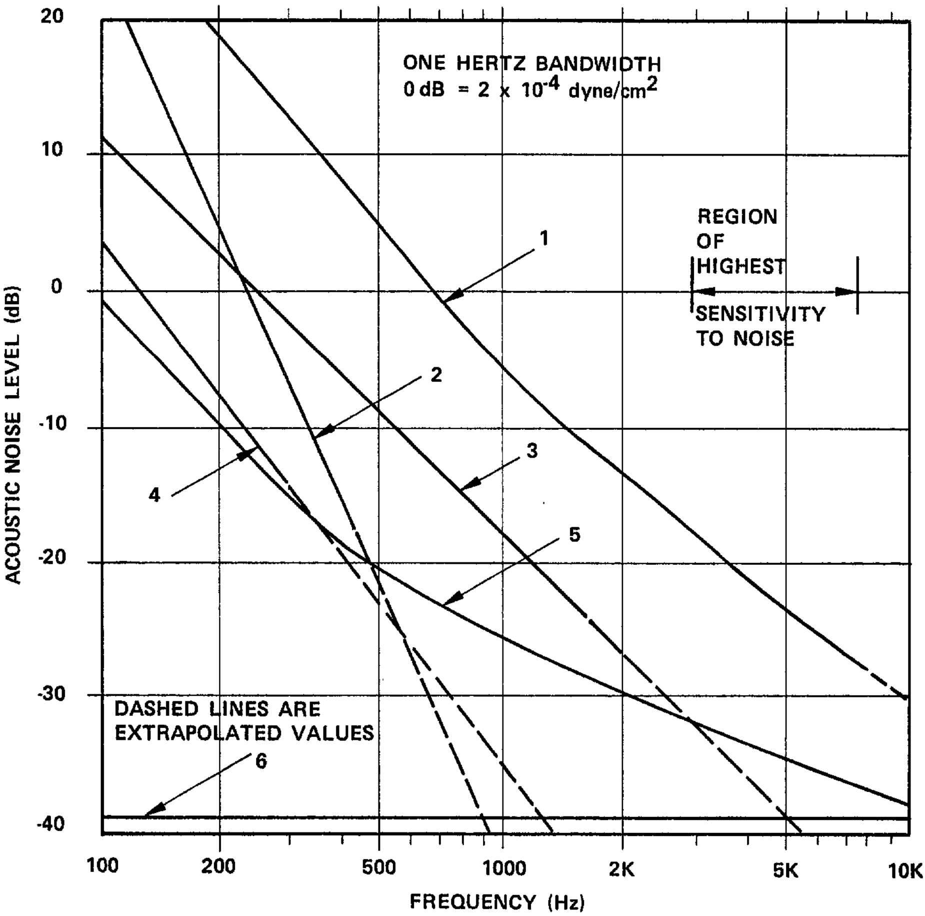

Figure 3 Noise Spectra for the Recording Environment

1. Quiet Noise Level During a Classicial Music Performance with Audience and Orchestra Present. (San Jose Center for the Performing Arts)
2. Music Recording Studio. (Automatt, S.F.).
3. BBC Music Recording Studio (BBC research report [11]).
4. BBC Drama Recording Studio (BBC report [11]).
5. Best Measured Recording Microphone Noise.
6. Threshold Level for White Noise.

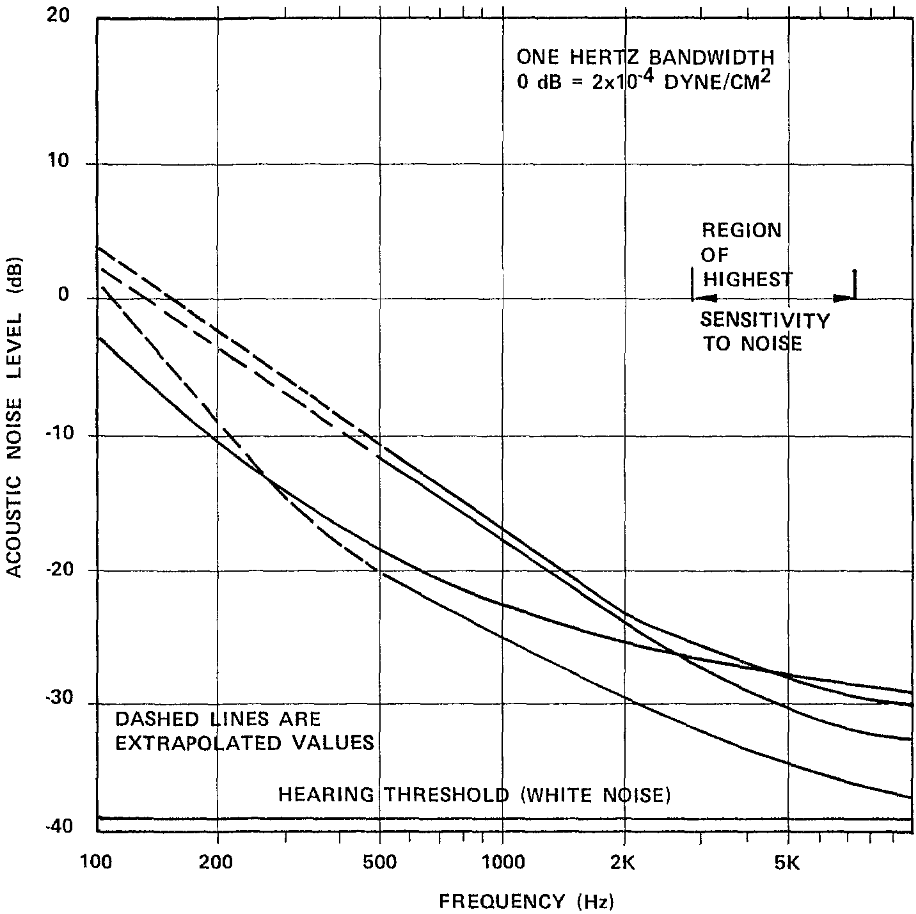

Figure 4 Equivalent Acoustic Self Noise for Various Recording Microphones

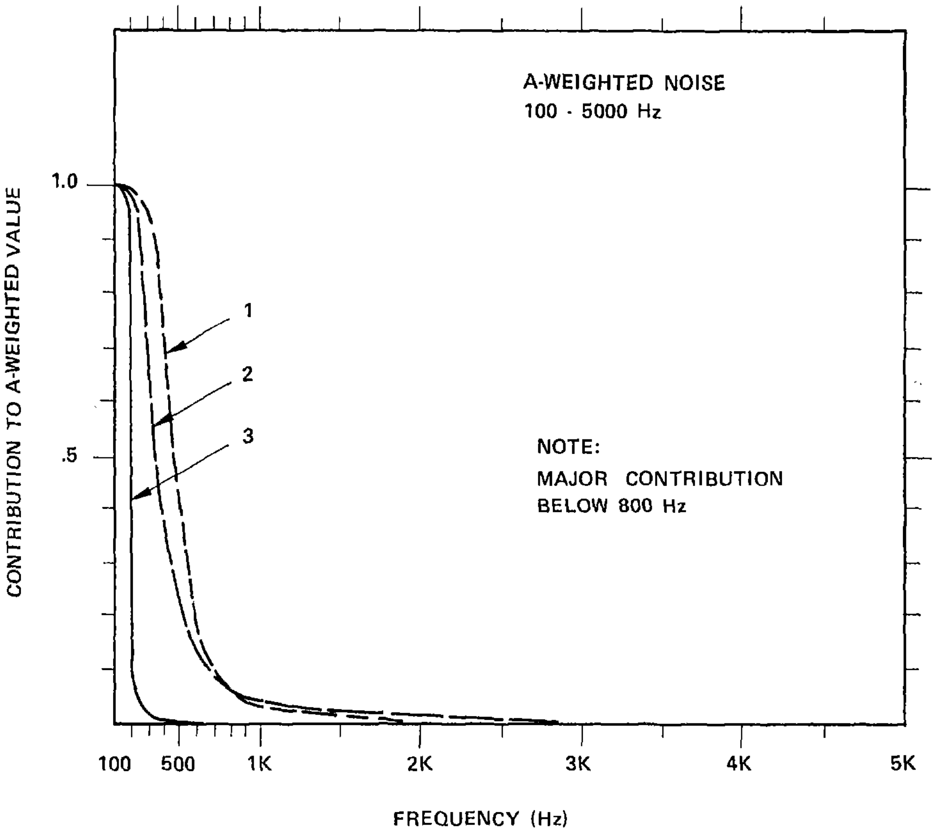

Figure 5 Contribution Versus Frequency in an A-Weighted Sound Level Measurement of Various Rooms

1. Professional Monitoring Room. (Automatt, S.F.).
2. Average Home Listening Rooms.
3. Music Recording Studio (Automatt, S.F.)