# SURROUND SOUND FORUM

Eine Gemeinschaftsinitiative von VDT, IRT und SRT

RECOMMENDED PRACTICE SSF – 01.1- E-2002

# Listening Conditions and Reproduction Arrangements For Multichannel Stereophony

Schule für Rundfunktechnik

## PREFACE

This “Recommended Practice” was published by the “SURROUND SOUND FORUM”, an interdisciplinary and supraregional working group that was founded in 1996 at the 19th Tonmeistertagung (Convention of Sound Design of the Association of German Tonmeisters, VDT). This forum is also supported by the IRT, Institute of Broadcast Engineering, and the SRT, School of Broadcast Engineering, and is open for anyone interested. In conjunction with other committees and institutions as well as the relevant industry, the SSF wishes to coordinate the development of multichannel stereophony, to take control of further standardizations, and especially, to support a move for fair practice in this field as well as encouraging an exchange of information and experience.

## CONTENTS



This document replaces the former “Recommended Practice SSF-01/1-E-2 (1999). Herewith proposals of the member of the SSF are considered as well as results of discussions with the working group “Multichannel Audio” of the AES. Moreover, printing errors are amended and editorial improvements have been made. The amendments are in yellow color. This designation is removed after acceptance by the SSF at 2003-08-01.


Chairman:

Dr. Günther Theile, IRT
Floriansmühlstr. 60
80939 München
Tel.: 089 32399-324
Fax: 089 32399-351
E-mail: theile@irt.de

Vice-Chairmen:

Uwe Krämer, SRT
Wallensteinstr. 121
90431 Nürnberg
Tel. : 0911 9619-487
Fax: 0911 9619-199
E-mail: UKraemer@srt.de

SSF E-mail Adresse:
surround-sound@tonmeister.de

http: / / www.tonmeister.de

## Contents in detail



## 1. Introduction / Scope

This recommendation is intended for the purposes of studios[1] (broadcasting, television, film and mastering studios). It should also be valid, to a large extent, for consumer equipment.

The information here is based largely on the ITU-R –Recommendation BS 775-1 [1], which gives the international standard for a hierarchy of uniform discrete multichannel systems. Other literature used includes recommendation BS 1116-1 [13], and the technical document EBU 3276 [3], which give the basic requirements for listening conditions for the reproduction of program material from monophonic recordings to multichannel stereophony. From here details on the application of multichannel reproduction are taken, that must be considered in order to ensure “basic audio quality” (in accordance with [13])[2].

In the studio domain high quality listening conditions are required in

- *Reference listening rooms*, i.e. listening rooms for critical assessment of the quality of both transmission equipment set up and items of program material, from which a selection will be used in a sound or television broadcaster’s program output and for recordings, as well as in

- *Control rooms*, for the production and critical technical/artistic assessment of the sound quality of items of program material.

***Reference Listening Rooms*** must have optimal listening conditions and should comply with the following rules, in order to allow comparable results in different places and to guarantee a satisfactory program interchange.

In ***Control Rooms*** the necessary technical installations should affect the quality of the sound field at the reference listening position as little as possible. Control rooms also should comply to a large extent with the following rules.

[1] The so-called *studio quality* is defined sufficiently by the present recommendation as well as by such of international standardization bodies of broadcasting and television organizations, mastering studios etc. The term ‘studio’ should not be falsified by additions (as e.g. ‘high-end studio, studio quality).
[2] See also Appendix 1.

Because of differing conditions (regarding the acoustical properties of the living room) in *home reproduction*, deviations from the studio and reference conditions must be permitted. However, these should not impair the multichannel stereo effect. With today's technical standards, most electro acoustic parameters do not require significant differentiation between the studio and the home; for some parameters, values for home use are given which are expected to be sensible.

For home reproduction, the former Hi-Fi-Standard (DIN 45500) is no longer adequate and replaced by documents of IEC 29 *(EN 60268*). That standard gives minimum conditions for a higher quality category compared with the available mass productions. From there some relations to the here described requirements can be derived.

# 2. Listening Conditions

## 2.1 General Notes

The term 'listening conditions' defines the complex characteristics of a 'sound field', that a listener in a listening room at the reference listening position is subjected to by the room-related[3] reproduction of sound over loudspeakers.

**The overall listening conditions and the achievable quality of the sound field associated with them are determined by:**

- **the geometric and acoustical properties of the listening room,**
- **the properties and arrangement of the loudspeakers in the listening room**
- **the listening position, or rather the listening zone, for the selected listening places.**

The listening conditions in a high quality listening room, should allow neutral and critical monitoring of the sound signal, that is, the characteristics and deficiencies should be clearly recognizable, and the listening events, i.e. the subjectively perceived sound events, at the defined listening place should not be impaired.

Furthermore, the reproduction of a technically and artistically perfect, high quality sound signal should give a technically and aesthetically satisfactory impression, and also include a definite acoustical response of the listening room (reproduction room) in form of a 'natural listening perception'.

In order to retain the same listening experience with home reproduction that the producer intended at the recording, the quality of home conditions should come close to the technical quality of the studio control, or rather the reference listening conditions.

For the standardization of such listening conditions for the reproduction of sound signals, it is necessary to clearly describe and define the characteristics of the sound field generated by means of loudspeakers in a listening room at the reference position at the height of the listener' ears (in accordance with the rules, given 1988 by the ITU-R/CCIR working group 10-C-1).

From this it follows, that **under reference conditions the sound field can be optimally designed in such a way, that it can be regarded as an reference sound field, resulting from the interaction of optimally arranged loudspeakers with high quality characteristics in accordance with a 3/2 configuration[4] (see figure 1, page …) and a high quality acoustically designed listening ("reproduction home") room.**

Such a definition applies to the creation of a reference sound field which is similar to a large extent, in different listening places, with differently designed reproduction rooms, and different loudspeakers. At present, however, this reference sound field cannot be completely defined. For this reason it is presented indirectly and approximately through parameters for listening rooms and monitor loudspeakers that can be fixed.

The parameters, tolerances, and measurement instructions contained here as well as in international documents, for recommended arrangements for listening conditions are therefore only the **necessary minimum requirements** to ensure that the necessary high quality of program exchange can be

[3] For the reproduction with headphones, the preparation of a separate recommendation is planned (SSF 03/…)

[4] With bass loudspeaker on the LFE channel, the 3/2/1-format is designed also as 5.1 configuration.

achieved. But they are not yet adequate, to exactly describe optimal arrangements or to guarantee an adequate conformity between different listening rooms.

In a similar way, the 'virtual listening room' offers a new approach for the optimization of listening conditions with headphones[5] .

## 2.2 Parameters and Values for Listening Conditions

The following statements, tables and explanations, present the state of international knowledge of reference listening conditions. The requirements for home reproduction should bear a sensible relationship to these (See also Appendix 1).

With today's state of knowledge as well as the available measurement technology, the required characteristics for the desired similar effects for multichannel reproduction in operation can be achieved occasionally by different measures. For example, in accordance with section 3.3 the operational sound level curve can be achieved with one full-range loudspeaker per channel or with a combination of one or more separate bass speakers (subwoofers) and the remaining loudspeakers, for which the lower limiting frequency could be increased to about 80...100 Hz, which allows smaller cabinet volumes (at least with home equipment).

### 2.2.1 Requirements for the reference sound field at the recommended listening position in studio conditions, and advice for home replay.

The requirements are presented for reproduction with loudspeakers; when replay with headphones are used the listening room has practically no influence (For the present requirements for headphones with multichannel reproduction see section 2.2.4.)

*Table 1: Requirements for the reference sound field at the reference listening position (studio conditions) and advice for home reproduction*

| Parameter | Units / Conditions | Studio | Home Replay |
|---|---|---|---|
| **Direct Sound** | | | |
| Amplitude/Frequency resp | Free field propagation measurem | Tolerance borders see tabl ence monitor) | See table 2 |
| **Reflected Sound** | | | |
| Early reflections | 0 ms… 15 ms<br>(in the region 1 kHz to 8 kHz) | < -10dB relative to direct so | < -10dB relative to direct |
| Temporary diffusion of the ant sound field | Avoidance of significant anoma sound field. | no flutter echoes, no sound coloration etc. | as for studio conditions with linear decay |
| Reverberation time | Tm (s) = Nominal value in the 200Hz to<br>V = Volume of listening room;<br>$V_0$ = Reference room volume of 1 | ≈ 0.25* $(V / V_0)^{1/3}$<br>**Reverberation time decay and tolerance bor diagram in figure 1** | ≈ 0.2 to 0.4<br>Tolerance field, see Figure 1. |
| **Stationary sound field** | | | |
| Operational sound level cu | 50 Hz… 2 kHz<br>2 kHz… 16 kHz | ± 3 dB<br>+ 3 dB/ from –3 –6dB in accordance with field, see figure 2. | Tolerance field, see figure decrease according to ITU [13], see appendix 1) |
| Background noise | | < NR 10 and/or <GK 10, se | |
| Reference listening level ( defined measurement sign | Input signal:<br>Pink Noise, -18 dBFS (RMS) | 78 dBA (RMS slow)<br>(per channel) | 78 dBA<br>( per channel) |

[5] See Theile, G. et al.: Design and Applications of a Data-based Auralization System for Surround Sound. Paper presented at the 106th Convention of the AES, Munich, and Preprint no. 4976.

The **Direct sound** is the sound without influence from the listening room in the form of reflections and reverberation. The quality is determined by the characteristics of the appropriate loudspeakers, in accordance with 2.2.2 (see table 2)

The **Reflected sound (reverberation sound field)** is split into

- **Early reflections**, within the first 15ms, in the region of 1 kHz to 8 kHz. Requirements and measurements according to EBU [3].

- **Diffuse sound** of the reverberant sound field (linear decay).

- **Reverberation time**: Figure 1 shows the tolerance field for the reverberation time, in accordance with EBU [3]. The measurements are made with the used loudspeakers and with 1/3-octave band filtering. *Tm* is the arithmetic average of the measured reverberation time *T* in the 1/3-octave bands from 200 Hz to 4 kHz. It should lie between 0.2 and 0.4 seconds, in dependence on the room size (see table 1), in order to allow a 'natural' spatial perception.

  The frequency response for the reverberation time should be steady and continuous; sudden or strong breaks influence the operational sound level curve. Therefore such deviations in adjoining 1/3-octave bands in the region of 200 Hz to 8 kHz should not exceed ± 0.05s, and less than 200 Hz less than 25% of the longest reverb time should not be exceeded[6].

The **Stationary sound field** is represented by the

- **Operational sound level curve** (see also tolerance field in figure 2), in accordance with EBU [3]. It represents an important criterion for interaction between the room and the loudspeakers and for the quality of the listening conditions achieved. It is measured as the frequency response of the sound pressure level at the reference listening position. Measurement signals are band filtered Pink Noise. The tolerances should be checked for each loudspeaker separately. Consistency with the operational sound level curve is particularly important for the front loudspeakers (see also further advice in [3]).

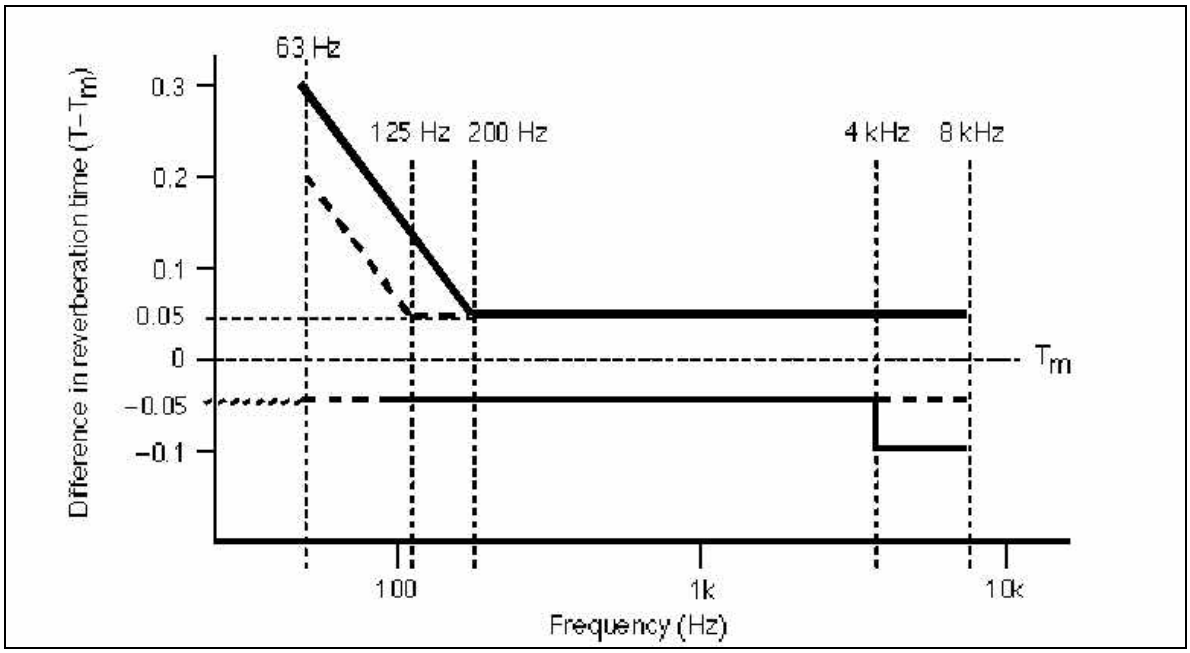



***Fig. 1 – Tolerance mask for the reverberation time, relative to the arithmetic average value $T_m$ [3].***
(The dotted line, the proposal to reduce the tolerances below 125 Hz is under consideration).

[6] see appendix 1 for the deviation to DIN 15996 [7]

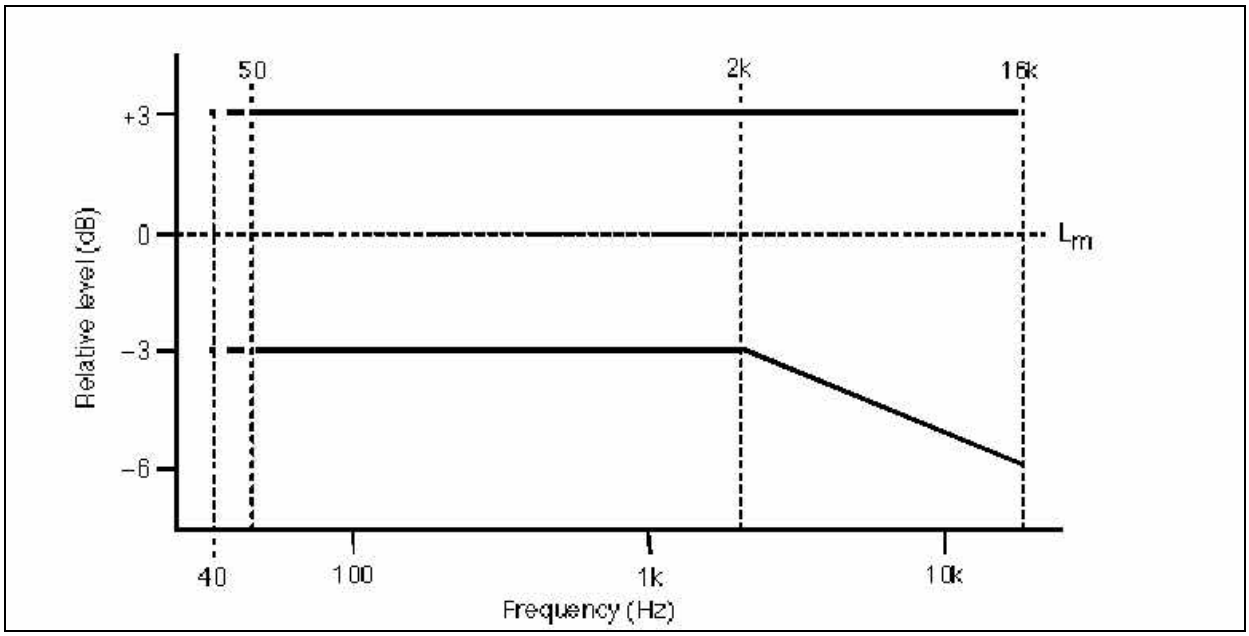



***Fig.2 – Tolerance limits of the operational room response curve, relative level [3][7].***

The region from 20 Hz to 30 Hz is yet to be completed.

**Reference listening level $L_{LIST\ ref}$ :**

It characterizes the sensitivity of the reproduction channels: The setting results with band limited pink noise with a digital level of –18 dBFS/RMS [25]. The reproduction channels are uniquely set so that at the reference listening position an A-weighted sound pressure level (RMS) of

$L_{LIST\ ref}$ = 85 – 10 log (n) dB(A) ,

Where *n* represents the number of reproduction channels.

For a 3/2 arrangement this is 78 dB (A) per channel; for 2/0 = two-channel stereo = 82 dB (A) is necessary. For individual cases see measurement instructions in SSF-02-10/98 [30].

The level difference between any two preferred channels should not exceed 0.5 dB (ITU [13] recommends ± 0.25 dB); this is particularly important for the front loudspeakers.

**Listening sound levels for the LFE channel**

For film reproduction it is usual to set the level of the LFE channel 10 dB higher for the emphasis of effect signals, in accordance with SMPTE. The measurement signals are available on the measurement and test recordings in accordance with 'Recommended Practice' SSF 002/98-10 [30] and the demo and test DVD.

[7] The diagrams in figures 1 and 2 are taken from EBU recommendation [3]. In documents from the ITU [13] and from DIN [7] greater tolerances or greater deviations are given regarding the reverberation times and the operational sound level curves (see appendix 1). For this recommendation the values of the EBU are given priority.

**Background noise**

The continuous noise level (from air conditioning or other external or internal sound sources) is given in form of **1/3-octave band sound pressure level** $L_{pFeq,\ T\text{-}30s}$ (RMS, slow), in accordance with DIN 45641, for the 1/3-octave band averaged frequencies from 50 Hz to 10 kHz by a table or a curve. The information for single values is not sufficient.

The recommended curves in accordance with DIN 15996/1996 [7] are derived from and in view of the internationally known ISO 'Noise Rating Curves'(NR) [27], corresponding to the real conditions, from sensible frequencies at a constant value. The curve GK 10 (respectively NR 10) should preferably not be exceeded; under no circumstances must the curve GK 15 (respectively NR 15) be exceeded, see fig. 3. AES Information Document AESTD 1001.1 [2c] prefers the international more accepted NR curves.

> **Beware**: In many international documents the NR curves are given as octave band averages (also in [4]). $NR10_{Oct}$ then gives a value at 1 kHz of 10 dB. Because otherwise room acoustic measurements are mostly carried out with 1/3-octave band averaged frequencies, the obtained values are on average 5 dB lower. The here recommended $GK_{1/3\text{-octave band}}$ curves given here are more appropriate for practice.

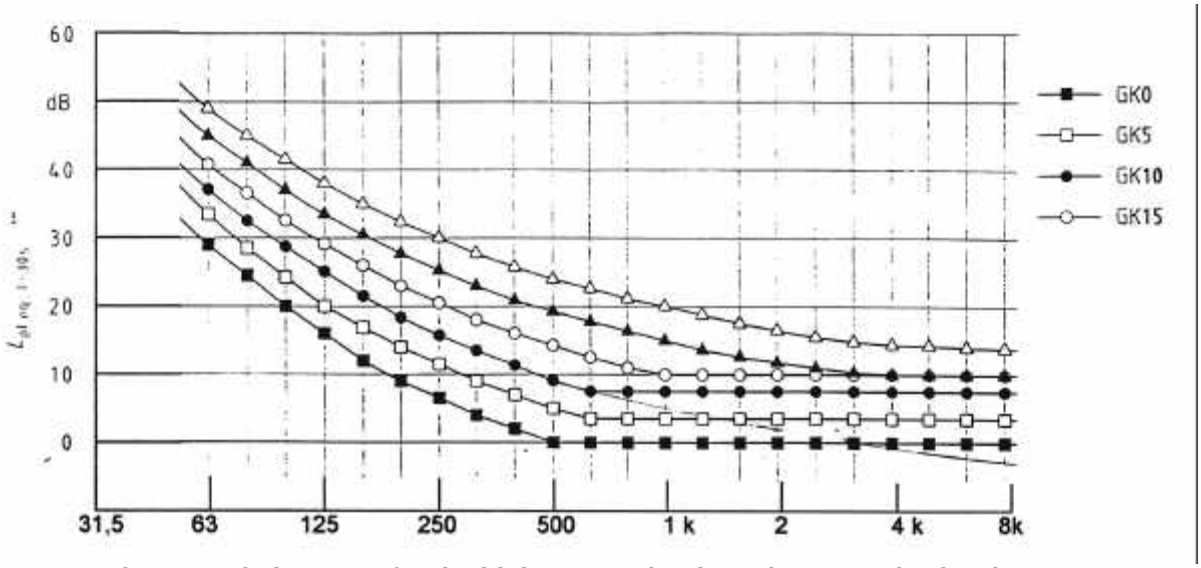



***Figure 3: Limit curves for the highest permitted continuous noise level***
*(1/3-octave band sound pressure level!) In accordance with DIN 15996 [7]*

GK0 corresponds up to and including 500 Hz to NR0; over 500 Hz the value is a constant 0 dB.
GK5 corresponds up to and including 630 Hz to NR5; over 630 Hz constant 3.5 dB.
GK10 corresponds up to and including 630 Hz to NR10; over 630 Hz constant 7.5 dB.
GK15 corresponds up to and including 1000 Hz to NR15; over 1000 Hz constant 10 dB.

### 2.2.2 Requirements for the reference monitor loudspeakers and advice for home loudspeakers.

The specifications in table 2 include the objective minimum conditions for a reference monitor loudspeaker. It must be mentioned that there are loudspeakers that comply with these regulations that are not necessarily suitable for all program genres as reference loudspeakers. To be able to fully perform these critical functions, the conclusive selection and decision is formed on the strength of investigative subjective tests and the resulting criteria and attributes.

For measurement conditions, known guidelines (i.e. [3]) are used to relate the measurement distances to the dimensions of the loudspeaker casing (usually dimensions are > 2m) etc. According to IEC 268-5 the result should be referred to the nominal distance of 1 m.

For electrical measurements the value should be guaranteed to be within ± 0.2 dB; for acoustic measurements the measurement error margin should be less than ± 1dB in the total frequency range.

The amplitude/frequency response is measured under free field conditions with pink noise for the 1/3-octave band averaged frequencies in the range 31.5 Hz to 16 kHz at 0°, ±10° and ±30°. The permitted tolerances and differences are given in the table. The correct characteristics should preferably be symmetrical around the reference axis.

The sound directivity index can also be derived from the 1/3-octave band measurements. It can either be calculated from the directional characteristics or derived from the difference between the free field measurements and the diffuse field measurements. The values according to ITU [13] of a directivity index of >6 dB with a steady slow increase towards higher frequencies seems no more to be sufficient. The so called omni directional radiators are totally unsuitable for the front loudspeakers. Amongst other things a diffuse radiation might be desirable for the surround speakers dependent on the program material, but at the time being the valid standards recommend only uniform (equal) loudspeakers for all the five channels.

Regarding the additional use of a separate bass loudspeaker (subwoofer), see section 3.3.

*Table 2: Requirements for reference monitor loudspeakers and advice for home loudspeakers.*

| Parameters | Units/ conditions | Studio | Home reproduction |
|---|---|---|---|
| **Amplitude/Frequency response**<br><br><br><br><br>Difference between front loudspea | 40 Hz…16 kHz<br>0°<br>±10°<br><br>horizontal ±30°<br><br>in the range > 250 Hz t | <br>Tolerance: 4dB<br>Deviation to 0°: 3dB<br>Deviation to 0°: 4 dB<br>0.5 dB | Tolerance: 4 d |
| **Directivity Index C** | 250 Hz… 10 kHz | 8dB ± 2 dB<br>(ITU [13]:> 6 dB) | 4 dB… 12 dB |
| **Non-linear distortion——a**<br>(SPL=96dB) | <br><100 Hz<br>>100 Hz | <br>-30 dB (= 3%)<br>-40 dB (= 1%) | <br><br>-40 dB |
| **Behaviour at Transients**<br>Decay time $t_s$ for reduction to a lev<br>i.e. 0.37 of the output level. | <br><br>$t_s$[s] | <br><br><5/f [Hz]<br>(preferable 2,5 f/f) | <br><br>preferable:<br>at 100Hz: ≤10ms |
| **Time delay**<br>Difference between stereo Loudsp | <br><br>δt [µs] | <br><br>≤ 10µs | <br><br>≤ 10µs |
| **Dynamic range**<br>Maximum Operating level<br><br><br>Noise level | <br>$L_{eff\ max}$<br><br><br>$L_{noise}$ | <br>>112 dB<br>(at IEC 268-1 Programme si<br>noise or specific measuring<br>≤10dBA | <br>>102 dB |

### 2.2.3. Requirements for the Reference listening rooms and home listening rooms

Table 3 contains the minimum requirements for smaller (about 60…80m³) and mean (about 80…150m³) listening rooms. A volume of 300m³ should not be exceeded for studio listening rooms. In order to obtain a sensible distribution of room modes, the dimensions according to the conditions in column 3 should be used.

The rooms should be largely symmetrical around the listening direction, and with regard to the distribution of the absorption material, especially around the speakers, doors, windows, and technical equipment, etc., so that any acoustical discontinuities can be avoided. The surface of the mixing desk should be designed to avoid disturbing reflections.

In the course of the further work it is intended to give more practical advices for living rooms.

*Table 3: Requirements for reference listening rooms and advice for living listening rooms.*

| Parameter | Units/ Conditio | Studio: according to EB | Home ~~Repla~~ Reproduction |
|---|---|---|---|
| ***Room size***<br>Floor surface<br>-mono/2-channel-stereo<br>-multichannel stereo | S [$m^2$] | <br>>30<br>>40 | <br>>25<br>>25 |
| ***Room Proportions*** | l = Length<br>w = Width<br>h = Height | 1,1w/h ≤ l/h ≤ 4,5w/h-4,<br>with<br>l/h <3 and w/h <3 | |
| ***Base width***<br>- 2-channel stereo<br>- multichannel- stereo | B [m] | <br>2.0…4.0<br>2.0…4.0 | <br>2.0…3.0<br>2.0…3.0 |
| ***Basis angle***<br>- 2-channel-stereo<br>- multichannel- stereo | [degrees]<br>referred to L/R | <br>60<br>60 | <br>60<br>60 |
| ***Listening distance***<br>- 2-channel stereo<br>- multichannel stereo | D [m] | <br>2m…1,7 [8] B<br>B (± 0.8m) | <br>2m…1,7 [8] B<br>B (± 0.8m) |
| ***Listening zone***<br>- 2-channel stereo<br>- multichannel stereo | R (radius) [m] | <br>0.8<br>0.8 | <br>0.8<br>0.8 |
| ***Loudspeaker Height*** [8]<br>- 2-channel stereo<br>- multichannel stereo<br>L,C,R,LS,RS | h [m] | <br>≈ 1.2<br>≈ 1.2 | (see also Fig.4)<br>>0.9…1,4<br>>0.9…1,4 |
| ***Distance to surrounding refl surfaces***<br>- 2-channel stereo<br>- multichannel stereo | d [m] | <br><br>≤ 1<br>≤ 1 | <br><br>≤ 0,5<br>≤ 0,5<br>(better: >1) |

[8] that is from the acoustic centre

### 2.2.4. Requirements for reference monitoring headphones

When listening on headphones, generally the listening room will be excluded. Different evaluations will be achieved when the multichannel recording was made with 3/2 reproduction format for loudspeakers.

Adequate examination results are not available. It is currently not possible in this document to define the requirements for reference monitor headphones for monitoring multichannel sound recordings. It is intended that a separate 'recommended practice SSF 03/…' will be made.

Principally the requirements should correspond to the diffuse field frequency response for monitoring headphones in recommendations ITU-R BS.708 [10].

A new approach for the optimization of listening conditions is given by 'virtual listening rooms' [9].

## 3. Reproduction configurations: advice on installation and use

### 3.1 Reference configuration

When fulfilling adequate listening conditions according to section 2 as well as the tables 1 through 3, a reproduction in accordance with standards is guaranteed as defined by tables 1 to 3 and figure 4, guaranteeing setup, arrangements, and settings of loudspeakers in the 3/2 format and other recommended compatible formats.

For clear differentiation and definition, the number of source signals should be used as the basis for the format code. In most cases it is identical to the number of channels; however it is not always identical to the number of connected loudspeakers. This can sometimes be greater.

The **reference arrangement** (Basic reproduction configuration) has the **3/2 format** with **3 front signals / channels L** = left, **C** = Center, **R** = Right plus **2** so-called **surround channels** (Room and ambience channels) **LS and RS.** This reference configuration offers as the world standard a number of speakers which can reasonably be expected in the studio as well as in the home, to give a compromise between the acoustical requirements of the process and the practical realizable conditions.

For setting up the five loudspeakers, an arrangement in accordance with **figure 4**, on the basis of the recommendations in ITU-R BS. 775-1 [1], and SMPTE [2] is recommended. This **principle reproduction standard** is totally independent on the applied transmission system and recording processes, and should not be confused with the different **coding formats** which are only partially standardized, as for example, ISO/MPEG. See appendix 2.

***Reference Color code***

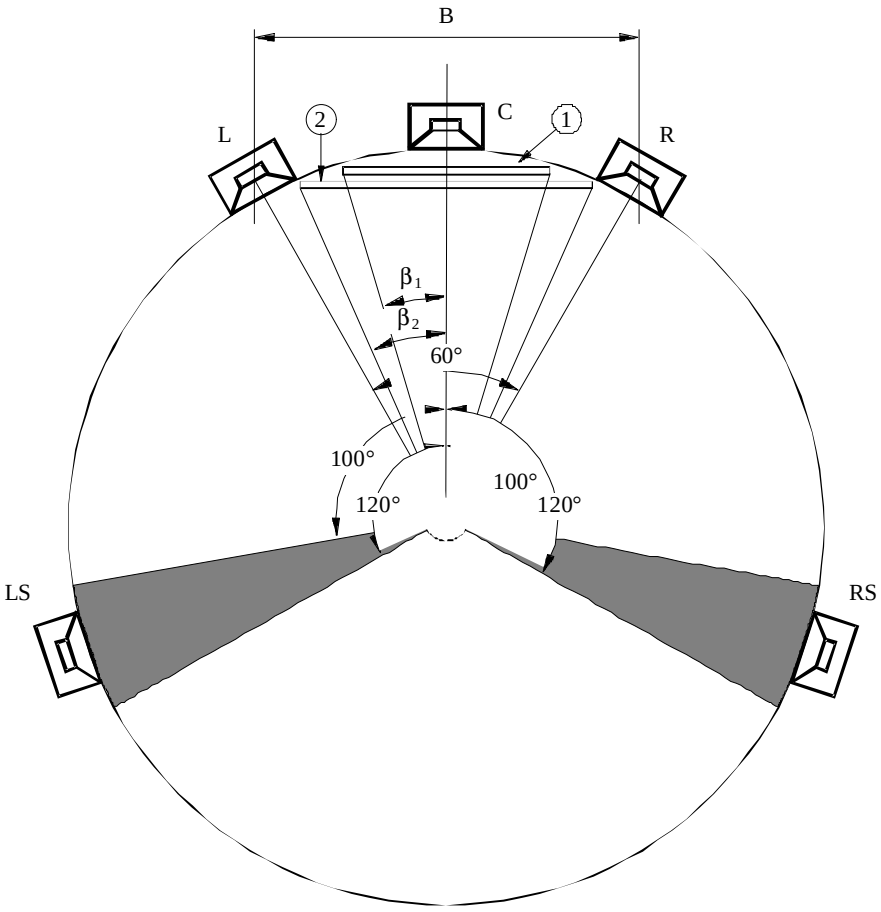
For avoiding any interchange between program signals and/or channels with the connection of the loudspeakers of multichannel equipment with cables and inputs and/or outputs the use of a color code (according to *Camerer/Hoeg*) is of large benefit, as it recommended for the designation and allocation of 8-channel recording formats [30]:

| ***Channel*** | ***Signal*** | ***Direction*** | ***Farbe*** |
|---|---|---|---|
| 1 | L | Left | Yellow |
| 2 | R | Right | Red |
| 3 | C | Center | Orange |
| 4 | Optional | LFE and/or free available | Grey |
| 5 | LS | Left Surround Signal | Blue |
| 6 | RS | Right Surround Signal | Green |
| 7 | Free available | Optional, e.g. left with 2/0 | Violett |
| 8 | Free available | Optional, e.g. right with 2/0 | Brown |

[9] see Theile. G.: Virtual Loudspeakers in a virtual room. Production Partner Special at the 20. Tonmeistertagung, 1998, Karlsruhe.

**FIGURE 4**

**Reference loudspeaker arrangement with loudspeakers L/C/R and LS/RS**




Screen 1 HDTV – Reference distance = 3 $H$ ($2\beta_1 = 33°$

Screen 2 = 2 $H$ ($2\beta_2 = 48°$

H: height of screen

B: loudspeaker base width

| Loudspeaker | Horizontal angle from centre (degrees) | Height (m) | Inclination (degrees) |
|---|---|---|---|
| C | 0 | 1.2 | 0 |
| L, R | 30 | 1.2 | 0 |
| LS, RS | 100 ... 120 | ≥ 1.2 | 0 ... 15 down |

D01

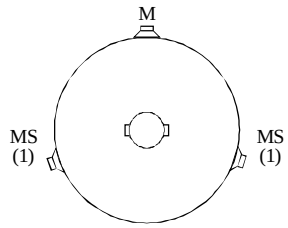
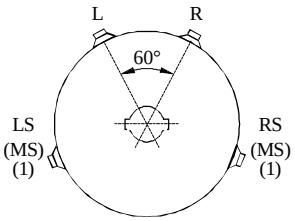
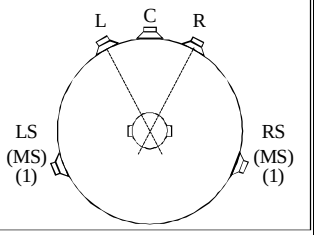
**Figure 4: Reference loudspeaker arrangement** with the loudspeakers, L/C/R and LS/RS in combination with picture reproduction installation (in accordance with ITU-R BS. 775-1 [1]; Figure 1).

For the exceptional reproduction of surround signals also positions outside the chosen radius can be used.
Regarding the deviations within the circle line see section 4.1.1.

### 3.2 Hierarchy of compatible multichannel sound systems for broadcasting and recording

The 3/2 Stereo-Surround System is plausibly embedded within a hierarchy of multichannel sound formats, which ensures the precondition of an uniform, flexible and compatible application with Television, Sound Broadcasting, Motion Pictures, Multimedia as well as with tape/disc video and sound carriers . For such a hierarchy, down-compatible up to the monophonic format, as shown in **Table 4**, appropriate matrixing conditions are valid, according to [1], for the addition of partial signals at the transmitting and/or reproduction side, depending on the reproduction situation; especially, if only individual signals have to be derived on the multi-signal bundles (e.g., the 2/0 format, the conventional two-channel stereophony).

*Table 4: Hierarchy of compatible multichannel sound systems of Broadcasting, Motion Pictures and Recording (according to [1]).*

| System | Channels | Code | Loudspeaker arrangement |
|---|---|---|---|
| Mono channel system | M | 1/0 |   |
| Mono plus mono surround | M/MS | 1/1 |   |
| Two-channel stereo | L/R | 2/0 |   |
| Two-channel stereo plus 1 surround | L/R/MS | 2/1 | |
| Two-channel stereo plus 2 surround | L/R/LS/RS | 2/2 | |
| Three-channel stereo | L/C/R | 3/0 |   |
| Three-channel stereo plus 1 surrond | L/C/R/MS | 3/1 | |
| Three-channel stereo plus 2 surround | L/C/R/LS/RS | 3/2 | |

(1) In the case of mono surround the signal feeding LS and RS should preferably be decorrelated.

D03

Also the current **3/1-matrix formats** (three frontal signals plus one surround signal) are integrated in the hierarchy and may reproduced with the 3/2 configuration, because the monophonic surround signal provides two loudspeakers, in which case the gain of the surround channels has to be reduced to 3 dB. (See in Appendix 2). It was the opinion of Japan to accept the 3/1-format within the ITU-Standard as an exception because it is used in Japan with the MUSE transmission system.

Moreover, also higher format systems above the 3/2 format are possible and can be matched – e.g., 5/2, 5/4 etc. These formats are not included within the ITU standard because they seem to be unrealistic for home conditions. On the other side, the format with 5 frontal speakers is used for longer time in the film domain (for which SMPTE is competent) and can be used also for the DVD. But nowadays it should be produced in such a way that it is also down compatible up to the 3/2 and 2/0 formats.

For all the other possible format combinations the reference configuration has to be the basis. Further loudspeakers can be attached to the reference configuration, but with the same number of source signals, to enlarge the effect of enveloping and/or to render discrete sound sources at a higher number of positions locations in such a manner that always up and down compatibility can be ensured.

## 3.3 Low-Frequency Extension

In order to avoid confusion a clear distinction is made here between a low-frequency-extension (LFE) signal that can be carried over a separate LFE channel in a transmission or recording system, and the separate radiation of low-frequency program content through so-called subwoofers. Although these may seem to be one and the same, they need not be. Indeed, it is this confusion about low-frequency management that causes a large number of problems in practical situations.

### 3.3.1 Additional bass channel: LFE signal and LFE channel = Low Frequency Extension)[10]

In the film domain the use of a special effect channel was introduced in the bass range from 20 up to about 80...120 Hz, as a 'Low Frequency Extension', LFE, which – according to ITU-Recommendation BS 775-1 [1] – **can be used optionally** as an supplement of the formats in studio and/or home. The designation is abbreviated with "0.1" respectively "./1", because of the small frequency range used. Therefore the nominations 3/2/1 or 5.1 and 5/2/1 or 7.1 are common.

According to [1], **optionally**, one additional channel for the enhanced bass region is permitted, with a frequency range of 20-80 Hz (up to 120 Hz maximum), which is the norm in cinemas with motion pictures. In **consumer audio systems, the LFE channel is also considered optional in reproduction**. Media should be prepared that conform to this recommendation – according to [2] - that they sound satisfactory even if the LFE channel is not reproduced.

Note: With the DVD-Audio the optional channel 4 can be used with full bandwidth, so that this signal can be used also for program signals, e.g. surround back (SB), instead of LFE, because a LFE signal is not necessary for high-quality music programs. But not any international agreement exists up to now.

EBU and SMPTE documents on multichannel sound [2a], [2b] contain sime remarks on the use of the LFE channel. This is from the SMPTE document [2a]:

> "When an audio program originally produced a as feature film for theatrical release is transferred to consumer media, the LFE channel is often derived from the dedicated theatrical subwoofer channel. In the cinema, the dedicated subwoofer channel is always reproduced, and thus film mixes may use the subwoofer channel to convey important low frequency program content. When transferring programs originally produced for the cinema over the television media [e.g. DVD] it may be necessary to re-mix some of the content of the subwoofer channel into the main full bandwidth channels. It is important that any low frequency audio which is very significant to the integrity of the program content is not placed into the LFE channel. The LFE channel should be reserved for extreme low frequency, and for very high level <120Hz program content which, if not reproduced, will not compromise the artistic integrity of the program".

With cinema reproduction the sound level usually is about 10 dB higher than the reproduction level of the other individual channels. According to SMPTE [2a] this will be compensated by a level increase of the reproduction channel, not by an increased recording level. This has to be observed in the studio domain and also with home reproduction, for reasons of compatibility. (It does not mean that the broad-band or weighted sound pressure level of the LFE loudspeaker should measure 10 dB higher than of any of the other channels when aligned using broad-band pink noise – in fact it will be considerably less than this as its bandwidth is narrower).

### 3.3.2. Separate low frequency loudspeakers (so-called sub-woofers) within the standard configuration

It may be useful, in addition to the main loudspeakers (L/C/R/LS/RS), to use separate bass radiators (Subwoofers) for the extension of the lower frequency range, so that the limit frequency of the 5 main speakers can be raised up to about 80 Hz and their volumes can be consequently reduced.

[10] In some international documents also other designations as 'Low Frequency Enhancement' or 'Low Frequency Effect' are usual.

In this case it is possible to use several subwoofers for intended individual channels (e.g. frontal and/or surround channels), or one single subwoofer supplements this low frequency range of all the five main speakers. All the loudspeakers are connected via crossover circuits (limit frequencies of 80 up to 160 are usual by the consumer industry; more efficient is in the vicinity of 80 Hz).

The configuration has to be regarded further on as a 3/2 format, but it is possible that such a separate bass equipment can be configured so that both 5.1 channel motion pictures and 3/2-format material without a separate LFE channel can be handled, according to **Figure 5.**

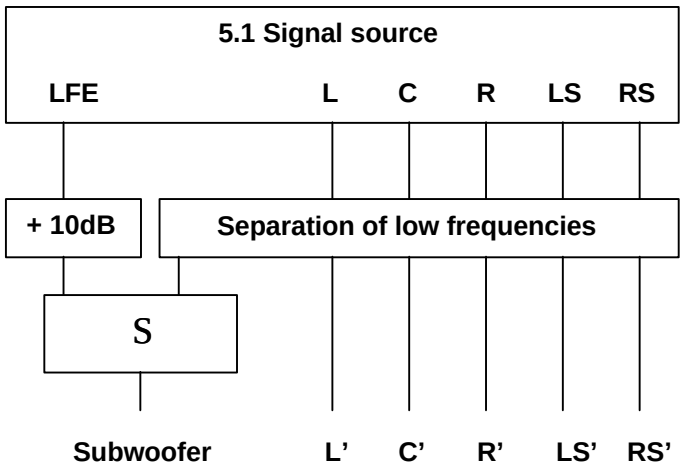



***Figure 5: Derivation of combined subwoofer and LFE signals***

The resulting quality, including the operational sound level response, is also dependent on the position of the relevant loudspeakers in relation to the listening position as well as on the non-linear distortions of the subwoofers, by which localization errors can occur (see more details in [3], [4]).

On the optimum arrangement of a single or several subwoofers within the listening room evidently only a small agreement of the opinions exist, but measurements are published, see in the documents of the EBU [3], AES [2c], and other publications as [40], [47], [48}, [49].

### 3.3.3. Enhancement of the listening zone by additional surround loudspeakers

For getting a larger listening zone and/or an improved enveloping by means of the reproduced room information with the 3/2 format, it is permitted, according to [1], to add a higher number of surround loudspeakers to both of the standard channels LS/RS. For larger reproduction rooms (as cinemas) it is necessary and usual anyway. In this case a sufficient de-correlation of the added loudspeaker channels has to be provided, e.g. by appropriate delay, and connected via special signal distributors (matrixes) or processors [41], [50], [51]. The result can be regarded as a pseudo 3/4-format.

# 4. Comments to alignment and measurement

## 4.1 Corrections of deviations of the positions of loudspeakers from the reference arrangement by delay

If it is not possible to install the frontal speakers on the recommended circular line, the appropriate delay differences have to be compensated by the equipment with relation to the reference position. Only for the exceptional reproduction of surround signals for LS and RS speakers positions may be chosen outside the circular line.

### 4.1.1. Reference listening rooms

Here sometimes the position of the center speaker is within the circular line because of the use of a larger high-perforated screen with projection in a large format, so that all the frontal speakers are arranged on a straight line.

In this case it is valid for the reference position:

With the listening angle of 60°, it is basis width B = reference radius; therefore the listening distance h (sometimes also named as D) = $\sqrt{3/4}$ x B = 0.87 B.
Therefore the distance to be compensated up to the circular line is 0.13 B.
So follows the delay time difference $\Delta t$ = 0.13 x B x 3 ms = **B/m x 0.39 ms.**

In general a level compensating is not necessary, if it is < 1 dB. Usually during the recording it will be observed that the center speaker C in any case cannot be dominant against R and L.

### 4.1.2. Home listening rooms

When optimizing the arrangement under home conditions and all the five loudspeakers have deviations to the hypothetic circular line, with relation to the center speaker all the other speakers may need additional and differentiated level and delay compensations.
The necessary corrections should be measured and aligned automatically by measuring impulses with appropriate so-called 'home processors' [46].

### 4.1.3. Larger multichannel presentation rooms (TV viewing sites, motion pictures theaters)

Also for TV Viewing sites the correction equation is valid according to section 4.1.1.
With theaters and mixing studios it may be not usual to use such delay compensation as described above; possibly here may exist a discrepancy against the TV viewing sites and the intended reproduction in cinemas of correctly recorded TV multichannel sound material. Further studies are necessary. Moreover, for such rooms the corresponding SMPTE regulations are valid; also the proprietary THX regulations (designed by Tomlinson Holman and/or Lucas film) are usual in film theaters.

Film sound track reproduction in living rooms with such delay compensated center speaker seems to be not problematically. Because of the fact that film dialog always occurs in the center, no detectable loss of presence may be expected (see also [41]).

## 4.2 Alignment of the reference listening sound level

The alignment of the reference listening sound level according to section 2.2.1. – i.e. 78 dBAS for the individual channel, therefore 85 dBAS for the 3/2 configuration – is made by alignment and test recording, see 'Recommended Practice SSF-02' [30]. An arrangement, measured in such manner, can be used also for the reproduction of 3/1 format recordings.

## 5. References - Standards

(Further relevant literature see Appendix 1, section 2 ).

**Multichannel Formats, Configurations:**

[1] ITU: Recommendation ITU-R BS. 775-1: Multichannel Stereophonic Sound System with and without accompanying Picture (Geneva, 1992-1994).

[1a] EBU-Recommendation R96-1999: Signal formats for production and delivery of multichannel audio programs. EBU Official Technical Texts, Geneva, 1999

[2] SMPTE Recommended Practice: Loudspeaker Placements for Audio Monitoring in High Definition Electronic Production (SMPTE N 15.04/152-300B) (December 1991); SMPTE RP-173.

[2a] Proposed SMPTE Standard for Television: Channel Assignments and Levels on Multichannel Audio Media. ITU Information document ITU-R. -10C/11 and 10-11R/24; 16 March 1998 (E).

[2b] EBU Document BPN 021-1999: Multichannel Audio: Report on different reproduction and delivery formats available on the market.

[2c] AES Information Document AESTD1001.1.01.10 : Multichannel Surround Sound Systems and Operations, March 2002, New York.

**Listening Conditions, in general, Sound field:**

[3] EBU-Recommendation: Listening conditions for the assessment of sound program material. EBU Technical Recommendation R 22 – 1998.
EBU document Tech 3276 - 1998 (second edition): Listening conditions for the assessment of sound program material: monophonic and two-channel stereophonic).
EBU document Tech 3276 – 1999: Supplement 1: Multichannel Sound.

As well as in: [7, 13,]

**Listening Rooms:**

[4] EBU-Recommendation: Listening conditions for the assessment of sound program material. EBU Technical Recommendation R 22–1998.
EBU document Tech 3276-1998 (second edition): Listening conditions for the assessment of sound program material: monophonic and two-channel stereophonic).
EBU document Tech 3276 – 1999: Supplement 1: Multichannel Sound.

[5] OIRT Recommendation 86/3: Technical Parameters of OIRT Reference Listening Rooms. Doc. TK-31-58. (Prague, 1990).

[6] Nordic Recommendation N 12A: Sound Control Rooms and Listening Rooms.
Technical Recommendation from the Nordic Broadcasting Corporations (1992).

[7] Deutsche Norm/German Standard: DIN 15996, April 1996: Elektronische Laufbild- und Tonbearbeitung in Film-, Video- und Rundfunkbetrieben. Anforderungen an den Arbeitsplatz (DIN 15996:1996-04). [Electronic moving picture and sound processing in film, video and broadcasting organizations. Requirements to the working location].

As well as in: [13]

**Monitor Loudspeakers:**

[8] OIRT Recommendation 55/2: Technical Parameters for Studio Monitor Loudspeakers.
Doc. TK-2-2092. (Prague, 1990).

[9] Nordic Recommendation N 12B: Studio Monitoring Loudspeakers.
Technical Recommendation from the Nordic Broadcasting Corporations (1992).

As well as in: [3, 4, 13]

**Monitor Headphones:**

[10] ITU-R Recommendation BS.708: Determination of the electro-acoustical properties of studio monitors headphones.

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

As well as in: [3].

**Subjective Assessments:**

[12] ITU-Recommendation ITU-R BS.562-3: Subjective Assessment of Sound Quality (Geneva, 1978-1990) – replaced by new recommendation BS.1284 of 12.12.96: see [14, 15].
[13] ITU-Recommendation ITU-R BS.1116-1: Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems (Geneva, Supplement 1 to Volume 1997).
[14] ITU-Recommendation ITU-R BS .1283: Subjective Assessment of Sound Quality - A Guide to existing Recommendations (Geneva, Supplement 1 to Volume 1997)
[15] ITU-Recommendation ITU-R BS. 1284: Methods for the Subjective Assessment of Sound Quality - General Requirements (Geneva, Supplement 1 to Volume 1997).
[16] ITU-Recommendation ITU-R BS. 1285: Preselection Methods for the Subjective Assessment of small impairments in Audio Systems (Geneva, Supplement 1 to Volume 1997).
[17] ITU-Recommendation ITU-R BS.1286: Methods for the Subjective Assessment of Audio Systems with accompanying Pictures (Geneva, Supplement 1 to Volume 1997).
[18] EBU Technical Recommendation R 90-1998: The subjective evaluation of the quality of sound program material. (See here: EBU document Tech.3286-1997: Assessment Methods for the Subjective Evaluation of the Quality of Sound Programs - Music. EBU Official Technical Texts (Geneva, 1997).
[19] OIRT Recommendation 63/3: Formation of Listening Groups and their Working Methods. Doc. TK-31-63 (Prague, 1992).
[20] OIRT Recommendation 91/2: Subjective Assessment of Sound Recordings. Doc. TK-31-64/1 (Prague, 1992).
[21] ISO/MPEG: Report on the MPEG/Audio Multichannel Formal Subjective Listening Tests. ISO/IEC JTC1/SC2/WG11 N0685 MPEG 94/063 (1994).

**Audio Quality Parameters:**

[22] ITU-Recommendation ITU-R BS. 644-1: Audio Quality Parameters for the Performance of a High-Quality Sound-Program Transmission Chain (Geneva, 1986-1990).

**Test signals, measurement technologies:**

[23] ITU-Recommendation ITU-R BS.468-4: Measurement of Audio-Frequency Noise Voltage Level in Sound Broadcasting (Geneva, 1986).
[24] ITU-Recommendation ITU-R BS.645-2: Test Signals and Metering to be used on International Sound Program Connections (Geneva, 1986-1992).
[25] EBU-Recommendation: Alignment Level in digital Audio Production Equipment and in digital Audio Recorders. EBU Technical Recommendation R 68 - 1992.
[26] EBU document Tech. 3282: Alignment signals for digital coding levels and listening conditions – Handbook for the EBU R-DAT alignment tape.
[27] ISO Recommendation 1996 (1972): One-third octave band background noise level limits noise Rating curves (NR).
[27a] ITU-Recommendation ITU-R BS.1387: Method for objective measurements of perceived audio quality (PEAQ). (Geneva, 2000)

**Recording parameters:**

[28] ITU-Recommendation ITU-R BR. 1384: Parameters for international exchange of Multi-Channel Sound Recordings. (Geneva, 1998).

[29a] EBU-Recommendation: Exchange of Sound Programs as digital Tape Recordings. Technical Recommendation R 64 - 1993.

[29b] EBU Technical Recommendation R 91 - 1998: Track allocations and recording levels for the exchange of multichannel audio signals (Geneva, 1998).

[30] Surround Sound Forum: Parameter für Mehrkanal-Tonaufzeichnungen im 3/2-Format (Parameter für Programmaustausch und Archivierung, Einstellung von Wiedergabeanlagen. SSF-2/10/98 (Edition October 1998).
English version: SSF-2.1- E - 2002: Multichannel Recording in 3/2 Format (Parameters for Program Exchange and Archiving, Alignment of Reproduction Equipment; Edition July 2002).

**Picture / Sound allocation:**

[31] EBU Technical Recommendation R 37: The relative timing of the sound and vision components of a television signal.

[32] Surround Sound Forum: SSF Information Document SSF – 05.1-2002: Multichannel Stereophony: Relations of Viewing and Listening Arrangement (in preparation).

**Appendix 1 to Doc. SSF – 01.1- E-2002**

**Remark 1: On the parameters and values of section 2 of the current „Recommendation of Practice“of the SSF.**

The elaboration and determination of listening conditions as well as of reproduction arrangements for the subjective assessment of sound programs with broadcasting and television is of a long tradition. The first international listening groups were established, beginning with 1960, within the frame of the OIRT and then standards for listening rooms and monitor loudspeaker were established. These methods very soon were overtaken by the Nordics Broadcasting Organizations. The last results were demonstrated in the recommendations [5, 6, 8, 9, 19, 20].

Later such investigations have been done also with the CCIR (nowadays ITU-R) [12-17], with ISO/MPEG [21] as well as with the EBU [18], especially since the introduction of perceptual coding for bit-rate-reduction for transmission and recording which needed intensive listening tests with higher requirements of listening conditions beside the non-sufficient objective measuring methods. In the last time on the basis of such investigations the expert group 10/4 of ITU could develop with the Recommendation “PEAK”, ITU-R BS.1387, a combined measurement tool for the objective, perceptual based evaluation of compressed speech and audio signals [27a].

Beside this a lot of new findings have been made, which resulted in actual recommendations of listening conditions, especially of listening rooms and the employed reference monitor loudspeakers [3,13]. The successful realization of new coding technologies gave the possibility also to introduce the discrete multichannel stereophony with broadcasting, television and optical recording and to make it handling for the user.

Therefore it seemed to be useful, to summarize in this ‘Recommendation of Practice’ of the SSF in your hand the essentials of the documents of international standardizations bodies, in particular the parameters and values of listening conditions. Because of the tolerances of the necessary reference conditions used up to now have been varying world-wide, now the latest findings could be considered which were elaborated by EBU [3]. In particular, there the tolerances of reverberation time (figure 1) and operating sound level response (figure 2) now require smaller tolerances compared with the modified version of Recommendation ITU-R BS.1116-1[13], published a short time before. But experiences gained with newer realizations of reference listening rooms underline the necessity to get more stringent requirements. This can be derived also from the fundamental remarks in section 2.1., that it is possible to reach the definition of the target of the “reference sound field” only during the course of permanently gained experiences. On the basis of further considerations within the AES in the course of the preparation of the common information document AES TD1001 „Multichannel Surround Sound Systems and Operations“ it was possible to agree on stronger values with the studio monitors, which now were included in that reediting of July 2002, according to [2c].

In spite of manifold efforts for the realization of ‘ideal’ listening rooms and the use of ‘uniform’ monitor loudspeakers in studios as well as in the home, it has been proved to be unrealistic. Finally, only the sufficient defined sound field can be the precondition to create similar ‘listening events’. But because of non-sufficient cognition of the optimum sound field - with relation to subjective criteria and the relations to objective parameters and measuring methods – and because of the complex and not yet explained context between the required sound field measures and the physical properties of rooms and monitor loudspeakers, this reference sound field at the time being cannot be defined completely and realized. Moreover, the definition of that can be made only for one reference listening position respectively for a few listening positions and not yet for a larger listening zone of chosen listening places.

At the present time it is only possible to approximate iteratively to the mentioned target ‘reference sound field’ with every new approach, by means of known sound field measures (quantities?) as well as by reproducible properties of high-quality loudspeakers and listening rooms. In any case careful subjective assessments are necessary in particular with given methods [13 – 18] because the usual known objective measuring methods cannot ensure sufficient quality and/or conformity.

It is well known that the influence of the listening room can be reduced or suppressed in general with head-related reproduction. Results of evaluations of test recordings can differ between the perception via loudspeakers or via headphones. Here special marginal conditions have to be observed (it is intended to prepare a special ‘Recommendation of Practice’ of the SSF – [SSF 03]).

Finally it should be noted that the threshold curves gained from DIN 15996 [7] seem to be more realistic than the noise rating curves by ISO [27], but they are not so well-known, therefore they were not referenced in the AES Information Document (see above, [2c]).
With raising requirements of reference conditions determined for formal subjective tests it might be that also curve GK 10 have to be remain under. On the other side, for this recommendation of practice the strong limit deviations of <±10% by DIN [7] could not be overtaken as not yet secured.

Altogether, the intention was to present with the recommendation in your hand also a contribution for further discussions with other organizations, as it is the AES. In several organizations the published standards might be determined by a strong influence of the industry (e.g. in the AES-17, in which the measurement of audio-frequency noise voltage level is defined just in open contrast to the international recommendation ITU-R BS. 468, see remark in Appendix 3).

## 2. Further bibliography:

**To section 3.3:**


[40] ZACHAROV, Nick; BECH, Søren; MEARES; David: The Use of Subwoofers in the Context of Surround Sound Program Reproduction; paper presented at the 102nd AES-Convention, Munich, 1997, March 22-25. Preprint Nr. 4411.

[41] STEINKE, G.: Vom Referenz-Hörraum zur Tonsystemlösung für das Filmtheater der Zukunft. (From the reference listening room to the sound system approach of the film theater of the future).In: Fernseh- und Kino-Technik, Berlin, 48 (1994), H.11, S.583-592; H.12, S.667-678.

[42] KÖNIG, Florian M.: Beschreibung von Klangfarbenveränderungen beim 3-D-räumlich varianten Hören mittels bestimmter Kopfhörer-Beschallungsverfahren (Description of sound colour alterations with the 3-D spatial variant listening by means of particular headphone methods). Paper presented at the DAGA, 1997 (there see further references).

[43] THEILE, G.: Zur Theorie der optimalen Wiedergabe stereofoner Signale über Lautsprecher und Kopfhörer (On the optimum reproduction of stereophonic signals via loudspeakers and headphones). In: RTM, 25 (1981), September, S.32-47 (Special issue).

[44] HOEG, W., CHRISTENSEN, L., WALKER, L.: Subjective assessment of audio quality: Means and methods in the EBU. EBU Technical Review, No. 274 – winter 1997, pp. 40 – 48 (Geneva, 1998).

[45] EBU document Tech. 3287-1998: Parameters for the Subjective Evaluation of the Quality of Sound Programme Material – Music (PEQS); with Compact Disc (Geneva, 1998).

[46] SCHUBERT, P. Die Wahrnehmbarkeit von Rückwürfen bei Musik (On the perception of reflections with music)., Z. Hochfrequenztechnik u. Elektroakustik, 78 (1969), 230-245, Tech Mitt. d. FZ 10 (1966), 3, 124 – 127.

[47] T. Nousaine, Multiple subwoofers for home theatre. Paper presented at the AES 103th Convention, J. Audio Eng. Soc. (Abstracts), 46 (1998), 4, 276 – 287.

[48] KÜGLER, C., THEILE, G.: Loudspeaker reproduction: study on the Subwoofer concept. Paper presented at the AES 92nd Convention, J. Audio Soc. (Abstracts), 40 (1992), 5, 437, Preprint 3335.

[49] GRIESINGER, D.: Spatial impression and envelopment in small rooms. Paper presented at the AES 103rd Convention, J. Audio Eng. Soc. (Abstracts), 45 (1997), 11, 1013 – 1014, Preprint 4638.

[5o] FELS; P., STEINKE, G., WÜSTENHAGEN; U.: Method and Apparatus for Multichannel Sound Reproduction. Paper presented at the 100th AES-Convention, Copenhagen, May 1996. Preprint No: 4272.

[51] STEINKE, G.: Quality Monitoring of Multichannel Signals for Surround Sound - An Overview. Paper presented at the IBC'99, Amsterdam, September 1999 (www.ibs.org.uk/November'99)


(To be continued, see also the collection of references 'Multichannel Stereophony' by the Surround Sound Forum)

**Appendix 2 to Doc. SSF – 01.1 – E - 2002**

# On the discrimination of reproduction formats and coding formats

At present with multichannel sound engineering different coding formats and methods are on the market, in particular for moving pictures theaters. They are not standardized mostly but introduced as proprietary formats. Therefore the impression is given that these formats are system solutions of recording and reproduction formats, as it is presented in Section 3 of this recommendation in your hand, on the basis of international standards, according to ITU-R BS.775-1. But this is not correct.

For shorter connections, e.g. between recording rooms and control rooms, in each case are discrete, transparent direct lines are available, without any coding in between. Therefore the format to be reproduced can be defined clearly according to Section 3. On the other side, for delivery and recording of multichannel sound signals for consumer application, bit-rate reduction coding methods are necessary, in particular with the constraints of limited delivery and data capacities. This is valid with all media as sound broadcasting, television, film, and recording.

Therefore it has to be distinguished [2b]

- **Multichannel reproduction formats**
  as the result of special sound recording and reproduction technologies, as they are represented by the reference format 3/2 or 3/2/1 respectively, and considered here according to section 3 (but also 2/2, 3/1, 5/2 etc.), and

- **Coding and delivery formats for recording, delivery/transmission and connection**
  of multichannel signals with different media.

In the last case always the number of delivery channels should be considered in addition to the format. E.g., the code '4-2-4' means a matrixing format, with which the 4 signals (L, C ,R, S) are delivered or recorded respectively, and reproduced later in the 3/1 format.

Falsely and for advertising the coding formats very often are designated as to be "Surround Sound Systems". However, this is not correct because codec's can be seen actually as **auxiliary equipment only**, servo-mechanisms for the delivery of multichannel sound programs which in general should be independent on the relevant sound recording and reproduction method. For the listener at home anyhow it is of no importance on which way and by which method the multichannel signal bundle arrives to his final device, he is interested only in its quality.

The former matrix methods for the additional delivery of surround information via two-channel connections (on the basis of matrix inventions by L. Keibs, P. Scheiber, Cooper etc. in the sixties ) have been the original basis for a practical introduction of multichannel sound with film, recording, television etc., but could come to an success only after the Dolby Inc. brought them to the market. In this connection these coding methods very often might be changed with the multichannel recording and reproduction methods themselves. But such methods independently exist from former times (e.g. 3/1 Cinemascope 1953; 4/2-Todd-A-O 1955; 2/2 or 3/2 Stereo-Ambiophony respectively, 1960 etc.) and are developed further long ago.

The large disadvantage of the analogous matrix technology, the simplest coding method, is the impairment of signal quality. Therefore, the production permanently has to be monitored comparatively via such Codecs. For many applications with moving picture production these methods are sufficient and very successful, but not satisfying for music. In the meantime they are developed further, but now and with increasing amount they can be exchanged by discrete (digital) coding methods.

**Applied Coding Formats at the present time:**

**Analogous matrix formats for 3/1 reproduction (non-standardized, proprietary)**
**4-2-4:**

- *Dolby Stereo* (Motion Picture Matrix Coding) = cinema matrix
- *Dolby Surround/ Dolby ProLogic Surround* = consumer version
- *Virtual Dolby Surround* = multimedia application
- *MUSE*, in the meantime digital too, = TV in Japan.

**Matrix formats for 3/1 and 3/2 reproduction (non-standardized, proprietary)**
**4-2-4 / 4-2-5, 5-2-5:**

- *Circle Surround* = consumer and cinema matrix
- *Lexicon Surround* = consumer application

**Digital (lossy) coding formats for the reproduction of**
**3/2; 3/2/1 = 5.1 up to 5/2/1= 7.1; as well as 5-5-5 up to 8-8-8:**

- *MPEG-2-Audio* (backwards-compatible to MPEG-1 for two-channel reproduction)
= for Broadcasting, Television, Recording (DVD)
= international ISO-standard.

- *MPEG-2-AAC* (Advanced Audio Coding), non backwards-compatible.
= Up to now optimum solution for bit-rate reduced coding.
= international ISO-Standard.

- *Dolby Digital* (synonym: AC-3 = Audio Coding No. 3)
= for film and consumer application (US-TV/ITU; Recording, DVD). US-standard.
In Japan used in 3/1-MUSE TV.
- *Virtual Dolby Digital* = multimedia applications. Non-standardized.

- *DTS* (Digital Theater System) = for film and recording, DVD, non-standardized

- *SDDS* (Sony Dynamic Digital System) = for film, DVD, non-standardized.

**Digital (loss-less) coding formats for reproduction up to 7.1**

**- PCM = for all recording methods**

- *DSD* (Direct Stream Digital) = = for recording (archives, Super-Audio-CD).

**Appendix 3 to Doc. SSF – 01.1 – E - 2002**

# Subjective Threshold Values of the Reproduction Quality of Sound Signals

The basis of the following table 1 is the ITU-Recommendation R BS.644-1[22] (Table II). It was elaborated between 1986 and 1990 does not consider the now reached state-of-the-art of digital engineering and bit-rate reduction. Newer quality parameters and/or more stringent values are not yet covered. In the meantime, also an objective measuring method and special measuring equipment for any impairments with perceptual coding are available by international standardizing 1999 [27a].

Nevertheless, this table can be helpful during the standardizing work and the elaboration of "SSF Recommended Practice" documents and shows the perception measures which the sophisticated listener at the end of the chain can perceive and evaluate. From this, target conditions for the whole delivery/transmission chain can be derived (see also ITU-Recommendation 644-1, [22] Table I) as well as for the listening conditions and reproduction configurations within the studio and the home domains. Reference conditions should be always better than the subjective threshold values, but should not be very far above or below respectively, otherwise the expenditure is unjustified high.

Conditions for the home situation should be in the same order or might have perceptible deviations respectively.

The table can be an important information source, how far efforts of the industry are necessary and/or where discussions are useful on the justification of further requirements. Missing parameter values should be collected (via literature) and registered in appropriate manner.

By means of newer digital equipment it is possible without any problem to exceed the subjectively founded values, but might cause other subjectively perceptible impairments.

Some data in the table are questionable at the time being. Some comments in this direction are given – from the point of view of the editors – which are not coming from the original document,

For a long time the audio-frequency noise voltage level is measured international uniformly according to the Recommendation ITU-R BS.468-4 [23]. Unfortunately, there are some industry interests to circumvent this standard since longer time. E.g., for getting "beautiful figures" in the AES "Recommended Practice"-No. 17 only the CCIR-weighting curve was overtaken, but the zero-point was displaced by about 5.6 dB; and then dispensed with the quasi peak value measurement instead of r.m.s.-measurement. The source of this deviation is an older proposal by Dolby Inc. of 1978 to the CCIR in Geneva (now ITU-R), but this opinion was **none accepted unanimously** at that time.
So the results of the different measuring methods (ITU-R-468 with „dBdpq/468"and AES-17) can differ about 10 dB and more, and they are no more comparable.
Of course, it is the matter of the AES which method they will use and recommend. But nevertheless it is not correct to designate the AES/US measurement as "according to CCIR-468", as given in various measurement devices, so that very often confusion occur. In future, this wrong designation should be avoided.

For the limits of the "Early Reflections" in listening rooms at first the investigations of the RFZ by P. Schubert has given the fundamentals [46].

## Table 1 - Subjective threshold values (according to [22])

- Valid for the reproduction at the end of the transmission/delivery chain! -

| Parameter | Frequency of test signal | Subjective threshold value |
|---|---|---|
| Amplitude/frequency response[8]<br>(relative to level at 1 kHz) | 40 Hz - 125 Hz<br>125 Hz - 10 kHz<br>10 kHz - 14 kHz<br>14 kHz - 15 kHz | ± 1,0 dB<br>± 0,5 dB<br>± 1,0 dB<br>± 2,0 dB |
| Group delay variation | 40 Hz - 15 kHz | to be defined |
| Non-linear distortion THD (at maximum program level)[9] | 40 Hz - 15 kHz | - 52 dB |
| Error in reconstituted frequency (after transmission) | Any | 0,25 Hz |
| Error in amplitude/amplitude response | 1 kHz | to be defined |
| Level stability (over a 24-h period) | | 1 dB[10] |
| Noise and single-tone interference:<br>- idle channel conditions<br>- test signal level: + 9 dBu0s<br>- test signal level: -31 dBu0s<br>- single-tone interference | <br>-<br>60 Hz<br>60 Hz<br>- | <br>70 dB<br>to be defined<br>to be defined<br>80 dB |
| Level difference between A and B channels<br>(with stereo reproduction)[11] | 40 Hz - 125 Hz<br>125 Hz - 10 kHz<br>10 kHz - 14 kHz<br>14 kHz - 15 kHz | 2,0 dB<br>0,5 dB<br>1,5 dB<br>2,0 dB |
| Phase difference between the channels [11)] | 40 Hz<br>40 Hz - 125 Hz<br><br>125 Hz - 10 kHz<br>10 kHz - 15 kHz<br><br>15 kHz | 45°<br>to be interpolated)<br>30°<br>(to be interpolated)<br>90° |
| Cross-talk between A and B channels[11)]<br>- linear cross-talk<br><br><br><br><br>- non-linear cross-talk | <br>40 Hz<br>40 Hz - 300 Hz<br>300 Hz – 4 kHz<br>4 kHz – 15 kHz<br>15 kHz<br>Sound-program simulating signal | <br>15 dB<br>(to be interpolated)<br>20 dB<br>(to be interpolated)<br>15 dB<br>to be defined |

[8] The justification to higher sampling rates should have been reflected here; also the lower limit frequency should be reduced to 20 Hz.

[9] Here the conventional harmonic distortion (THD) is given; more effective would be to note the 'difference tone factor' = double-tone measurement.

[10] A change of 1 dB is perceptible only if it is a sudden change.

[11] It should be supplemented analogous to multichannel stereo conditions.

**Appendix 4 to Doc. SSF – 01.1 -E- 2002**

# Glossary of used termini and definitions in Section 2.

**(**See also the terminology collection of the Surround Sound Forum)

### *Listening conditions*

The term "listening conditions" describes the complex acoustic requirements for a reference sound field affecting in a listening room at the reference position, for sound produced by loudspeakers ('room-related reproduction').
This includes:

- the geometrical and acoustical characteristics of the listening room,
- the characteristics and the arrangement of the loudspeakers in the listening room,
- the location of the reference listening position or area,

which are producing the resulting sound field characteristics at the point or area.

Because the state-of-the-art does not yet allow the description of the reference sound field completely and uniquely by acoustical parameters only, some geometric and room acoustic requirements for a reference listening room and a reference monitor are given to ensure the viability of the listening conditions described.

### *Reference monitor loudspeaker*

"Reference monitor loudspeaker" means high-quality studio listening equipment which is use for technical and artistic quality monitoring and for the subjective assessment of sound program signals. The equipment comprises an integrated unit of loudspeaker systems in a specifically dimensioned housing, combined with special equalization, high-quality power amplifiers and appropriate crossover networks.

Possibly separate bass loudspeakers (subwoofer) can also be used in connection to such equipment. To be accepted as reference monitors the requirements according to section 2.2.2. and table 2 have to be fulfilled.

The electro-acoustical characteristics are measured with free-field conditions. Further measuring conditions see also in [3].

### *Directivity index D*

The directivity index is defined as the ratio in dB of the acoustical power of a loudspeaker, produced in the direction of the main axis, to the power, which is produced in that direction, if the sound source would radiate spherically.
The directivity index D is measured with one-third octave band noise, over the frequency range up to 16 kHz at the standard frequencies according to IEC-Publication 268-5. It should be as constant as possible over 500 Hz.

### *Transient fidelity*

The transient fidelity of monitor loudspeakers is measured by means of the decay time of an sinusoidal tone burst, to a level of 1/e (approximately 0.37) of the original level (on the main axis only) checked on an oscilloscope.

The decay time may not exceed 2.5 times the period of the corresponding sine wave.(the value of 5, required in ITU and/or Japanese documents seems not be sufficient).

### *Maximum sound level of monitor loudspeakers*

The maximum operating sound level will be measured for a time period of at least 10 min without thermal or mechanical damage and without overload circuits being activated. It is measured with a program simulating signal, according to IEC-Publication 268-1, and a sound level meter, set to flat response and R.M.S. slow).

### *Equivalent acoustic noise level of a reference monitor loudspeaker*

The equivalent acoustic noise level $L_{noise}$ generated by a single reference monitor loudspeaker and associated amplifier, referenced to a distance of 1m from the acoustical centre measured as R.M.S. (slow) in dBA, with the input short-circuit.

### *Acoustical centre*

The acoustical centre is the reference point for measuring purposes. It usually corresponds to the geometrical mid-point of the surface radiating the highest frequencies of the loudspeaker. It should be indicated by the manufacturer.

### *Operational room response curve*

The operational room response curve is defined as the frequency response of the sound pressure level produced by each monitor loudspeaker in the listening zone, using pink noise over the frequency range 40 Hz to 16 kHz. The one-third octave frequency response curve is measured at the reference listening point.

The difference between the operational room response curves produced by each of the multichannel monitor loudspeakers should not exceed the value of 2 dB within the whole frequency range.

(This term should not be changed with the 'percentile level', used in the domain of noise suppression, which very often is also designated as 'operational sound level').