Listening conditions for the assessment of sound programme material: monophonic and two–channel stereophonic

EBU Tech. 3276 – 2nd edition May 1998

CONTENTS

European Broadcasting Union
Case Postale 67, CH–1218 Grand–Saconnex (Geneva) Switzerland

Introduction

"Listening" is an integral part of all sound and television programme–making operations. Despite the very significant advances of modern sound monitoring and measuring technology, these essentially objective solutions remain unable to tell us what the programme will really sound like to the listener at home. The human ear alone is able to judge the aesthetic or artistic quality of programme material and, indeed, certain aspects of the technical quality. However, to permit the ear to make fair, unbiased judgements it must be allowed to function under favourable conditions.

The listening conditions specified here are essentially relevant in the following two situations:

– *Reference listening rooms:* Listening rooms used for the critical assessment and selection of programme material for inclusion in a sound or television broadcaster's programme output.

– *High–quality sound control rooms:* Sound control rooms used for the critical assessment of sound quality as a part of the sound or television broadcast production process.

The accuracy and quality of listening conditions depend on the relevant sound field parameters affecting the ears of the listener. Their definition will pose a number of design constraints upon the characteristics of the loudspeakers used for monitoring and the properties of the listening room. It may be noted that, in the case of headphone listening, the properties of the room have practically no influence on the listening conditions.

The main part of the present document sets out the basic requirements for the sound field parameters. Four *Appendices* then give recommendations regarding ways in which these requirements can be met.

To a greater or lesser extent, the listening conditions given here may inspire the implementation of listening facilities in other contexts, such as programme fairs, the assessment of programmes in competitions and subjective testing of the technical quality of sound systems.

Sound field parameters for listening rooms

1. General considerations

1.1. Loudspeaker listening

The quality of the listening conditions in a listening room is defined by the properties of the sound field produced by the monitor loudspeaker(s) in the listening area (see *Appendix 1*) at the height of the listener(s) ears (about 1.2 m above the floor level).

The main components of the sound field are the direct sound, the early reflections and the later reflections which form the reverberant field. All these components are time and frequency dependent.

1.2. Headphone listening

When headphones are used, the room has practically no influence. The requirements for monitor headphones are given in *Appendix 4*.

2. Acoustic parameters

2.1. Direct sound

The direct sound is defined as the sound field which would be measured, using the same loudspeakers, under anechoic conditions, i.e.: without the early reflections and the reverberation caused by the listening room (see *Sections 2.2. and 2.3.*).

The quality of the direct sound is determined by the relevant loudspeaker parameters. Specifications for reference monitor loudspeakers suitable for the applications considered in this document are given in *Appendix 3*.

2.2. Early reflections

Early reflections are defined as reflections from boundary surfaces or other surfaces in the room which reach the listening area within the first 15 ms after the arrival of the direct sound. The levels of these reflections should be at least 10 dB below the level of the direct sound for all frequencies in the range 1 kHz to 8 kHz.

The amplitude and frequency responses of individual reflections may be derived from a measurement of the room impulse response, using Fourier transform methods. The fundamental constraints of time and frequency will limit the resolutions which can be obtained. It is important that the effective time window and bandwidth of the measurement are appropriate[1]. The measurement of early reflections of sounds including components in the frequency range below about 500 Hz may be difficult.

1. Significant underestimates of the reflection amplitude will occur if the time window or the bandwidth exceed those of the reflection. In such a case, the effective measurement would include some parts of the time/frequency domains which did not contain the reflection energy and, by the averaging process implicit in any measurement, tend to reduce the estimate of the reflection amplitude.

The effect of early reflections may also be observable, as a comb filter effect, in the operational room response curve (see *Section 2.4.*).

It should be noted that, in control rooms, the top surface of the mixing desk (console) is a potential source of strong early reflections.

2.3. Reverberation field

Reverberation is the effect caused by multiple reflections from the boundary surfaces of the listening room, which reach the listening area after the early reflections (time delays more than about 15 ms).

The reverberation field should be sufficiently diffuse over the listening area to avoid perceptible acoustic effects such as flutter echoes

The reverberation time is an important characteristic of the reverberation field; it is defined as the time taken for the sound to decay to 60 dB below the initial level. It is usually measured over the range from 5 dB to at least 25 dB below the initial value. The decay time of the measuring instrument and the filters should be shorter than the decay time of the reverberation field. The reverberation time should be measured in the listening room with 1/3rd octave filtering [1] using the listening loudspeakers as the sound sources.

Reverberation time is frequency–dependent. The nominal value, T_m, is the average of the measured reverberation times in the 1/3–octave bands from 200 Hz to 4 kHz.

The nominal reverberation time, T_m, should lie in the range:

$$0.2 < T_m < 0.4 \text{ s}$$

To ensure that the acoustic environment remains “natural”, the value of T_m should increase with the size of the room. The following formula is given as a guide:

$$T_m = 0.25(V/V_0)^{1/3} \text{ s}$$

where: V = room volume in cubic metres
V_0 = reference room volume of 100 m^3.

The reverberation time T, measured in 1/3–octave bands over the frequency range from 63 Hz to 8 kHz, should conform to the tolerance mask shown in *Fig. 1*.

In addition, sudden changes in reverberation time with frequency should be avoided and the differences, ΔT, in reverberation times between adjacent 1/3–octave bands should not exceed the following limits:

ΔT	< 0.05 s	for 200Hz $\leq f \leq$ 8 kHz
	< 25% of longer time	$f \leq$ 200 Hz

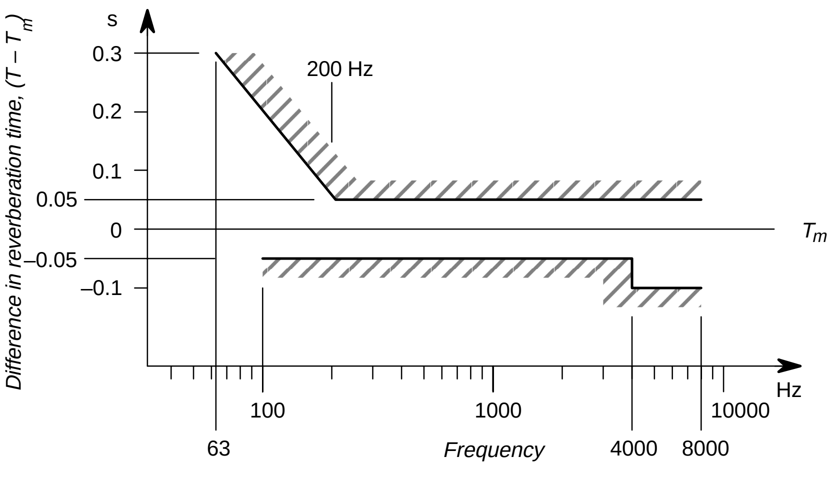

Fig. 1 – Tolerance limts for reverberation time.

2.4. Operational room response curve

The operational room response curve is defined as the frequency response of the sound pressure level produced by the loudspeaker(s) at any listening point. The test signal for this measurement is 1/3–octave filtered pink noise, as specified in [1].

The operational room response curve is an important criterion in the evaluation of the mutual influence of the loudspeaker and the listening room, and hence for the assessment of the listening conditions. It corresponds well with the subjective assessment of reproduced sound.

Tolerance limits for the measured operational response curves are given in *Fig. 2*. L_m is the mean value of the levels of the 1/3–octave bands with centre frequencies from 200 Hz to 4 kHz. The tolerances should be met for each (main) channel separately. For stereophonic reproduction, the close matching of the room response of each channel is important.

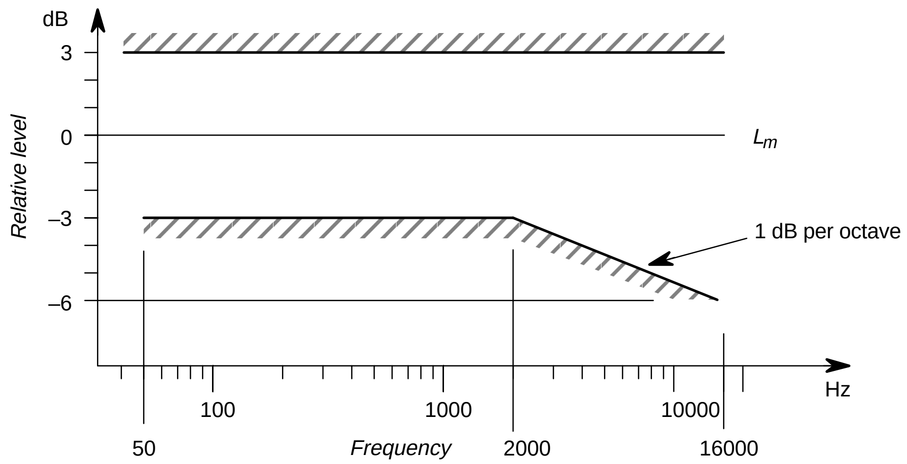

Fig. 2 – Tolerance limits of the operational room response curve.

Notes:

– It may be difficult to achieve the tolerances for the room response curves in the listening area at low frequencies, especially in small rooms. They should be satisfied preferably by acoustic treatment of the room and/or by adjustments to the positions of the loudspeakers and listening positions.

– Electrical correction (equalization) may be necessary when linearity of the operational room response curve cannot be achieved by the above–mentioned methods. The adjustment can be carried out using frequency response adjustments of the monitor loudspeakers or by means of external equalizers.

– To avoid degrading the quality of reproduction, electrical equalization should be used carefully. It is advisable to make the corrections in the low–frequency range ($f < 300$ Hz) only. All channels should be adjusted in the same way.

2.5. Listening level

The reference listening level, $L_{LISTref}$, characterizes the sensitivity of a reproduction channel. It is used to set the reference gain (0 dB) for level adjustments in listening sessions.

A loudspeaker reproduction channel generally consists of a volume control and a monitor loudspeaker which may be a complex device including amplifiers, filters and other elements (see *Appendix 3*). The channel is adjusted to the reference listening level as follows.

A pink noise test signal is fed to each reproduction channel separately. The RMS level of the test signal shall be at “Alignment Signal Level”, that is:

– –9 dB with respect to PML (Permitted Maximum Level) in analogue devices (ITU–R Recommendation BS.645 [2]);

– –18 dB with respect to dB–FS (full scale digital level) in digital devices (EBU Technical Recommendation R68 [3]).

The gain of the reproduction channel is adjusted so that the A–weighted sound pressure level (RMS, slow) [5] at the (reference) listening point is:

$$L_{LISTref} = 85 - 10\log(n) \quad \text{dB(A)}$$

where: n = number of reproduction channels (main loudspeakers) in the total configuration.

The difference between the levels of any two channels should not exceed 1 dB. For two–channel stereophony close matching of the loudspeakers is especially important.

Notes:

– The measuring signal is available on the EBU R–DAT alignment tape [4].

– By increasing the gain by 10 dB, sound levels in excess of 100 dB(A) are available (see *Appendix 3*).

It should be noted that high sound levels may cause temporary or permanent hearing damage. The provisions of statutory and employer's workplace hearing protection codes must be observed.

No satisfactory method for measuring sound pressure levels produced by headphones can be recommended. The level should be adjusted in such a way that a perceived loudness equal to a reference sound field produced by loudspeakers is achieved.

2.6. Background noise

The sound pressure level (RMS, slow) [5] of the continuous background noise from air conditioning systems or other external or internal sources, measured in the listening area at a height of about 1.2 m above the floor level, should preferably not exceed NR 10 [*Table 1, Fig. 3*]. Under no circumstances should the background noise exceed NR 15.

The background noise should not be perceptibly impulsive, cyclical or tonal in nature.

Table 1 – Noise Rating curves.
Octave band sound pressure level values in dB re. 20mPa.

NR curve	Frequency (Hz)								
	31.5	**63**	**125**	**250**	**500**	**1k**	**2k**	**4k**	**8k**
10	62.2	43.4	30.7	21.3	14.5	10.0	6.6	4.2	2.3
15	65.6	47.3	35.0	25.9	19.4	15.0	11.7	9.3	7.4

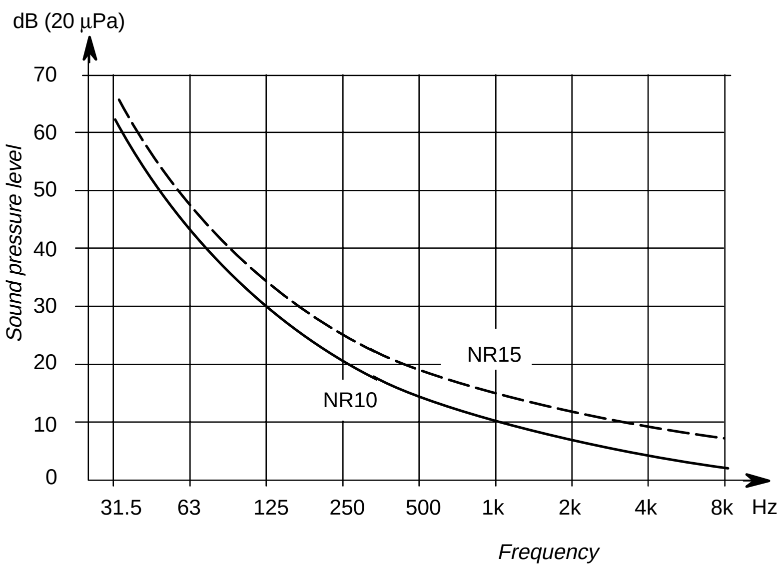

Fig. 3 – Noise Rating curves

Appendix 1

Monophonic and stereophonic loudspeaker listening: positioning of sound sources and listeners.

1. General

The height of the acoustical centre of the monitor loudspeaker should be at least 1.2 m. The height and the inclination of the monitor loudspeakers should be determined in such a way that their reference axes at the reference listening point intersect at the height of the ears of a seated person (approx. 1.2 m above the floor level). The inclination angle of the reference axes in relation to the horizontal plane should not be more than 10°.

If the monitor loudspeakers are not installed into the wall, the distance of their acoustical centres from the surrounding walls should be at least 1 m.

The propagation of the direct sound should not be affected by any obstacles between the loudspeakers and the listening area (this is not so important at the lowest frequencies, below 150 Hz).

Time delay differences between the channels of a stereophonic system should not exceed 100 µs[2].

All listening positions should be situated at least 1.5 m from the side walls and the back wall of the room.

2. Stereo listening

Two monitor loudspeakers should be placed in the listening room for stereophonic reproduction according to the layout given in *Fig. 4*.

The monitor loudspeakers should be located symmetrically with respect to the perpendicular passing through the centre of the stereo base, *b*. The reference listening position is represented by the location of the

2. These requirements remain to be studied.

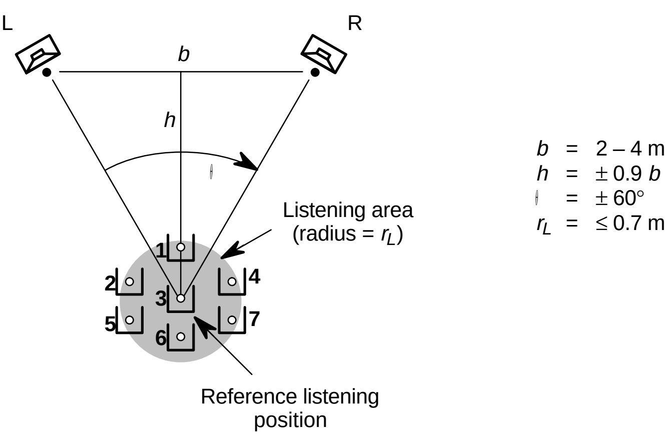

Fig. 4 – Typical layout of stereo listening arrangement.

listener's head (about 1.2 m above the floor) at the centre of the listening area. The reference listening position is specified by the optimal base angle = 60°, corresponding to the optimal listening distance $h \approx 0.9\ b$.

Admissible limits of the base width, *b* are:

$$2.0 < b < 4.0 \quad \text{m}$$

The radius of the listening area, r_L, shown in *Fig. 3*, will not usually exceed 0.8 m.

For the evaluation of stereo impression, a listening position on the perpendicular line passing through the middle of the stereo base is preferable.

For control of the compatible mono signal (M = L + R), a third monitor loudspeaker might be provided. If used, this loudspeaker should be in the middle of the stereo base at the same distance from the reference listening point as the stereo loudspeakers.

The quality parameters concerning the loudspeaker responses are given in *Chapter 2* of the main text and in *Appendix 3*.

3. Separate bass loudspeakers

The use of separate loudspeakers for reproducing the lowest part of the frequency spectrum has become commonplace. The potential advantages include:

- freedom to locate the bass sources optimally in relation to the low–frequency room mode pressure distributions;
- reductions in the sizes of the main loudspeakers – this is of primary importance for multi–channel listening, where the combined size of many full–range loudspeakers can be difficult to accommodate, but may also be a significant consideration in two–channel stereophony;
- reduction in distortion because the main loudspeaker driver displacements can be reduced.

The optimum cross–over frequency between the bass and the main loudspeakers depends on many factors, including the positions of the loudspeakers in the room, the room acoustics and the desired overall frequency response (see *Appendix 3*).

To prevent the separate bass source locations from becoming perceptible, lower cross–over frequencies will be required for bass loudspeaker positions which are located further from the main loudspeakers.

Appendix 2

Design of listening rooms and sound control rooms

1. General

This *Appendix* is concerned with room design features which favour the establishment of the sound field parameters described in the main text.

Two types of listening room are specifically covered by the specifications in this *Appendix*:

– Reference listening rooms used for listening tests and listening group activities for between three and seven listeners.

– High–quality sound control rooms such as sound control rooms for music and drama recordings and post–production operations.

In these rooms, high–quality loudspeakers of the type referred to as reference monitor loudspeakers are normally used. The performance characteristics of such loudspeakers are discussed in *Appendix 3*. The listening room has practically no influence when headphones are used; indications regarding the use of headphones for listening are given in *Appendix 4*.

The walls, floor and ceiling of the listening room define the boundaries of a controlled acoustic environment, and to achieve consistent results, care must be taken in the design of the room itself and in the placement of the loudspeakers and the listeners. The quality of the listening environment is governed by the properties of the sound field produced by the loudspeakers in the listening area, at the height of the listeners' ears.

2. Listening room dimensions

The minimum floor area should be:

40 m^2	for a reference listening room;
30 m^2	for a high–quality sound control room.

The size of a listening room will be determined as much by the operational aspects (installation of technical equipment, etc.) as by the acoustic aspects. In any case, the volume should not exceed 300m^3.

To ensure a reasonably uniform distribution of the low–frequency eigentones, the proportions of the room must lie within controlled limits. The following limits for the length to height and the width to height ratios should be observed:

$$1.1w/h \leq l/h \leq 4.5w/h - 4$$
$$l < 3h$$
$$w < 3h$$

where:

l	=	length	(larger dimension of floor plan, irrespective of orientation)
w	=	width	(shorter dimension of floor plan, irrespective of orientation)
h	=	height.	

In addition, ratios of l, w and h which are within ± 5% of integer values should also be avoided.

All dimensions are overall, measured to the internal structural surfaces.

3. Other design considerations

Careful design and good workmanship in the construction of listening rooms can greatly enhance the acoustic environment.

a) The room should exhibit sufficient acoustic symmetry about the listening direction. This requirement concerns not only the shape but also the absorption of the surfaces. Surfaces in the neighbourhood of the loudspeakers are particularly important in this respect.

b) The location of sound absorbers and reflecting surfaces on the walls, ceiling and floor of the room should be chosen to avoid flutter echoes and disturbing single reflections.

c) Any resonating structures in the room must be sufficiently damped (attenuation time shorter than the room reverberation time). The room must not contain any structures which cause audible rattling during loudspeaker listening.

d) The acoustic requirements and recommendations should be taken into account in regard to all surfaces and objects in the room, such as the door and window surfaces and the technical equipment. In particular, the working surface of the mixing desk can cause a disturbing early reflection.

e) For airborne and impact noise insulation, the background noise requirements in *Section 2.6.* of the main text should be considered.

Appendix 3

Reference monitor loudspeakers

1. Terms and definitions

Reference monitor loudspeaker	A monitor loudspeaker is a high–quality sound reproducing device used for technical and artistic quality control and subjective evaluation of sound programme signals. It consists of electroacoustic transducers, cabinets and/or wave guides, electronic equalizers, crossover networks and power amplifiers. A reference monitor loudspeaker system also includes any separate bass loudspeakers. To qualify as a reference monitor, the loudspeaker should at least fulfil the requirements given in *Section 2* of this *Appendix*.
Normal setting	Setting of frequency response and other adjustments, indicated by the manufacturer, which are usually applied in normal operational conditions.
Acoustical centre	A reference point used to define the location of a loudspeaker. The acoustical centre should be indicated by the manufacturer.
Main axis	The principal axis of the loudspeaker as defined by the manufacturer.
Measuring distance	The distance from the acoustical centre of the loudspeaker to the measuring point.
Directional angle	The angle between the main axis and the straight line connecting the acoustical centre with the measuring point.

2. Technical requirements

2.1. General

The specifications in this *Section* are the minimum objective requirements for a reference monitor loudspeaker. However, a loudspeaker which meets these requirements may not necessarily be suitable as a reference monitor loudspeaker; an evaluation by subjective listening tests is necessary when selecting a type of monitor loudspeaker for this critical function.

Measurements on loudspeakers are taken in an anechoic chamber or other measurement environment allowing the loudspeaker to operate at its normal settings under free–field conditions.

The measuring distance is selected taking into account practical conditions such as the size of the loudspeaker and the characteristics of the measuring room. The measuring distance should be great enough (about 2 m) to ensure that the conditions at the measuring point correspond to the direct sound conditions at the normal listening distance.

In the specifications which follow, all the sound pressures, voltages etc. are given as RMS values. When an absolute sound level value is of importance it is given for a reference distance of 1 m.

An accuracy of the order of ±0.2 dB should be obtained for the measurement of the electrical parameters. The measurement inaccuracy caused by imperfect sound field conditions in the anechoic measuring chamber should be less than ±1 dB over the whole response range. It may be appropriate to use a time–windowing measuring technique in order to achieve this requirement.

2.2. Frequency response

The frequency response curve is measured in 1/3–octave bands, with a pink noise test signal. The measurements are taken on the main axis (directional angle = 0°). The curve should fall within a tolerance band of 4 dB over the frequency range from 40 Hz to 16 kHz.

Frequency response curves measured at directional angles ±10° and ±30° should not differ from the frequency response measured on the main axis (directional angle = 0°) by more than the following permissible deviations:

±3 dB	for directional angles in the range	±10°
±4 dB		±30°

Requirements for directional angles of ±30° refer to the horizontal plane only.

2.3. Directional characteristics

2.3.1. Polar patterns

It is recommended that the polar patterns in the horizontal plane are measured in 1/3–octave bands, with pink noise, at least for the standard octave frequencies from 250 Hz to 16 kHz.

The radiation maximum should be in the direction of the main axis at all frequencies. The directivity of the monitor loudspeaker should be symmetrical with respect to the reference axis. If geometrically non–symmetrical monitor loudspeakers are used, mirror symmetry with respect to the perpendicular line passing through the middle of the stereo base is required.

2.3.2. Directivity index

The directivity index, D, is defined as the ratio, expressed in decibels, between the acoustic power delivered in the direction of the main axis and the power that would be delivered in that direction if the source were omni–directional. The directivity index should be measured with 1/3–octave band noise at standard frequencies according to IEC Publication 268–5 [6]. Two measuring methods are available:

a) Calculation of the directivity index from the directional characteristics.

b) Measurement of the directivity index through sound pressure level measurements in the free field and in the diffuse field.

The directivity index should be within the following limits over the frequency from 250 Hz to 16 kHz:

$$4 \leq D \leq 12 \text{ dB}$$

Large, frequency–dependent variations are undesirable, especially in the high–frequency range ($f > 500$ Hz).

2.4. Distortion

Harmonic distortion is measured with sinusoidal signals. A constant voltage input signal is supplied to the loudspeaker. This signal should produce an average sound pressure level of 90 dB in the frequency range from 250 Hz to 2 kHz.

With respect to this level, no harmonic distortion component shall exceed the following values:

–30 dB (3%)	for	40 Hz $< f <$ 250 Hz
–40 dB (1%)	for	250 Hz $< f <$ 16 kHz

Requirements for other kinds of distortion (intermodulation distortion, etc.) are still under study.

2.5. Decay time

The decay time t_s is defined as the time taken for a reduction of the sound pressure to a value $1/e$ (about 0.37) times the initial value. A sinusoidal tone burst shall be used as an input signal.

The decay time should not exceed the following limit:

$$t_s \leq 2.5/f$$

where: f is the frequency.

It is usually sufficient to carry out measurements of the decay time on the reference axis only.

The frequency response curves, measured with a sinusoidal signal, may also give some information on possible defects in the transient response.

2.6. Time delay

When sound accompanies pictures, the time delay of the loudspeaker systems should not cause the relative delay of the sound and vision components at the listening position to exceed that defined in EBU Technical Recommendation R37 [7]

2.7. Dynamic range

2.7.1. Maximum operational sound pressure level

The maximum operational sound pressure level $L_{eff–max}$ is defined as the maximum sound pressure level of a continuous programme signal which the monitor loudspeaker can produce for a period of at least 10 minutes without thermal or mechanical damage and without overload circuits being activated. The maximum operational sound pressure level which a monitor loudspeaker is able to produce is verified using a simulated programme signal according to IEC Publication 268–1 [7].

It should be noted that if the protection circuits are not working properly it is possible that the loudspeaker may be damaged during this test.

The maximum operational sound pressure level, measured using a sound level meter set to flat response and RMS (slow) [5] at the reference distance of 1 m is:

$L_{eff–max} \geq 108$ dB.

2.7.2. Self–generated noise level

The maximum allowable self–generated noise level L_{noise} measured as an A–weighted sound pressure level (RMS, slow) [5] on the main axis of the monitor loudspeaker with the input short–circuited, at the reference distance of 1 m is:

$L_{noise} \leq 10$ dB(A)

2.8. Separate low–frequency loudspeakers

The overall requirement is that the locations of the separate bass loudspeakers must not be audibly apparent.

The use of separate bass loudspeakers affects the requirements for the main loudspeakers, as given in *Sections 2.1.* to *2.7.* of this *Appendix*. The requirements for the main loudspeakers will then only apply over the frequency range for which they are the principal reproducers.

To satisfy the requirement that the location(s) of the separate low–frequency loudspeakers is (are) not audibly perceptible as different to those of the rest of the audio spectrum, the crossover frequency must be below some limit. That limit will depend on the locations of the low–frequency loudspeakers relative to the other loudspeakers, the size of the room, the harmonic distortion produced by the low–frequency loudspeakers and the crossover filter order.

Lower crossover frequencies will result in more freedom in the choice of location but less benefit in the reduction in size of the main loudspeakers.

For example, for almost complete freedom in the choice of location in typical size of room, a bass loudspeaker specification would be:

– cross–over frequency	80 Hz
– out–of–band harmonic distortion levels	≤ –50 dB (0.3%)
– filter order	fifth

In contrast, if the separate low–frequency loudspeaker was located near to the wall, in between the main front–left and front right–loudspeakers, then a suitable bass loudspeaker specification would be:

– cross–over frequency 160Hz
– out–of–band harmonic distortion levels ≤ –34 dB (2%)
– filter order fourth

It will also be necessary to adjust the gain of the separate bass channel, using the test signal specified in *Section 2.4* and one–third octave band analyser, to achieve an overall uniform response at one or more of the listening positions. In general, it will not be possible to achieve uniform responses at all listening positions and the setting will have to be a compromise between several positions, depending on the number of listeners.

Appendix 4

Reference monitor headphones

In those cases when headphone monitoring is advisable, the frequency response of the monitor headphones should meet the requirements specified in ITU–R Recommendation BS.708 [8].

All other characteristics of the monitor headphones should meet the requirements specified in IEC publication 581, Part 10 [9].

Bibliography

[1] IEC Publication 225 (1966): *Octave, half–octave and third–octave band filters intended for the analysis of sounds and vibrations*

[2] ITU–R Recommendation BS.645: *Test signals and metering to be used on international sound programme connections*

[3] EBU Technical Recommendation R68: *Alignment level in digital audio production equipment and in digital audio recorders*

[4] EBU document Tech. 3282: *Alignment signals for digital coding levels and listening conditions – Handbook for the EBU R–DAT alignment tape*

[5] IEC Publication 651 (1979, 1993): *Sound level meters*

[6] IEC Publication 268: ***S**ound system equipment.* Part 1 (1985): *General*; Part 5 (1989): *Loudspeakers.*

[7] EBU Technical Recommendation R37: *The relative timing of the sound and vision components of a television signal*

[8] ITU–R Recommendation BS.708: *Determination of the electro–acoustical properties of studio monitor headphones*

[9] IEC Publication 581: *High fidelity audio equipment and systems; minimum performance requirements.* Part 10 (1986): *Headphones*

Note concerning IEC Publications:

As from 1997, all new IEC publications and parts, as well as new editions, revisions and amendments to existing publications, are being issued with a designation in the 60000 series. It is necessary to add 60000 to the existing base number.

Publications printed before 1997 will continue to carry the old series of numbers on printed copies until they are revised, but they will appear with the new 60000 numbers in the IEC Catalogue.

Editeur responsable : P.A. Laven, Ancienne Route 17A, CH–1218 Grand–Saconnex/Genève, Suisse.