
Surround Live 5
Creating the Live Sports Experience

Tom Sahara

Sr. Director of Remote Operations and IT

Turner's HDTV History

- TNT-HD began broadcasting May 2004
  - NBA – 52 Regular Season
    - "All Star Weekend"
    - "Inside the NBA" studio show
  - NASCAR NEXTEL Cup "Summer Series"
  - MLB Division Series & NLCS
- TBS-HD began broadcasting Oct 1, 2007

Define an audio philosophy

- Choose an audio viewpoint as your reference, a *viewer perspective*.
- Select sounds and effects that recreate the experience of being there.
- Keep it simple to protect your primary audience.

Viewers' Perspective

- Pick a seat in the arena that will serve as your reference point. *Where the viewer is sitting within the virtual venue you are creating*.
- Place the sound sources within the sound field that coincide with the main camera view.

Pick a seat…

- Baseball – Behind home plate
- Football – Nearside 50yd line
- Basketball – Nearside mid-court
- Golf – Front row of gallery
- NASCAR – Start/Finish line

Creating a sound space

- The proper selection and placement of microphones works better than the pan knob.
- Keep game action in front of the viewer, don’t bounce a ball “inside their head”.
- Mono music should be placed equally into left and right or use a surround synthesizer.

Placing the sound sources

- Dialog should always be placed in the center channel
- Live field of play effects are spread across the front right and left channels. The spread should match the viewpoint of the main camera.
- The LFE is for “effect” it is not a “channel”

Placing the sound sources

- Multi-capsule microphones create a fully enveloping sound field but the spatial separation degrades rapidly when down-mixed to L_t / R_t and mono.
- Use the proper mic and placement to capture the desired timbre

Placing the sound sources

- Be careful of low frequency build up in multi-channel effects, the low frequency response may sound fine in discrete channel systems but may tend to be muddy and over-bearing in bass managed systems.

Keep it simple!

- Caution must be exercised when using delayed and synthesized sounds in the surround channels.
- Monitor the level of sounds in the surround channels to prevent buildup and phantom sound images.

Stay awake!

- Monitor the downmixes often. The fold-down of the surround channels coupled with the phase differences of front to rear channels matrix encoding often cause problems in the resulting downmix (L_t / R_t) and mono mixes.

Know your audience

- HDTV surveys revealed that respondents felt the HD picture looks better when the sound is good.
- The majority of viewers listen to the audio that comes from their TV and not component surround systems.

What went wrong….

“What metadata?”

Aldous Huxley

- *“Technological progress has merely provided us with more efficient means for going backwards.”*

Aldous Huxley

- *“Technological progress has merely provided us with more efficient means for going backwards.”*
- Surround sound has advanced audio technology to the point where we can no longer reliably produce a program that sounds good in stereo and mono.

I can’t hear the announcers….

- Misunderstanding of the equipment
- Improper set up of the encoding equipment
- Improper set up of the listening environment
- Mishandling of metadata
- Abuse of equalization & processing

All I hear is the crowd….

- Improper encoding equipment set up
- Poor dynamic range control
- Improper set up of the listening environment
- Mishandling of metadata
- Improper set up of viewers' equipment
- Keep the "atta-boys" coming…

The LFE is too loud/soft….

- Improper set up of the equipment
- Improper set up of the listening environment (not calibrated)
- Mishandling of metadata
- Improper set up of viewers' equipment
  - "What is a subwoofer?"

The lip sync is off….

- Account for all video and audio delay paths.
- Be sure to include conversion delays of flat screen monitors.
- Use embedded audio whenever possible to maintain video and audio synchronization, especially during transmission.

Understanding the technology

“When all else fails, read the manual…”

Arthur Schlesinger, Jr.

- *“Science and technology revolutionize our lives, but memory, tradition and myth frame our response.”*

Arthur Schlesinger, Jr.

- *"Science and technology revolutionize our lives, but memory, tradition and myth frame our response."*
- We need to educate our staff so they do not rely upon, *"this is the way we always do it"* methods. They need to know what the technology offers and how to use it properly.

What am I listening to?

- Mixing in surround takes more than just watching the meters, trust your ears.
- Multiple speakers makes for a more complex listening environment.
- Multiple speakers reduces sweet spot.
- Analog troubleshooting techniques no longer work in digital.

Verify your listening environment

- Avoid large reflective surfaces
- Don’t forget about standing waves
- Proper speaker placement
  - Especially distance and direction
- Calibrate your sound system
  - Match phase and level
- Eliminate extraneous noise sources

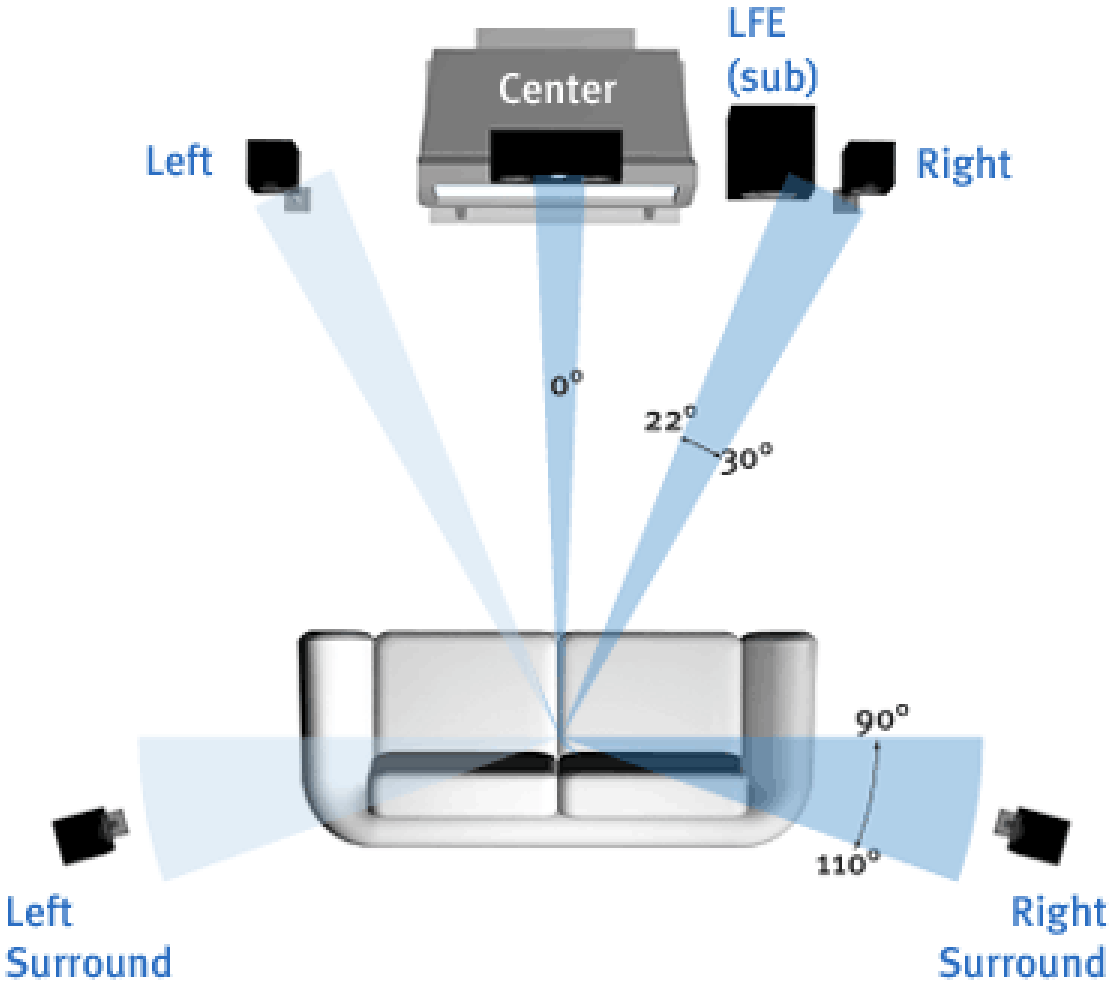
Verify your listening environment

Pink Noise Generator

Audio Analyzer

Sound Level Meter

Verify your listening environment



Dolby E

- Is a Transport (contribution) format
- Does not alter the audio that is put into it
- Does not add the LFE to any channels
- Is locked to video, has 1 frame delay
- Is not acted upon by the metadata

Dolby Pro Logic II (PLII)

- A matrix based encoder
- Analog stereo compatible
- Bass management
- Feeds the NTSC output for most networks

Dolby Digital (AC3)

- A Distribution format
- Digital and is a surround format
- Metadata directs the hardware to manipulate the AC3

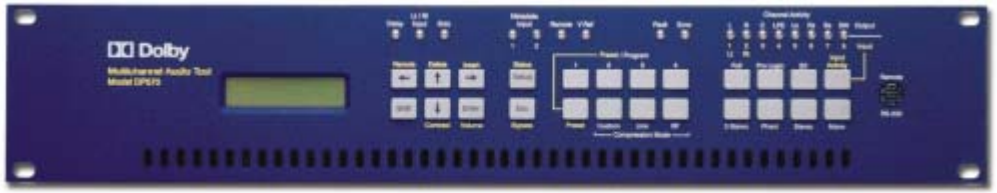
New tools are needed

- Digital source & monitor
- Dolby E analyzer
- Multi-channel meter with phase
- Loudness meter (ITU BS.1770)

New tools are needed

Dolby LM100 Loudness Meter

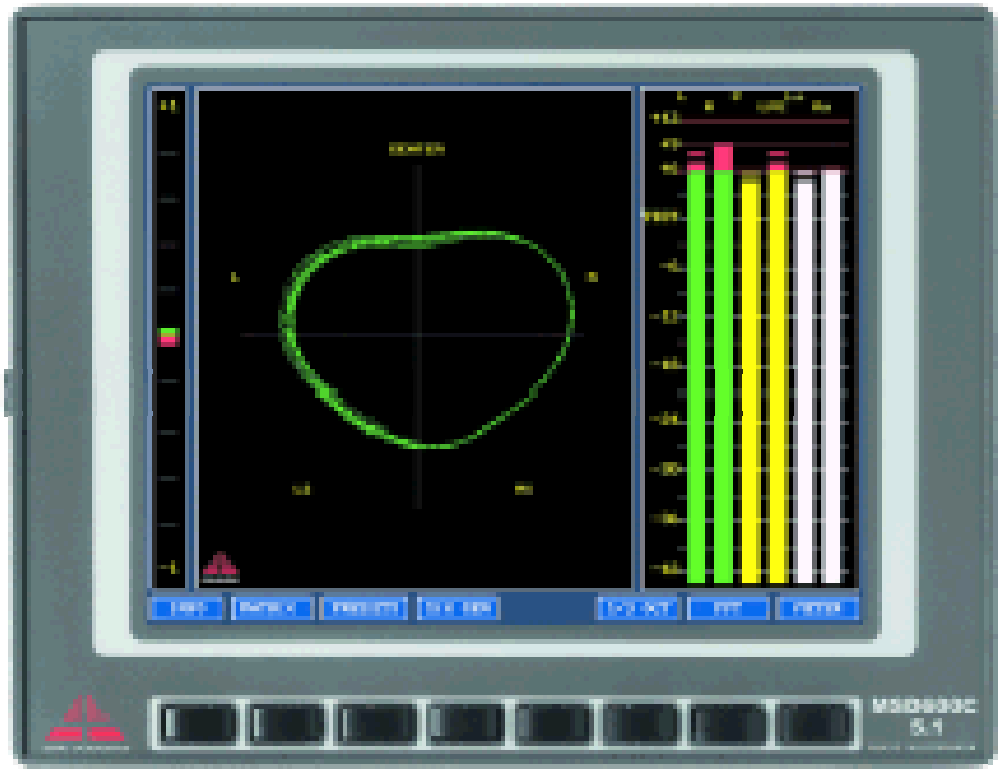
- Dolby DP570 Monitor



- Dolby DM100 Portable Analyzer

New tools are needed

- DK Audio – MSD600



Understanding the audience…

Not everyone is surround savvy!

Mike Trout

- *“Technology is ruled by two types of people: those who manage what they do not understand, and those who understand what they do not manage.”*

Mike Trout

- *“Technology is ruled by two types of people: those who manage what they do not understand, and those who understand what they do not manage.”*
- There are many variables that are out of the mixer’s control. He must focus on the aspects that he can affect through mix, mic placement, equipment and process.

The “Real” Viewer

- Often listens to audio from their TV
  - or at best the built-in stereo speakers
- Talks while watching TV
  - unlike movies, TV is still a social activity
- Do not always sit in the “sweet spot”
  - speakers placed by aesthetics and convenience SAF (Spouse Approval Factor)

Viewers' Listening Environment

- Room size – usually small
- Room type – acoustical properties
  - Wide variations in viewers' room acoustics
- Noise sources
  - Many noise sources and locations
- Listening position
  - Usually not in optimum location, "sweet spot"

Viewers' Speaker Arrangement

- Monaural – single source
  - TV set speaker, narrow bandwidth
- Stereo – built into TV
  - Slightly bigger speakers, limited bandwidth
- Multi-channel – components
  - Small home theater system
    - Bass management

Room performance

- Critical distance – D_c
  - Distance where direct sound from source equals reverberation and noise
- Frequency spectrum
  - Noise in the same spectrum affects speech intelligibility
- RT_{60} – Reverberation time
  - Reducing the reverberation time improves intelligibility (typically not a factor in small viewing spaces)

Critical Distance

- Where the direct sound is equal to the reverberant sound
- Greatly affected by reverberation and noise
- Intelligibility is highest when distance from source is less than D_c

Frequency Spectra

- Intelligibility improves when the noise is of a different spectra from speech
- Consonant sounds are most affected in the band 2kHz to 9kHz
- Crowd effects contains high frequency content that masks announcers

Consonant Sounds

- Consonants B - C - D - F - G - H - K…
  - Short duration
  - Low energy
  - Higher in frequency
  - Smaller packets
  - Most information of speech

Frequency Spectra

- Vowel Sounds A - E - I - O - U - (Y)
  - Longer duration
  - Higher energy
  - Lower frequency
  - “Glue” that hold consonants together
  - Little speech information

Reverberation

- Lower RT_{60} improves intelligibility
- Reflective viewing rooms decreases the distance that intelligible speech can be heard
- May cause localization problems because of reflections and phase cancellations
- Typically not a factor in home environs

Psycho-acoustics

- Minimum Audible Angle (MAA)
- Spatial unmasking
  - "Cocktail party effect"
- Precedence effect
  - "Haas effect", first impulse sets direction
- %ALCONS (Articulation Loss of CONSonants)
  - Measure of intelligibility of speech

Minimum Audible Angle

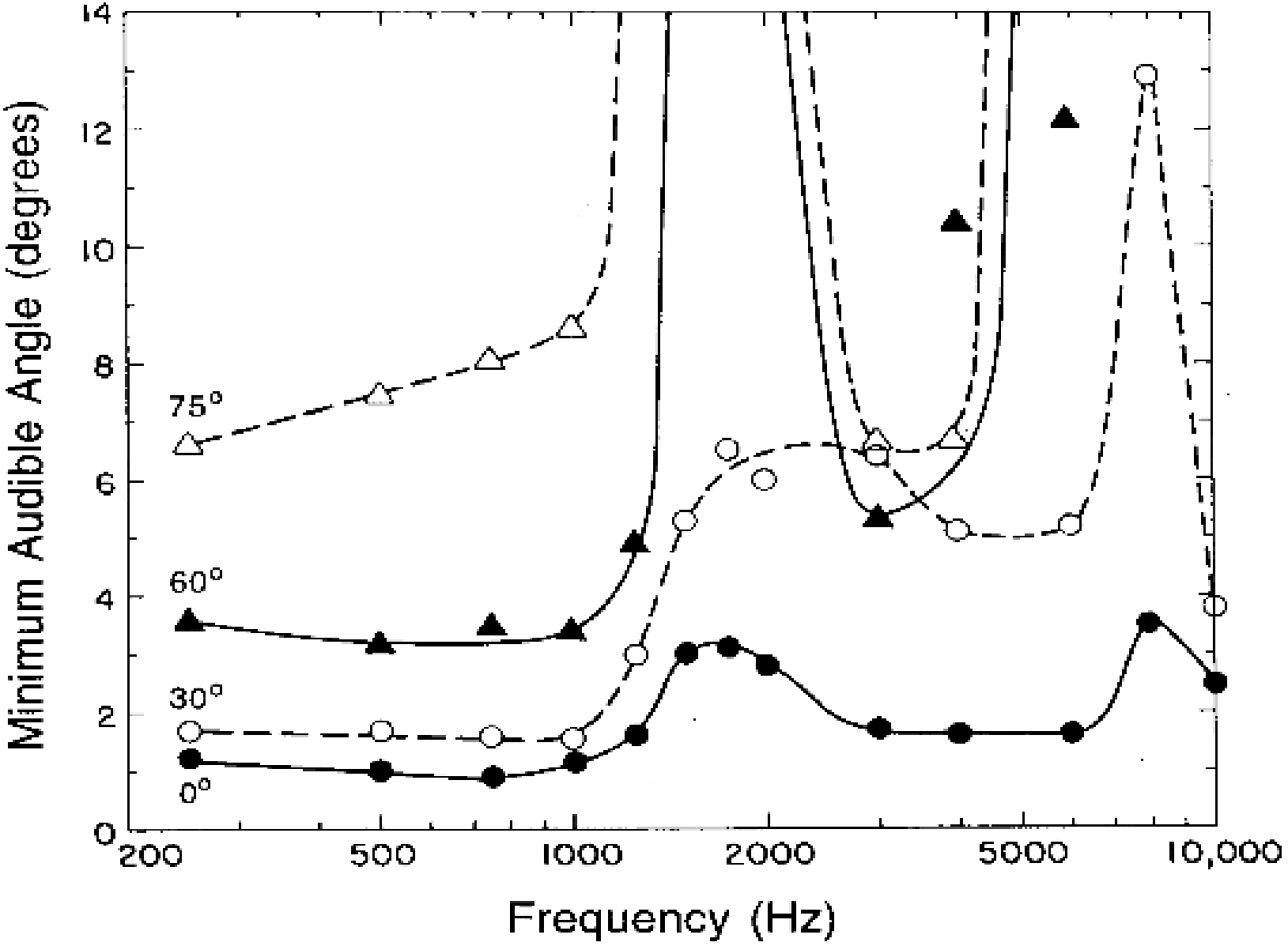
- Binaural - two ear listening
- Most accurate straight ahead ±30° az
- Maximum accuracy ±1° @ 0° az
- Least accurate at ±10° @ 90° az
- Less sensitive to vertical location 7°

Minimum Audible Angle

- At straight ahead 0° azimuth
  - Least precise 6,000 - 10,000 Hz
  - Less precise 1,000 - 3,000 Hz
  - More precise 3,000 - 6,000 Hz
  - Most precise 250 - 1,000 Hz
- “Cone of Confusion” @ 90° az

The minimum audible angle (MAA)

(Mills, 1972)



“Cocktail Party Effect”

- Ability to understand a conversation in an environment with many conversations.
  - Highest with familiarity of sounds
  - Highest when sounds are placed greater than 10 degrees apart
  - Decreases by 6dB for three sources
  - Drastically decreases with four or more souces

also “Spatial Unmasking”

- Spatial separation of sound sources improves intelligibility
  - improves with binaural hearing
  - +6dB @ 30 degrees
  - +3dB @ 90 degrees
- Down-mixing of reduces the spatial separation

“Haas Effect” (precedence effect)

- Early echoes that return in less than 35ms are ignored by the brain, sound is localized by the initial sound
- Echoes that are closely spaced in time are heard as a single sound
- The perceived direction of the sound is determined by the first sound pressure wave

%ALCONS Articulation loss of Consonants

- Intelligibility of speech measured by loss of consonant sounds
  - -3% Excellent
  - -7% Good
  - -15% Fair
  - -23% Poor
  - <-24% Unacceptable
- Affected by reverberation and noise

%ALCONS

- +25dB s/n for -3%ALCONS
  - Emergency alert systems
- +12dB s/n for -15%ALCONS
  - Typical public address systems
- Noise masking research indicates a +15dB s/n ratio for “good” reception of binaural sources

Putting it all together….

Where does it all lead …

What you understand…

- Intelligibility is how well you hear it, not how good it sounds
- Downmixing acts as the reverse of the “cocktail party effect”, it reduces the spatial separation of sounds
- Consonant sounds contain the speech information and are higher frequency but lower in energy

“ understand what you cannot manage”

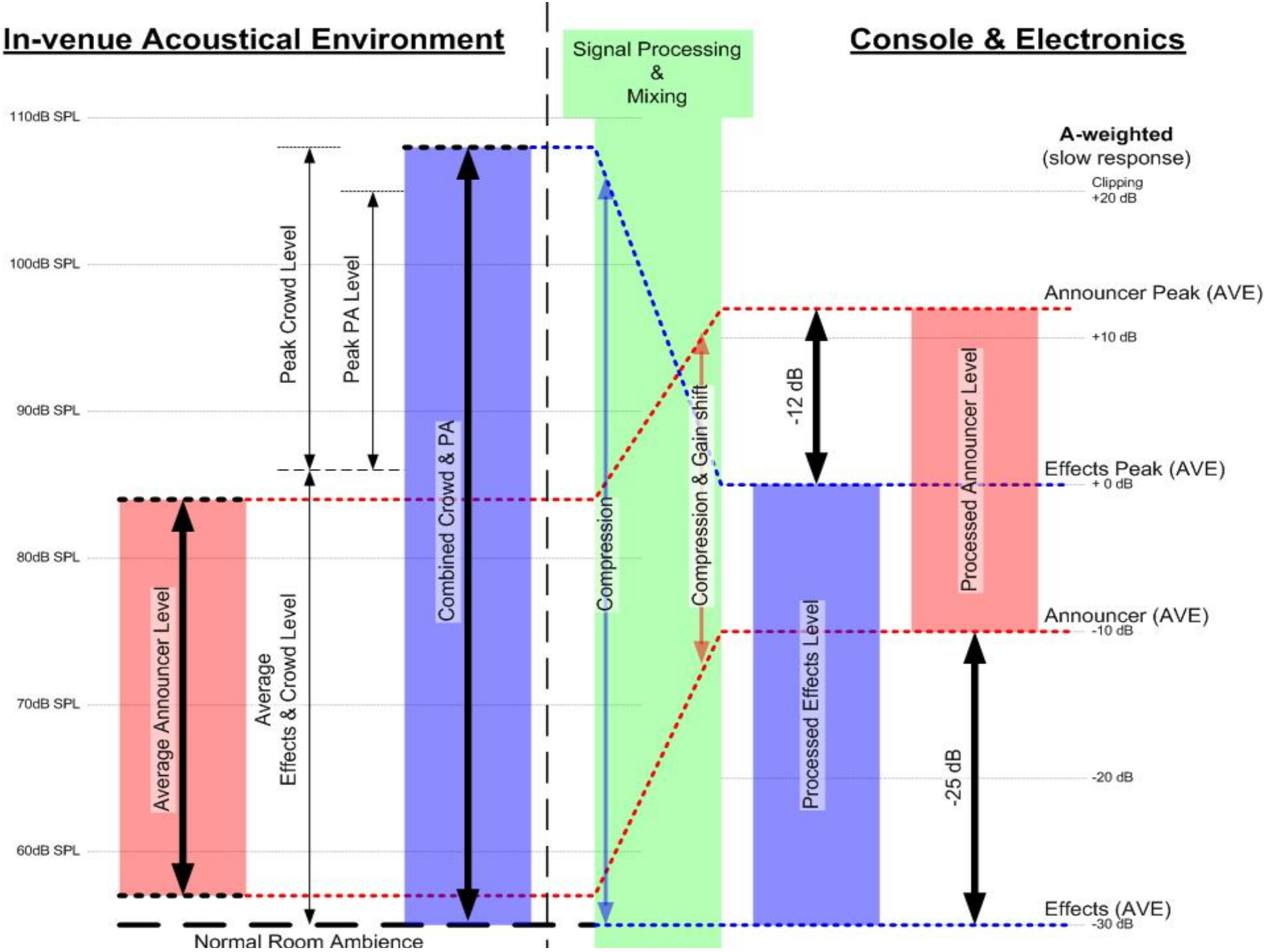
- The listener’s environment
- The listener’s equipment
- The listener’s age and hearing health
- Effects of downstream encoding & processing

Maintain Intelligibility

- Viewer environment is less than ideal
- Viewer equipment is less than ideal
- Noise - system or environmental
- Viewers want to hear the announcers whether they like them or not
- Use equalization sparingly, use mic selection and placement to capture the desired sound

Maintain Intelligibility

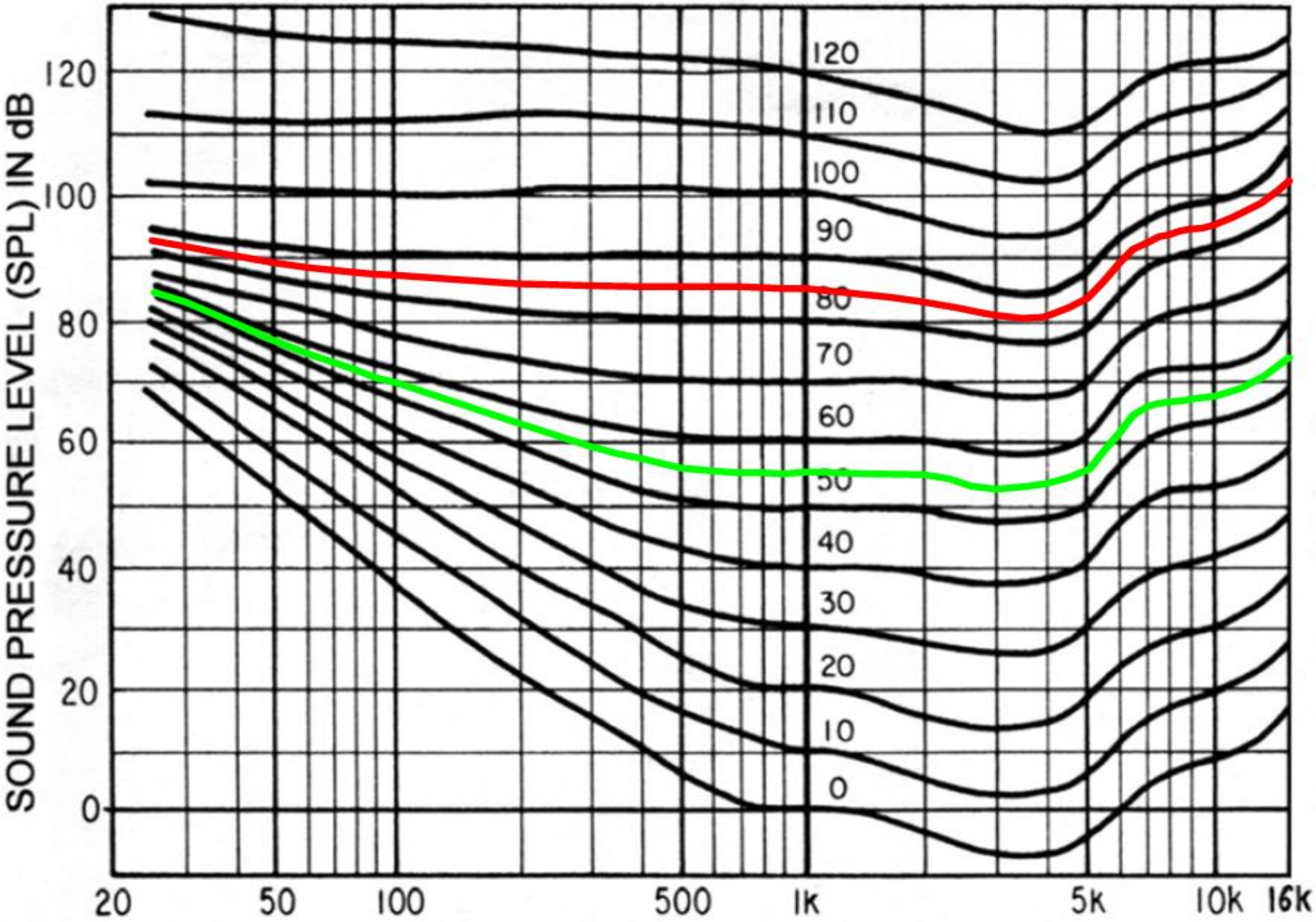
- Announcer should peak average +12dBA over effects (A weighted, slow)
- Surrounds should peak average –6dBA below left and right effects
- LFE should remain even relative to average program level
- Processing must be carried out with the final product in mind

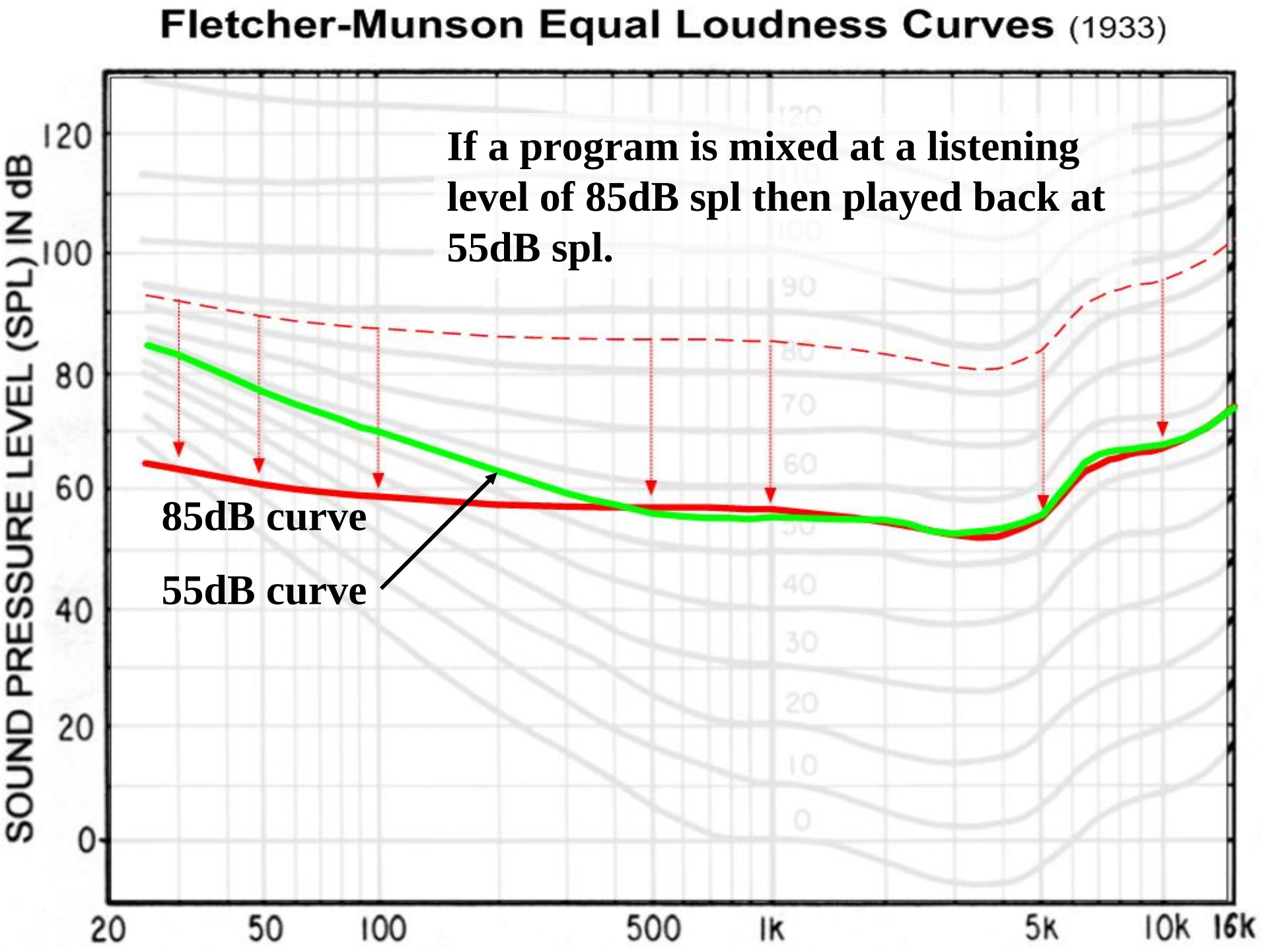

In-venue Acoustical Environment
Signal Processing & Mixing
Console & Electronics
110dB SPL
100dB SPL
90dB SPL
80dB SPL
70dB SPL
60dB SPL
A-weighted
(slow response)
Clipping
+20 dB
Announcer Peak (AVE)
+10 dB
Effects Peak (AVE)
+ 0 dB
Announcer (AVE)
-10 dB
-20 dB
Effects (AVE)
-30 dB
Peak Crowd Level
Peak PA Level
Combined Crowd & PA
Compression
Compression & Gain shift
-12 dB
Processed Announcer Level
Processed Effects Level
-25 dB
Average Announcer Level
Average Effects & Crowd Level
Normal Room Ambience

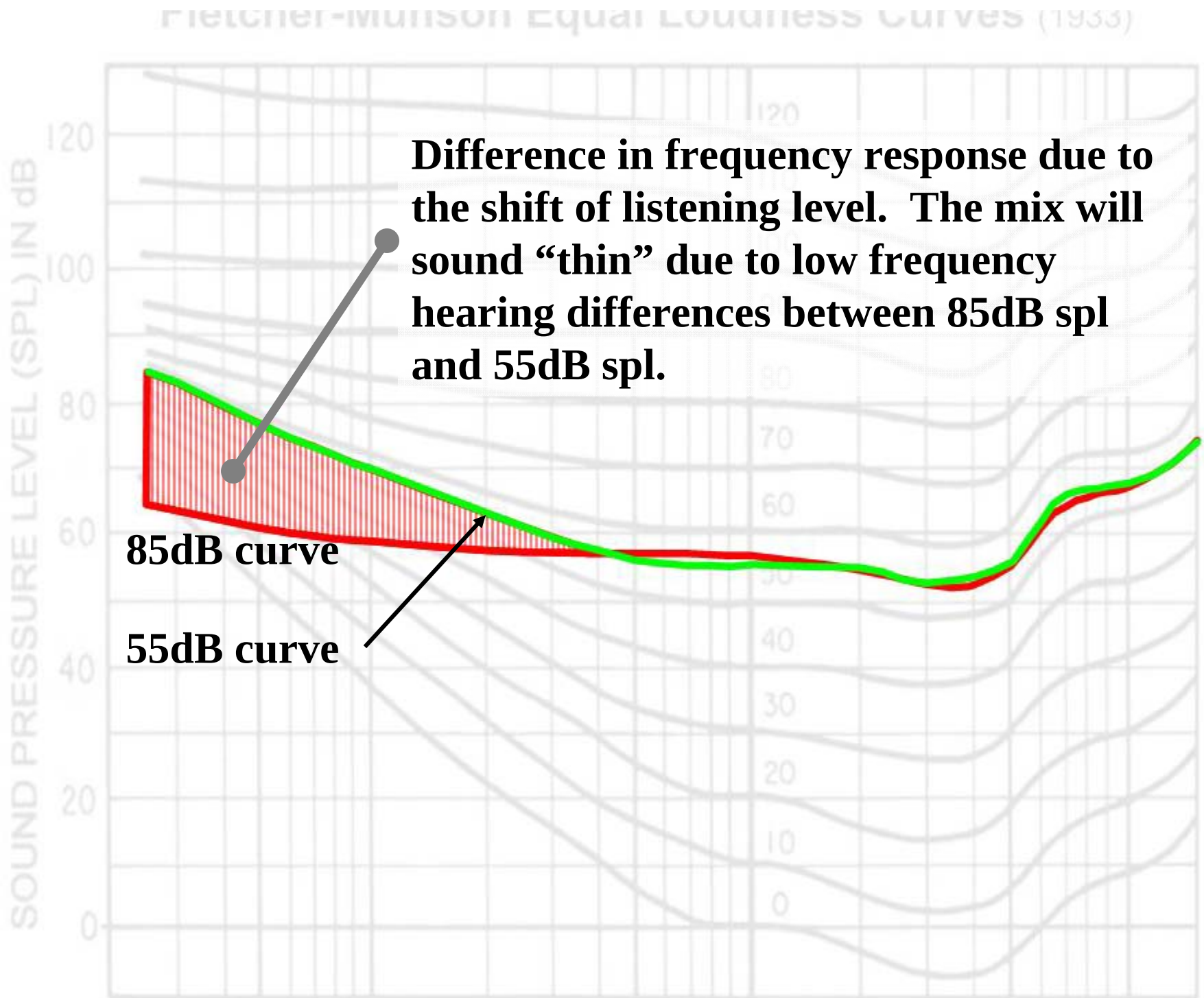
Listening levels

- Does the mixing room listening level affect the mix?
- Are there any standards?
  - Film mix stage 85db spl
  - TV proposed 79dB spl

Fletcher-Munson Equal Loudness Curves (1933)




Fletcher-Munson Equal Loudness Curves (1933)
SOUND PRESSURE LEVEL (SPL) IN dB
120
100
80
60
40
20
0
20
50
100
500
1k
5k
10k
16k
If a program is mixed at a listening level of 85dB spl then played back at 55dB spl.
85dB curve
55dB curve


Fletcher-Munson Equal Loudness Curves (1933)
SOUND PRESSURE LEVEL (SPL) IN dB
120
100
80
60
40
20
0
Difference in frequency response due to the shift of listening level. The mix will sound "thin" due to low frequency hearing differences between 85dB spl and 55dB spl.
85dB curve
55dB curve

Loudness

- ITU BS.1770 for loudness metering
- Maintain consistent perceived loudness from channel to channel across the TV channels
- Aim to control the “Loud Commercials”
- -24dB LKFS as “reference level”
- Viewer satisfaction is the key

Bringing it home…

Keeping all of the bits…

Transmission goals

- Maintain signal integrity and quality
- Maximize bandwidth efficiency
- Maintain audio/video synchronization
- Flexibility of transmission mediums
- Industry standard compatibility

Maintain signal integrity & quality

- Digital from end to end
  - HD-SDI for HD video
  - SDI for SD video
  - AES for audio
- Maximize bits and sample rate
- Maintain the reliability necessary for an on-air operation

Maximize bandwidth efficiency

- Transmit SD and HD signals within a single bit stream
- Compress the video and audio data rates to fit within a workable bit rate
- ASI multiplexed bit stream that can be sent either on fiber or satellite

Maintain audio/video sync

- Transmission system could not introduce synchronization errors
- Audio must remain frame locked to the accompanying video
- Accept embedded audio

Flexibility

- 60MB ASI is compatible with both satellite and terrestrial fiber
- Can be packetized over standard TELCO transports such as DTM, ATM and MPLS

Industry standard compatibility

- MPEG2 4:2:2 for HD video @ 45MB
- MPEG2 4:2:2 for SD video @ 15MB
- Dolby E for 5.1 surround audio @ 2MB
- MPEG Layer 3 for PCM audio @ 384kB
- 60MB ASI multiplexed bit stream
- DVB-S2 modulation scheme (satellite)