

Monitoring Surround-Sound Audio



Introduction

A picture may say a thousand words, but without audio to accompany the picture the impact of the material is subdued. In quality test it has been found that viewers can perceive a loss of visual quality if the sound quality is inferior, even when no change has been made to the picture quality. Therefore it is just as important to ensure the quality of the audio that is associated with the video material as it is to ensure the video quality of the material.

For a wide range of audiovisual content, the development of multi-channel, surround-sound audio technology has greatly enhanced the viewing experience.

Surround-sound technology has merged with digital television and other digital video technologies to create the home theater experience. The combination of enhanced picture quality and surround sound gives viewers a sense of total immersion and complete involvement in the program.

This technology evolution creates the need for multi-channel audio monitoring solutions. In particular, audio and video professionals need monitoring displays that help them visualize the auditory image that the viewer will experience.

Monitoring Surround-Sound Audio

► Application Note

In audio production, a visual representation of the sound image complements the auditory experience, helping audio engineers create the desired audio mix or more precisely adjust audio content in post production. In broadcast facilities, such visual displays help operators notice problems in multi-channel audio content more quickly and assist engineering in rapidly isolating the problem source. To describe the nature and use of this visual representation of the auditory image, we first need to review some of the principles and terminology associated with creating surround-sound audio.

The Surround-Sound Experience

Our visual sense operates over a relatively narrow angle directly in front of us. In contrast, our auditory sense detects sound from all around us and uses this total sensory immersion in creating our perception of real-world events. Limitations in this sound immersion can compromise realism.

Monophonic and stereophonic sound systems rely on reflections from the surrounding physical environment to immerse the listener in a sound space. Depending on the environment's acoustics, these systems can create a realistic perception of sound originating in front of the listener, e.g., an orchestra performing on a stage in a concert hall.

However, these systems cannot create sound that seems to emanate from the side of or behind the listener. Surround-sound systems use additional audio channels to drive speakers placed in these locations. These added sound sources create a more realistic sound immersion.

Multi-channel audio systems have another advantage that helps create a more realistic sound experience. In a monophonic sound system, listeners cannot distinguish the relative positions of the different sound sources. Stereophonic and surround sound systems can create the perception of sound sources at different locations by exploiting the cues that the human auditory system uses to determine the location of a sound source.

Sound Localization

Our brains use level, phase, inter-aural time delay, and spectral characteristics of the signals from our two ears to localize a sound's source. If the left and right speakers of a stereophonic system emit exactly identical signals, a listener positioned equidistant from each speaker will perceive identical, in-phase signals in each ear. The brain interprets these signals as the first reflection from a sound source located in front of the listener at the mid-point between the two speakers.

Two direct sound sources, the left and right speakers, have created a phantom sound source. Varying the signal level coming from each speaker will move the phantom source. For example, increasing the signal level in the left speaker will create a level difference between the signals perceived at each ear. For a listener positioned in the "sweet spot," i.e. equidistant from each speaker, the brain will use this level difference as a cue to position the sound source to the left of center. In stereo mixing, engineers use this technique, called intensity panning, to determine the left-to-right position of the various sound tracks.

Phase Difference and Correlation

Phase differences between the signals reaching each ear cause a different effect. The brain interprets in-phase signals as originating from a source at a specific location determined by the level differences between the signals. It interprets a phase difference between the signals as a cue that the sounds come from separate sources. Hence, a phase difference between the signals diminishes the phantom sound source illusion by "smearing" the source location. The listener perceives the sound source as originating over a wider range of positions, rather than a specific location. A large phase difference destroys the illusion and the brain can no longer localize a phantom sound source.

We can describe this effect in terms of the correlation between the signals driving the left and right speakers in a stereophonic system. Identical, in-phase signals have correlation values equal to +1. These signals will create precisely located phantom sound sources. Very similar signals, including identical signals with a small phase shift, have correlation values near +1. These signals will create a phantom sound source, but with less certainty about its location.

If the differences between the signals increase or they have a larger phase shift, the correlation value moves to 0. Driving the speakers of a multi-channel system with these uncorrelated signals will not create localized phantom sound sources. Listeners will perceive these signals as diffuse, non-localized, ambient sound. Negative correlation values arise when signals move towards an out-of-phase condition. These signals destructively interfere with each other and the sound may appear to come from outside the speaker locations. Identical signals that are 180° out of phase have a correlation value of -1.

Other Considerations

The level and phase relationships described above form the basis of the surround sound experience. A deeper study of sound reproduction would consider several other factors affecting fidelity and realism, including:

- Frequency- and directional-dependent responses of the human auditory system;
- The acoustics and the microphone placement in the sound recording environment;
- The effect of reflections and other acoustic characteristics of the listening environment;
- The electrical and acoustical characteristics of speakers, and their placement;
- Intentional or unintentional artifacts introduced by the compression and coding used to create multi-channel audio signals; and
- Other psychoacoustics phenomena associated with sound reproduction.

Typically, these factors affect the aesthetic assessment of the sound reproduction, which listeners characterize by subjective terms like “thin,” “dry,” “warm,” “harsh,”

“bright,” and “spacious.” Their significance depends on the desired quality level, the listening environment, and the capabilities of the sound recording, editing, and reproduction equipment involved.

While these other considerations play important roles, good stereophonic and surround sound reproductions depend on establishing the correct level and phase among the electrical signals driving the speakers in these systems. Modern monitoring instruments can help audio professionals verify these critical relationships. In particular, they can help verify audio level and phase in 5.1 multi-channel audio, the most common surround sound format.

Audio Channels in 5.1 Surround Sound

For several years, the film industry has used a multi-channel audio system as a standard format for cinema-based audio. Increasingly, to reproduce this surround sound experience in the home and give consumers a more cinematic effect, 5.1 multi-channel audio has replaced stereo in home entertainment systems. DVD's typically have 5.1 audio, and the television industry has started distributing and broadcasting this audio format in DTV systems. In conventional use, a 5.1 multi-channel audio system does not try to locate sound at precise, arbitrary locations. Rather, the different channels have particular roles (see Figure 1).

- The left (L) and right (R) channels drive the speaker pair in front of the listener (the mains) and carry most of the music. They typically operate like a stereo system.
- The center channel primarily carries dialog and drives a speaker positioned in front of the listener and between the mains.
- The left surround (Ls) and right surround (Rs) channels drive the left and right speaker pair placed to the side or behind the listener (the "surrounds"). They typically handle sound effects or ambient sounds that create the aural illusion of a particular environment or space.
- The low frequency effects (LFE) channel delivers low-frequency special effects, e.g. explosions, and drives a higher power, restricted frequency speaker (a sub-woofer), typically positioned in front of the listener.

Monitoring Surround-Sound Audio

► Application Note

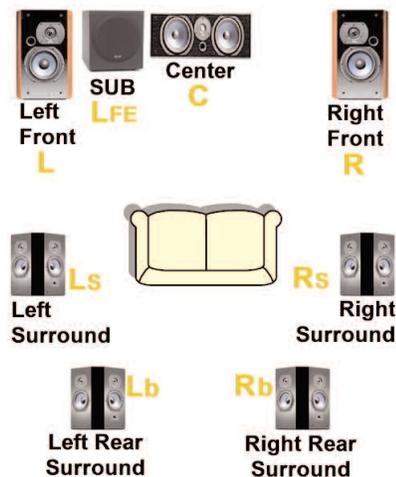
Dialog generally appears in the center channel because film and video producers usually want listeners to perceive this critical audio element in the center of the video field. They could place identical dialog signals in the L and R channels to create a phantom sound source in the center. However, this illusion only works for listeners located in the “sweet spot” between the left front and right front speakers. Using a dedicated center channel insures that listeners perceive dialog originating from the center of the video field, regardless of their location.

The L, R, C, Ls, and Rs channels form the “5” part of 5.1 multi-channel audio. They create the overall surround sound experience and handle the dialog and many special effects. They also exploit the sound localization characteristics of the auditory system to create appropriately located phantom sound sources. Below 150 Hz, the sound localization cues described earlier become much less effective. The L_{FE} channel (the “.1” in 5.1 audio) has a relatively restricted role in creating these dramatic, non-localized effects.

Although the speaker device is called a Subwoofer, in a surround sound system it is referred to as a Low Frequency Effects channel because, depending on the size of the speaker system being used by the viewer, the L_{FE} will have different responses. For instance a system with small satellite speakers will not have enough response to provide all the bass sounds and in this case these sounds can be directed to the L_{FE} channel. In the other case of large speakers in the room, they have more dynamic range to allow them to carry the lower frequency response of the bass sounds and there is less need to direct them to the L_{FE} channel.

Continuing extensions to the multi-channel audio system add further channels to the configuration.

Systems are now being used which are 6.1 or 7.1 channel systems. In 6.1 channel audio, an additional speaker is added to provide a mono back surround



► Figure 1. A Multi-Channel Surround Sound Speaker Placement.

channel. In 7.1 audio systems two speakers are used to carry the mono back surround channel to the Left Rear Surround (Lb) and a Right Rear Surround (Rb). Additionally, it may be necessary to monitor the down-mix of the multi-channel audio to a stereo pair. This can be denoted as Lo (Left only) -Ro (Right only) for a standard stereo mix or as Lt (Left-total) - Rt (Right-total) for a stereo down-mix which is Dolby Pro-Logic™ encoded.

Monitoring Multi-channel Audio Signals

Correctly producing and distributing multi-channel audio content requires a varied collection of monitoring tools. These include:

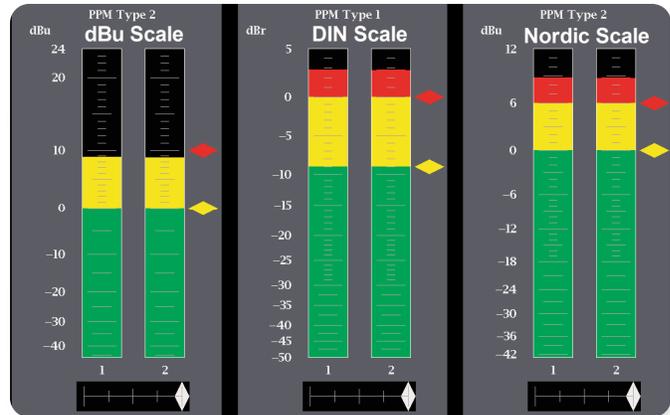
- Level bar displays with selectable meter ballistics and scaling for monitoring audio levels in multiple audio channels
- Lissajous*1 displays for checking the phase relationship between channels
- Displays for visualizing multi-channel level and phase relationships in surround-sound audio content.
- Data displays for verifying the metadata content contained in digital audio signals

*1 Also known as audio phase displays or audio vectorscope displays.

Tektronix Tools

Tektronix has developed several audio monitoring tools within their waveform monitors to provide solutions for measurements on video and audio signals.

- The WFM700 is a waveform monitor for monitoring High Definition (HD) and Standard Definition (SD) serial digital video signals. The addition of option DG allows the WFM700 to monitor various parameters of the audio signal including level and phase. The audio display can monitor up to eight channels either by extracting embedded audio within SDI (Serial Digital Interface) signal or applied as separate AES/EBU digital inputs via four BNC connectors on the rear of the instrument.
- The WVR series is a multi-format, multi standard rasterizer which comes in two configurations. The WVR7100 supports High Definition (HD-SDI Serial Digital Interface) serial digital input with options to support Standard Definition (SD-SDI) serial digital and analog composite inputs. The WVR6100 is a SD format unit which supports SD-SDI and an option for composite input. Each of the WVR series instruments supports a variety of different audio options. The base audio option allows monitoring of eight channels of AES/EBU digital audio or embedded audio from the SDI signal. Two banks of AES inputs (A & B) are available with this option. Bank B can be configured to be an active output for de-embedded audio or as a pass through of Bank A audio. The next options support analog and digital audio, with two sets of six channels inputs and the ability to output eight audio channels. These eight analog outputs are specifically useful when the Dolby decoder options are installed. There are two different Dolby options. Option DD is a Dolby Digital (AC-3) decoder only with limited two channel decoded output available. Option DDE allows support of both Dolby Digital and Dolby E with full decode to either analog or digital audio outputs.



► **Figure 2.** Analog Audio Scales for dBu, DIN and Nordic.

Level Bars

The level bars provide audio level indications based upon various audio scales. Within the Tektronix WFM700 series and WVR6100/7100 series, the audio bar scales are user configurable and allow the user to change the scale to suit the user's particular requirements. The overall dynamic range of the audio level can be displayed within the bar display. This can range from (decibels Full Scale) to 0dBFS in the WFM700 and -90dBFS to 0dBFS in the WVR series. Within this range a test level is also specified. This is the level of an audio tone that is designated for system level alignment within a complete facility. Depending on the facility, this test level will typically be at -20dBFS or -18dBFS. A peak level can also be set for the audio level bar. Typically this is at -8dBFS. Facilities with different audio practices can define different levels for the peak and test levels. The audio measurement device therefore has audio scales which are user selectable within the audio configuration menu. To provide more accurate level and phase display, the audio option 4X oversamples the audio stream. Several analog audio scales have pre-defined scales selectable within the WVR series, for instance dBu, DIN and Nordic as shown in Figure 2.

Monitoring Surround-Sound Audio

► Application Note

Visual indication of operation within desired levels is also provided. Below the test level, the bar will be shown in green, above the test level the bar will change color to yellow and above the peak level the bar will change to red. In-bar indicators can also provide information on clips, mutes, silence and over-scale alarm conditions present within the audio signal. These are defined as:

The **Clip** alarm is determined from the number of consecutive samples at full scale. The number of consecutive samples required before the alarm occurs is user configurable. When a Clip occurs this will be indicated at the top of the bar.

The **Mute** alarm is determined from the number of consecutive all zero samples. The number of consecutive samples before the alarm is triggered is user configurable. The in-bar indicator will display "MUTE" when present.

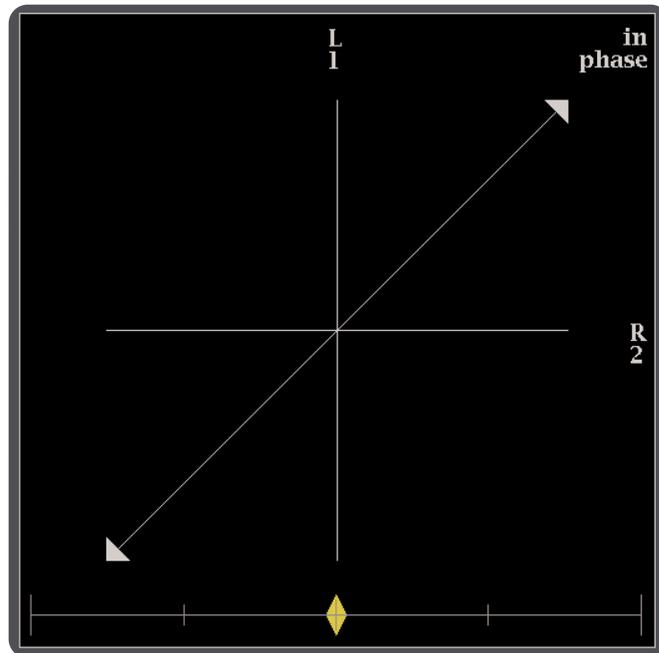
Silence – is a user selectable level. When the audio signal falls below this value for a period of time (in seconds), the in-bar indicator will display "SILENCE".

Over – is a user user selectable level. When the audio signal goes above this value for a period of time (in seconds), the in-bar indicator will display "OVR".

Ballistics

The ballistics of the audio meter is the speed of response of the meter to changes of audio levels. The ballistics are defined by various standards and provide information on the attack time (how quickly the bar level responds to an increase in the audio signal level) and the decay time (how slowly the bar response falls with a decrease in the audio level). There are several different types of meter responses which the user can select. These are defined as:

– **True Peak** - Shows actual signal peaks regardless of their duration. Rise times are essentially instantaneous. Fall time is like PPM Type 2, and requires 2.8 seconds to fall 20dB. The in-bar peak indicator will persist at the peak level for the "Peak Hold Time".



► Figure 3. Lissajous X-Y Display.

- **PPM Type 1** - Response equivalent to IEC Type I (essentially the same as DIN 45406 and Nordic N-9). The PPM Type 1 has a slightly faster attack time and a faster return time than Type 2, requiring 1.7 seconds to fall 20dB as opposed to 2.8 seconds for Type II.
- **PPM Type 2** - Response equivalent to IEC Type II (the same as defined in IEEE Std 152-1991). The PPM Type 2 has a slightly slower attack time and a slower return time than Type 1, requiring 2.8 seconds to fall 20dB as opposed to 1.7 seconds for Type 1.
- **VU** - VU meter ballistics are defined by IEEE Std. 152-1991, but with an extended dB-linear scale. The meter bars will also contain true peak indicators when VU is selected.

The WFM and WVR series audio monitoring displays will typically denote the type of ballistics being used within the audio display. The response time of the ballistic will often filter the instantaneous true peak level. Therefore, a true peak indicator is displayed as a tic mark within the audio bar.

The Lissajous (Phase) Display

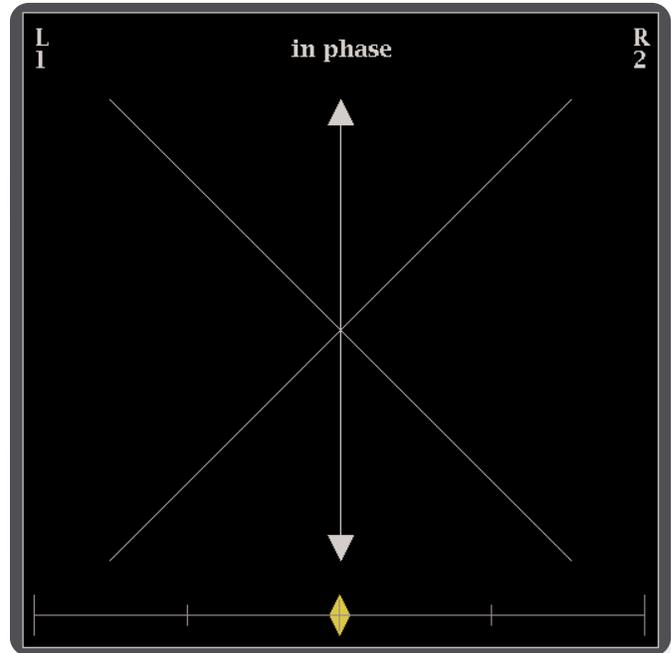
The left and right audio input signals are applied to an X-Y display similar to a vectorscope. The left channel is applied to the N-S axis and right signal content on the E-W axis similar to Figure 3. Many audio professionals are more familiar with the “Sound Stage” mode that simply rotates the display by 45 degrees to more easily visualize correct phasing of the channels (Figure 4). The Lissajous display provides instantaneous feedback of the overall energy distribution during a remix.

System phase errors

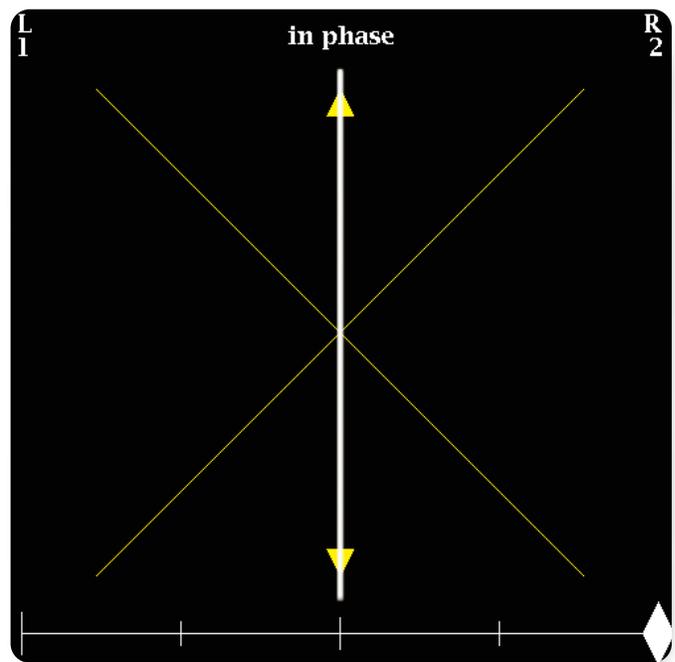
Phase errors can introduce any number of undesirable effects in an audio signal. A quick check with an audio phase monitor (Lissajous) display can help identify and quantify any significant amount of system phase error.

Ensure the Auto Gain Control (AGC) is enabled within the measurement instrument to make the edges of the trace ellipse just touch the phase tangent lines. If a test tone of equal amplitude, frequency and phase is applied to the audio system, a straight line that is coincident with the L=R axis will be observed. The left and right channels of the equipment under test will be exactly matched in phase and gain (Figure 5). If a slanted line is observed, the left and right channels match in phase but do not have the same amplitude. Figure 6 shows a signal with strong left content. A straight line falling on the L=R axis indicates reversed phase between channels as shown in Figure 7. Finally an ellipse whose major axis falls on the L=R line indicates equal amplitude but phase mismatch. When the amplitudes of the signal are identical but the frequency of the two channels are different then multiple ellipses will be displayed dependent on the frequency difference of the two channels as shown in Figure 8.

The Lissajous display can be used to identify various problems with the audio signals as shown in Figures 9 and 10. If a larger amplitude is present in one of the channels, this will move the sound axis towards that channel. Figure 9 shows a stereo signal with strong left content. Clipping distortion can occur within the audio signal path. In this case, the edges of the lissajous display will become square, as shown in Figure 10. The audio equipment needs to be adjusted to establish suitable levels to correct for this problem.



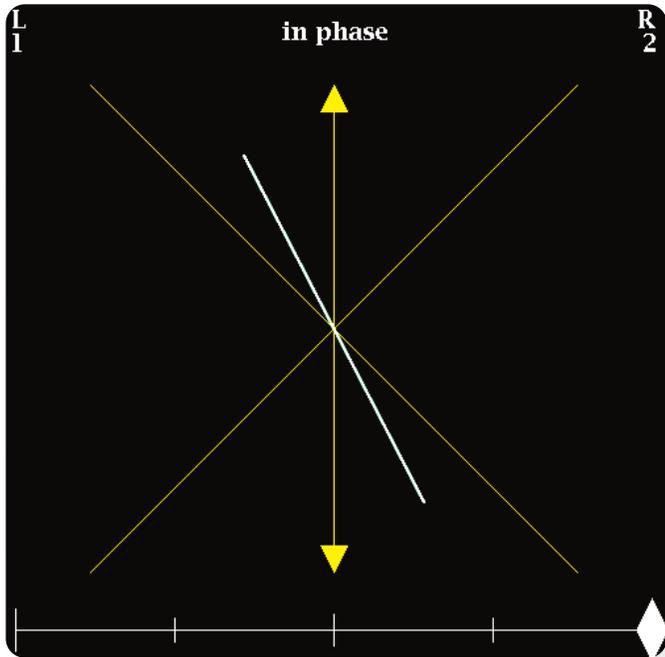
► Figure 4. Lissajous Display in Sound Stage mode.



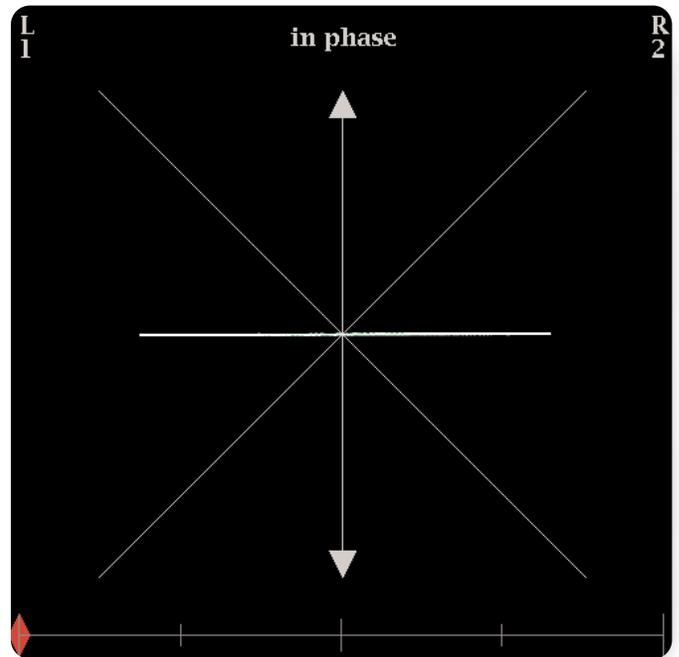
► Figure 5. Lissajous Display with Channel L1 and R2 identical.

Monitoring Surround-Sound Audio

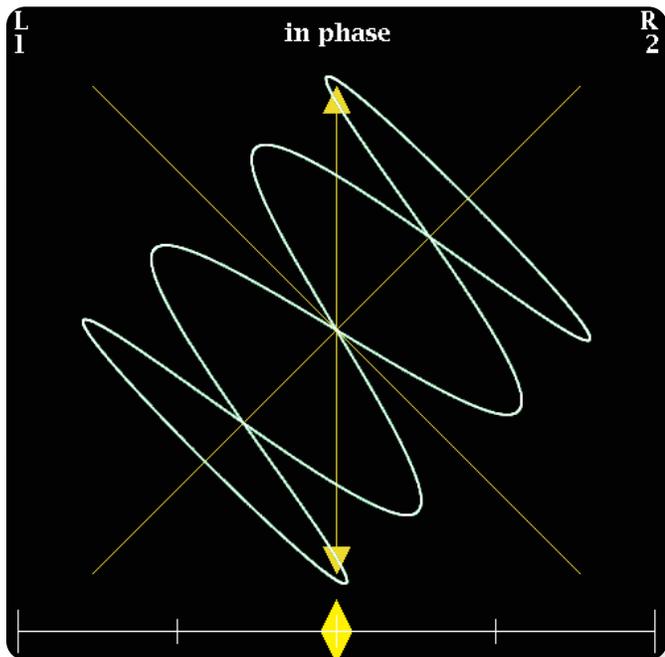
► Application Note



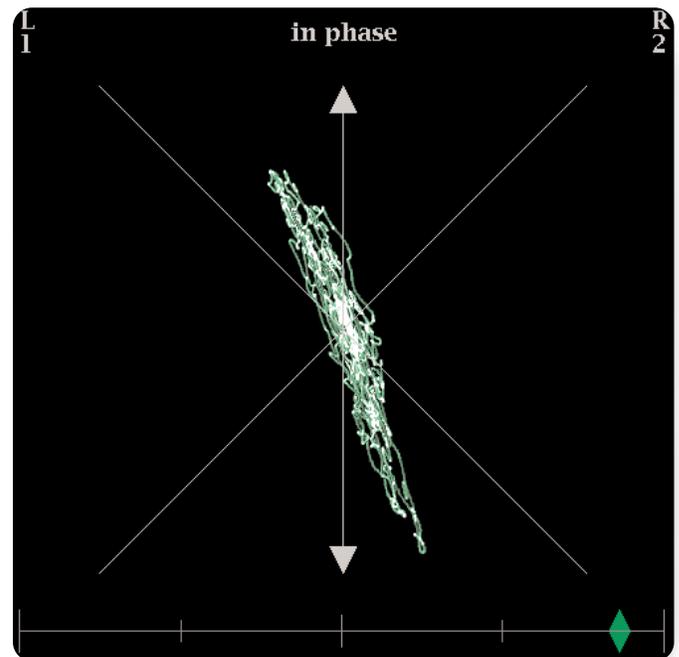
► **Figure 6.** Lissajous Display with different amplitudes for L and R.



► **Figure 7.** Lissajous Display with Channel L1 and R2 of equal amplitude and 180 degrees out of phase.



► **Figure 8.** Lissajous Display with Channel L1 and R2 identical amplitudes but different frequencies applied to input.

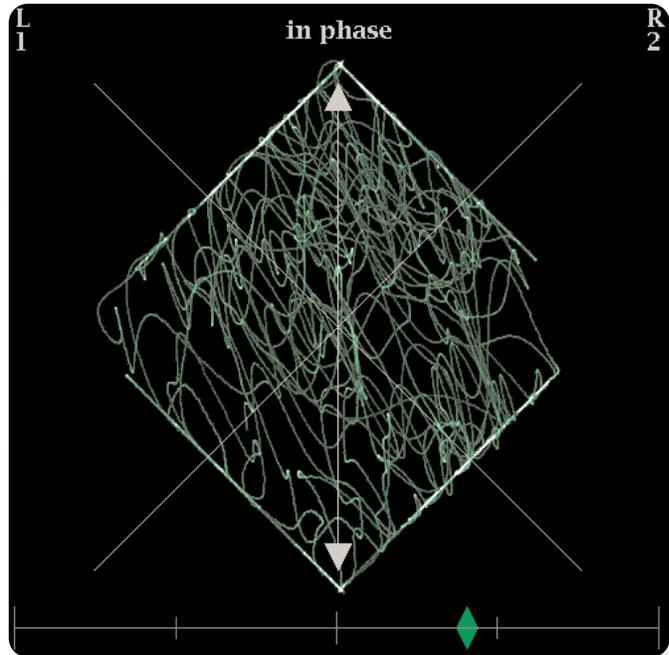


► **Figure 9.** Stereo Single with strong left content.

The Correlation Meter

The Correlation meter displays a true mathematical phase correlation (mono compatibility) between the two channels of the phase pair, independent of signal amplitude. This indicator can be found at the bottom of the Lissajous display and also beneath the appropriate bar level displays. The position of the diamond-shaped pointer indicates correlation between the channel pairs. A white indicator with a value of +1 indicates signals that are identical in frequency and phase (Figure 5). A green indicator indicates a correlation value of between 0.2 and 0.99, indicating highly correlated signals (Figure 9). A red indicator, with a value of -1 , indicates signals having the same frequency but that are 180 degrees out of phase (Figure 7). A yellow indicator at the center of the scale (0) indicates uncorrelated signals which are usually random signals (Figure 8).

Typically, a sound engineer in a multiple microphone recording session will apply a test tone of the same amplitude and frequency to ensure correct phasing and levels of each of the individual channels. The correlation meter helps the engineer to quickly identify and prevent any phase reversal between the channels before creating the stereo image of the final mix.



► Figure 10. Clipped Signal

In a modern digital surround sound system, the digital audio channels are carried on multiple AES/EBU pairs. For example - (L/R), (C/LFE), (Ls/Rs) in a 5.1 channel system. Therefore, three separate digital AES/EBU channels are required to carry the multiple channels of audio. It is easy for an audio engineer to compare the phasing of individual channel pairs present within a single digital AES/EBU audio channel pair by using the Lissajous display. But, how can the audio engineer check the complex multi-channel environment - the L-C and L-R phasing, for example?

Monitoring Surround-Sound Audio

► Application Note

To address this need, the audio options in the WVR series and WFM700 have a flexible Lissajous display. In the custom mode, the user can select any of the individual audio channels to be compared against any other channel present at the input to the monitor. The users can then select L-C, C-R, L-Ls, R-Rs and ensure the correct phasing of all of the channels. This is useful when using test tones to compare the individual channels, but is not completely satisfactory when monitoring live signals. A display appropriate for the unique needs and attributes of the modern multi-channel sound system is required.

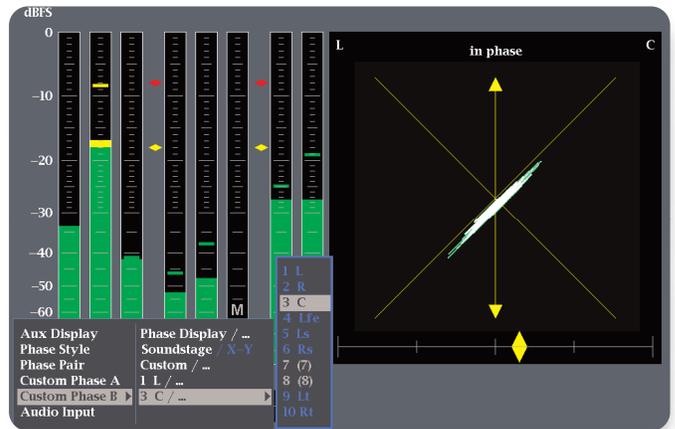
Figure 11 shows how the user can setup a custom Lissajous configuration, in this case to compare the Left and Center channels.

Surround Sound Display

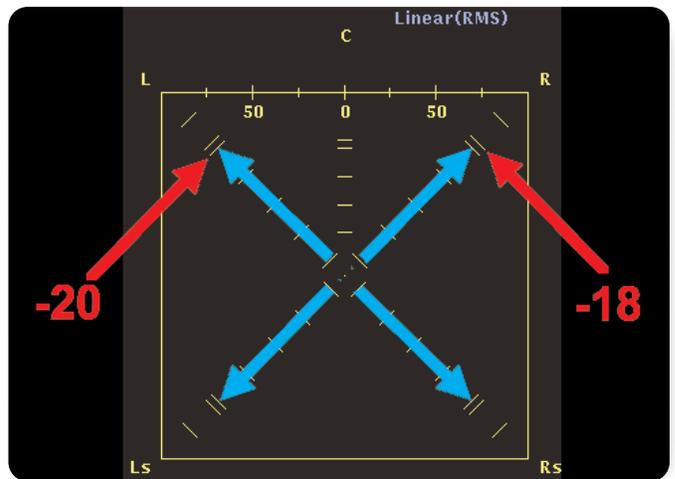
A surround-sound audio display was developed by RTW^{*2} for their Model 10800X Sound Monitor. They have licensed this display to Tektronix for use in video signal monitoring products like the WVR6100/WVR7100 and the WFM700.

Audio Level Indicators

The surround sound display associates an audio level with each of the five primary channels in a 5.1 audio system by determining the channel's RMS (Root Mean Squared) signal level. It can compute an un-weighted RMS value or can apply a filter that produces a frequency-weighted RMS value. Applying this A-weighting^{*3} filter adjusts for the frequency response of the human auditory system and yields an audio level value that better approximates the perceived loudness of the audio signal. As seen in Figure 12, the display shows the audio level in the L, R, Ls, and Rs channels on four scales originating from the display center and oriented toward the display corners. The upper left, upper right, lower left, and lower right corners of the display correspond to a 0 dBFS or 24 dBU (analog) level in the L, R, Ls and Rs channels,



► **Figure 11.** Flexible Lissajous display of WVR7100 showing Lissajous display of L and C.



► **Figure 12.** Audio Level Indicators.

respectively. The display center represents -65 dBFS (digital) or -41 dBU (analog). As the signal level in a channel increases, the cyan-colored level indicator lengthens from the center towards the display corner for that channel. Each scale has marks at 10 dB intervals, with a double set of marks at the commonly used alignment levels of -20 and -18 dB, as shown in Figure 12 for WFM700. In the WVR series, the mark is set by the test level.

^{*2} Radio-Technische Werkstätten GmbH & Co. KG, Cologne, Germany

^{*3} As defined in DIN IEC 651.

- As the correlation between the two signals increases toward +1.0, the line connecting the level indicators bends outward, away from the center and towards the potential phantom sound source.
- As the signals move towards an out-of-phase condition, i.e. correlation values approach -1.0, the line bends inwards, towards the center, indicating the destructive interference and reduction in total sound volume associated with out-of-phase signals.

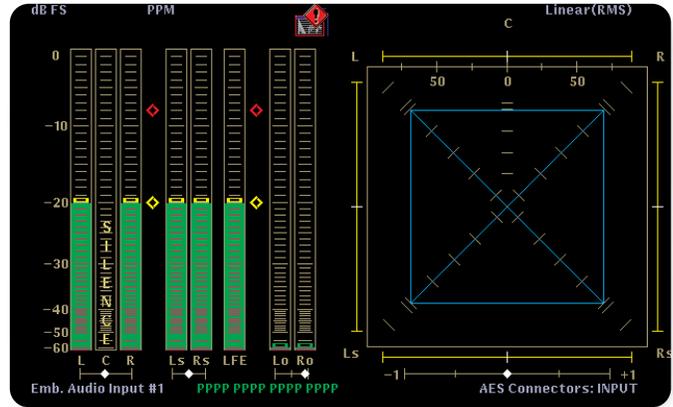
Total Volume Indicator

The display connects the ends of the audio level indicators to form a polygon called the Total Volume Indicator (TVI). The TVI indicates the level balance among the main and surround channels and gives an indication of the total surround sound balance. Figure 13 shows the characteristic square formed when each of the channels has un-correlated signals with the same amplitude. The TVI indicates the amount of correlation between signals in adjacent channels using the following conventions.

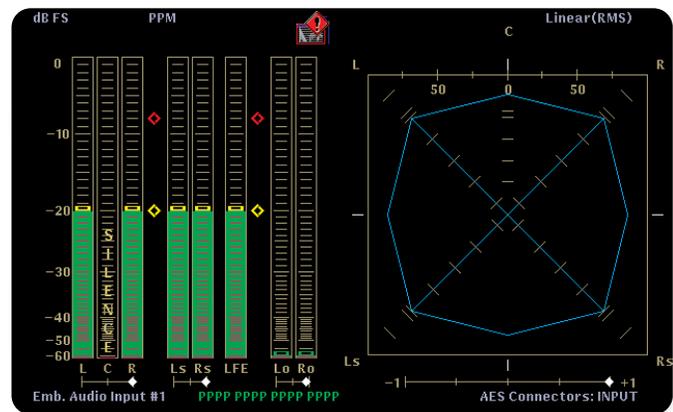
A straight line connecting the level indicators of two adjacent channels indicates that these channels have uncorrelated signals, i.e. a correlation value of 0.0.

As the correlation between the two signals increases toward +1.0, the line connecting the level indicators bends outwards, away from the center and towards the potential phantom sound source.

As the signals move toward an out-of-phase condition, i.e. correlation approach -1.0, the line bends inwards, towards the center, indicating the destructive interference and reduction in total sound volume associated with out-of-phase signals.



► **Figure 13.** TVI for uncorrelated signals of equal amplitude.



► **Figure 14.** TVI for in-phase test tones of equal amplitude and frequency.

Figure 14 shows the characteristic polygon formed by identical in-phase signals from each of the four channels.

The Center Channel

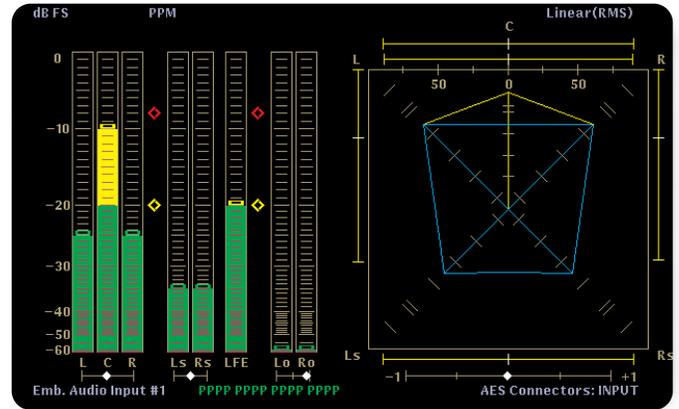
Recognizing the special role of the center channel in surround sound systems, the surround sound display handles this channel differently. The display indicates the center channel audio level as a yellow vertical line positioned between the left and right channel audio level indicators. The display forms a Center Volume Indicator (CVI) by connecting the ends of the L and C level indicators and the ends of the C and R level indicators (Figure 15).

The TVI and CVI operate independently. In Figure 15, the center channel has a strong presence, with dialog dominating the overall sound. Figure 16 illustrates a sound space with less center channel presence.

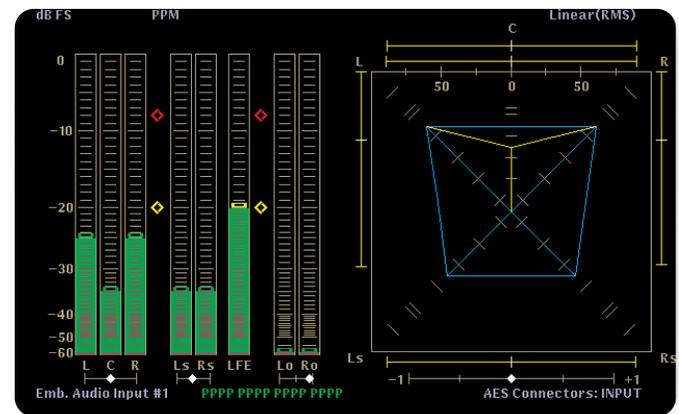
Phantom Source Indicators

Phantom Source Indicators (PSIs) positioned around the perimeter of the display offer additional help in visualizing sound localization. Four PSIs placed on each side of the display indicate the nature of potential phantom sound sources formed by the L/R, L/Ls, Ls/Rs, and R/Rs adjacent channel pairs.

These four PSIs operate in the same manner. Each PSI consists of a white tic mark, called the phantom source location pointer, which indicates the location of a potential phantom sound source. A variable length line extending on both sides of this location pointer indicates the typical listener's relative ability to localize this source.



► Figure 15. Surround sound program with strong center presence.



► Figure 16. Surround sound program with less center presence.

If the signals in an adjacent channel pair have a +1 correlation, they create a phantom sound source in a precise location between the two speakers. The phantom source location pointer appears on the side associated with the adjacent channel pair.

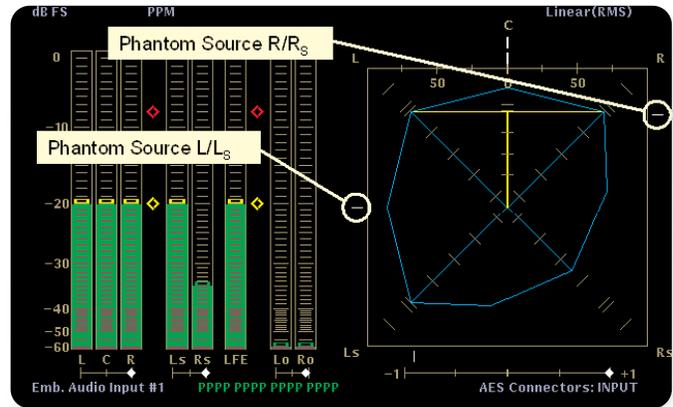
The position of the white tic mark depends on the level relationship between the signals in the adjacent channel. If the channels have equal audio level, the mark appears at the mid-point between the two display corners, indicating the listener will perceive the phantom sound source at the midpoint between the speakers (see L/Ls channel pair in Figure 17). For channels with different audio levels, the location of the phantom sound source moves towards the speaker with the higher level. In correspondence, the source location pointer moves towards the corner associated with the channel having the higher audio level (see the R/Rs channel pair in Figure 17).

A decrease in correlation between signals in an adjacent channel pair introduces some uncertainty as to the location of the associated phantom sound source. To indicate this, the PSI becomes a variable length line extending from the white tic mark toward the display corners associated with the channel pair.

As an additional visual aid, the line changes color as the correlation value crosses different threshold values. The fixed threshold values that trigger the color changes correspond to typical correlation values found in audio recording and reproduction.

Monaural sources generate correlation values above 0.9. The correlation values for signals in stereo recording primarily fall between 0.5 and 0.7, but can create an acceptable sound impression with correlation values over a wider range.

The uncorrelated signals that create diffuse, ambient sound typically have correlation values between 0.2 and -0.3. Because these signals do not create localized phantom source, these correlation values may indicate



► Figure 17. PSIs for in-phase test tones of varying amplitudes.

problems for signals intended to carry music or other more localized sound impressions.

Signals with correlation below -0.3 typically indicate an incorrect phase relationship that will create an undesired sound impression.

For signal correlations above 0.9, the PSI is a very short white line, indicating a highly localized phantom sound source. For correlation values below 0.9, the line becomes green. It continues to lengthen on each side of the phantom source location pointer as the correlation decreases, indicating increasing uncertainty in the location of the phantom sound source.

Once the line reaches a display corner, it will no longer lengthen with decreasing signal correlation. The location pointer remains in the position determined by the level balance between the adjacent channels. Consequently, unless the mark falls at the midpoint of a side, one end of the line will stop lengthening before the other.

Monitoring Surround-Sound Audio

► Application Note

For signal correlations below 0.2, the line turns yellow. When the signals become fully uncorrelated, i.e. the correlation value equals 0, the line will span the entire side of the display. This indicates that these adjacent channels will create a diffuse, ambient sound perception. Although the channel pair does not create a phantom sound source, the white tic mark still indicates the level balance between the channels (see Figure 18).

A further decrease in the signal correlation towards a -1 value does not change the length of the PSI or the position of the phantom source location pointer. The PSI will change color to red if the correlation falls below -0.3, indicating a possibly undesirable out-of-phase condition (see Figure 19).

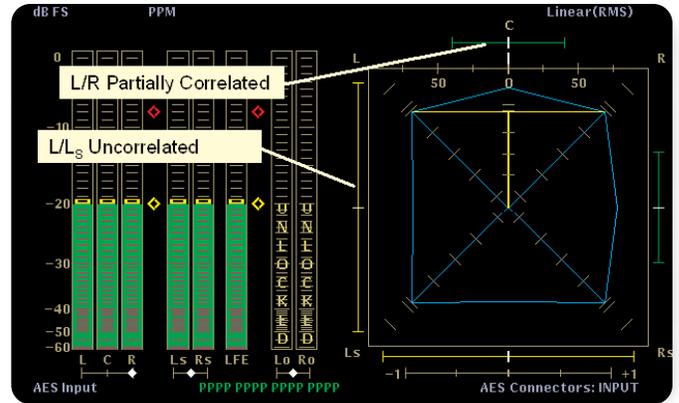
Figure 19 also illustrates an additional feature of the PSI for the L and R channels. For correlation values below -0.2, the display extends each end of the PSI at a 45° angle. This visualization aligns with the sound impression, i.e. listeners will perceive the source of these out-of-phase signals as outside the L and R speakers, not between them.

The L-C-R Phantom Source Indicator

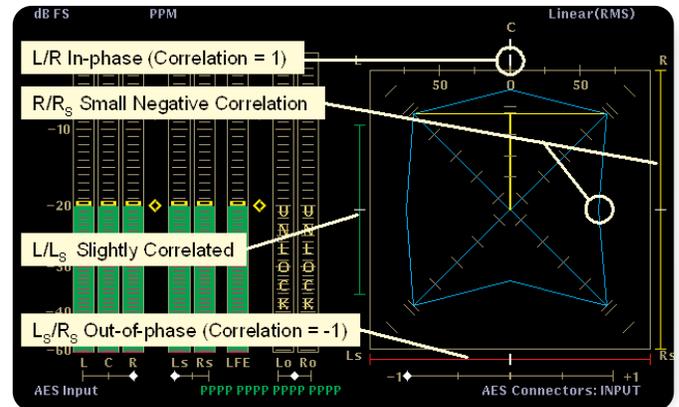
A fifth PSI located above the L/R PSI indicates potential phantom sound sources formed by the L/C and C/R channel pairs. This indicator behaves somewhat differently than the other four PSIs. The behavior of the phantom source location pointer on this PSI depends on the levels of the L, C and R channels. To describe this behavior, we let LL, LC, and LR represent these three levels. Then:

If $LL = LC$ and $LR = -65$ dB, the phantom source location pointer (white tic mark) for this PSI will appear above the 50 to the right of the 0 indicating a potential phantom sound source located midway between the R and C speakers.

If $LL = LC = LR$, the phantom source location pointer will appear above the 0 value on the top line of the surround sound display indicating a potential phantom sound source located midway between the L and R speakers.



► Figure 18. PSIs for test tones of varying amplitudes and phase.



► Figure 19. Surround sound display with LS and RS out of phase.

If $LR = LC$ and $LL = -65$ dB, the pointer will appear above the 50 to the right of the 0 indicating a potential phantom sound source located midway between the R and C speakers.

If $LL = LC = LR$, the phantom source location pointer will appear above the 0 value on the top line of the surround sound display indicating a potential phantom sound source located midway between the L and R speakers.

If $LR = LC = -65$ dB, the pointer will appear above the L on the top line of the surround sound display, indicating a direct source from the left speaker.

If $LL = LC = -65$ dB, the pointer will appear above the R to the top line of the display, indicating a direct source from the right speaker.

Changing the level balance among the L, C, and R channels will move the phantom source location pointer between the locations associated with the special conditions listed above. For example, if $LL = LC$ and $LR = -65$ dB, increasing the level in the R channel will move the location pointer from the 50 on the left towards the 0 in the center of the top line. This indicates that the location of the potential phantom sound source has moved from midway between the L and C speakers towards the center position (see Figure 20).

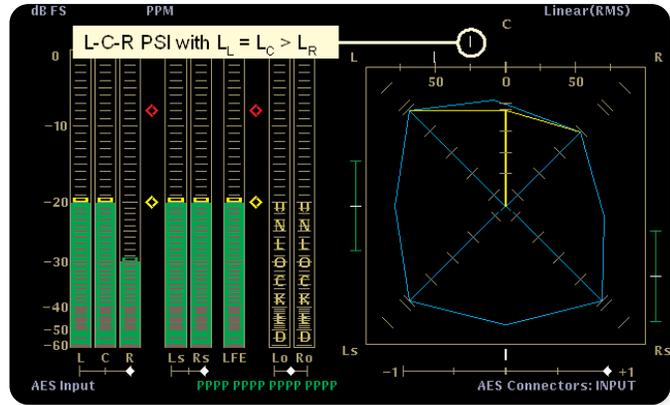
As with the other PSIs, the length of the line extending from the phantom source location pointer indicates the signal correlation. On this PSI, the lines to the left and right of the location pointer operate independently.

As the correlation between the L and C channel decreases from +1 to 0, the line to the left of the phantom source location pointer lengthens and changes color from white to green to yellow. The line extends to the upper left corner of the display when the L and C channels contain uncorrelated signals. A further decrease in correlation does not change the line's length, but will change the line's color to red and will extend the end of the line at a 45° angle, similar to the L/R PSI.

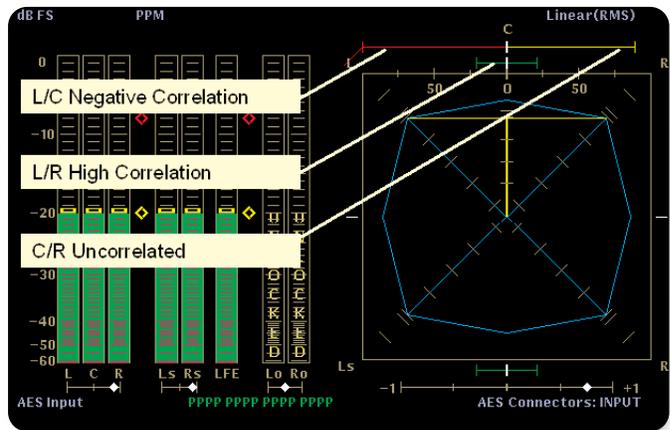
The line to the right of the phantom source location pointer behaves similarly to indicate the correlation between signals in the C and R channels (see Figure 21).

Correlation Meters

Figure 21 also illustrates another phase relationship indicator. Correlation meters appear beneath the audio level bars on the left of the display. The correlation meter beneath the L, C, and R audio level bars on the left shows the correlation between signals in the L and R channels. The correlation meter beneath the surround



► Figure 20. L-C-R Phantom Source Indicator (correlation = +1).



► Figure 21. L-C-R phantom source indicator (correlation < 1).

sound display, and beneath the Ls and Rs level bars, shows the correlation between signals in the Ls and Rs channels. The audio level bars display also contains a correlation meter for the Lo and Ro channels.

For uncorrelated signals, the diamond falls in the center of this single-axis meter. Increasing correlation between the signals causes the diamond to move right, towards the +1 position on the right end of the meter.

Decreasing correlation between the signals causes the diamond to move left, towards the -1 position on the left end of the meter. For identical, in-phase signals, the diamond will appear at the far right, +1 position. The diamond will appear at the far left, -1 position if the channels carry identical signals that are 180° out-of-phase.

Monitoring Surround-Sound Audio

► Application Note

The following table summarized the various correlation indicators in this display:

Correlation Value or Range	PSI Color	PSI Length	Position of Diamond on Correlation Meter	Typical Sound Impression
+ 1.0	White	No line	Right end	Localized sound
+0.9 to +1.0	White	Short	Near right end	Localized sound
+0.2 to +0.9	Green	Medium	To the right	Localized sound
+0.2 to 0.0	Yellow	Long	To right of center	Diffuse, ambient sound or poor quality localized sound
0.0	Yellow	Full length	At center	Diffuse, ambient sound or poor quality localized sound
0.0 to -0.2	Yellow	Full length	To left of center	Diffuse, ambient sound or poor quality localized sound
-0.2 to -0.3	Yellow	Full length with 45° line on L/R and L-C-R	To the left	Diffuse, ambient sound or poor quality localized sound
-0.3 to -1.0	Red	Full length with 45° line on L/R and L-C-R	To the left	Poor quality sound
-1.0	Red	Full length with 45° line on L/R and L-C-R	Left end	Poor quality sound

► **Table 1.** Correlation Indicators within the surround sound display.

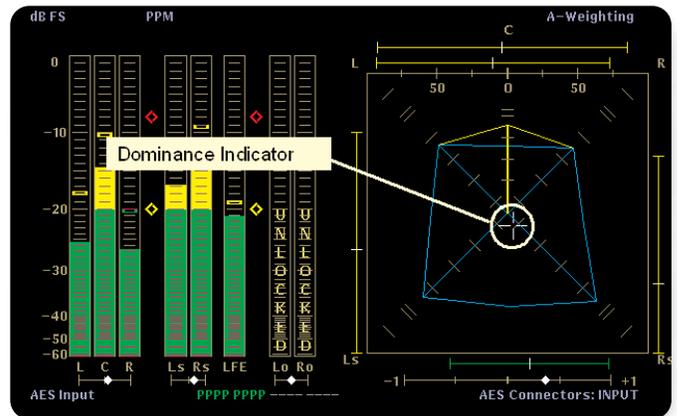
Dominance Indicator

The phantom source indicators help audio engineers visualize sound localization between adjacent channels. As a final visualization aid, the surround sound display has a white cross-hair dominance indicator that shows the location of the dominant sound created by the combined effect of the L, C, R, Ls, and Rs channels (see Figure 22).

A Real-life Example

Figure 23 illustrates the behavior of the surround sound display on DVD audio content instead of test tones. The cyan TVI indicates the sound volume and overall balance among the channels. The yellow CVI peaked above the TVI shows a strong center presence coming from the dialog in the scene.

The PSI shows some correlation between the signals in L/R, L/C, C/R, and Ls/Rs channel pairs. The L/R pair forms a somewhat localized phantom source; the other pairs do not form a localized phantom sound source. The phantom source location pointers show nearly



► **Figure 22.** Dominant Source Indicator.

balanced levels in the L, C, and R channels and that the mains dominate over the surrounds. This places the dominant sound in the front center, as shown by the dominance indicator. Finally, the level bar display shows a low signal level in the LFE channel. The surround sound display does not offer any additional information on this non-localized audio channel.

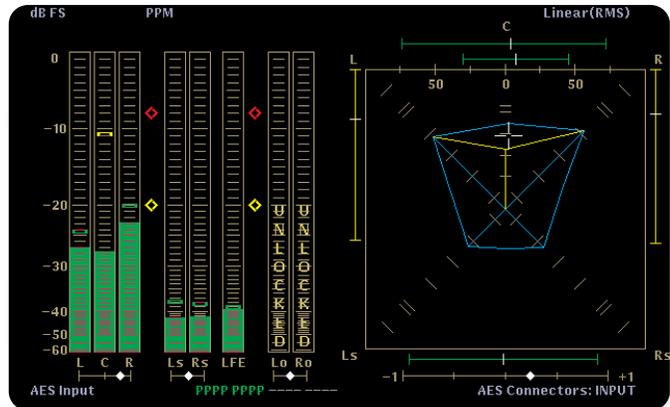
As a second example, we look at audio content in 3-1 format. This format has a stereo L/R pair, a center channel, and a single monaural surround channel. A 5.1-channel system carrying 3-1 surround sound audio has the characteristic display shown in Figure 21. The Ls and Rs channel carry identical signals, i.e. the monaural surround channel. This creates a highly-localized phantom sound source at the midpoint between the speakers. Once the user understands the basics of the surround sound display it becomes an essential tool for monitoring of multi-surround sound systems and it becomes very easy to interpret the interaction between channels: - all from one display.

Audio Compression

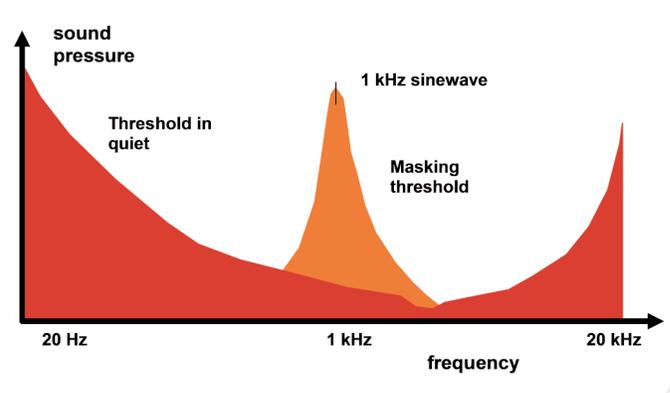
An increase in the number of audio channels being delivered drives a corresponding increase in the overall bandwidth required to carry these multiple channels of audio. For instance two channels of AES/EBU audio recorded with a sampling rate of 48kHz, at 20 bits per sample requires a data rate of 1.92Mb/s. For 3 AES/EBU pairs to carry a 5.1 surround signals a total data rate of 5.76Mb/s is required. Therefore, there is a need to reduce the overall bandwidth and compress the multiple audio signals into a more compact, single entity.

Different compression schemes use the same basic principles to remove sound that the listener will not hear, based on psychoacoustic models. The human ear requires a certain loudness of the sound before we will interpret the sound being produced. This produces a frequency curve below which the loudness of the sound at a certain frequency will not be heard, shown in Figure 24 as the “Threshold in quiet”. The human ear is most sensitive to frequencies of sound that are present within speech. Our ears are also not very selective; we are not able to detect individual frequencies in the presence of a dominant frequency. A loud sound at a specific frequency will be surrounded by a “masking threshold”. Other quieter frequencies in close proximity to the loud sound will not be heard because they will be masked by the more dominant louder sound.

Various compression scheme take advantage of these basic human acoustical factors to reduce the overall



► Figure 23. Surround sound display with live material.



► Figure 24. Psychoacoustic modeling of the human ear.

bit rate required to send the audio data. With multiple channels of audio, the data rate can also be reduced across the group of channels by looking for similarities between the channels, and only sending that set of data once with a pointer to the number of channels which require the same data.

Within this application note we are going to discuss monitoring tools for two particular forms of audio compression:

- Dolby E - Primarily used within broadcast and post production facilities, for professional use and not intended for consumer applications.
- Dolby Digital (AC-3) - Used for transmission of the signal to the home and for various consumer applications such as DVDs (Digital Versatile Discs).

Monitoring Surround-Sound Audio

► Application Note

For detailed information on these audio compression schemes please see further technical information at the Dolby website www.Dolby.com.

The WVR series option DDE provides Dolby E and Dolby Digital (AC-3) decoding, Option DD supports only Dolby Digital (AC-3) decoding.

Dolby E provides a light compression scheme, which allows for up to 8 discrete channels to be encoded onto a single AES/EBU digital audio signal. The Dolby E stream can be encoded either as 20-bit or 16-bit data.

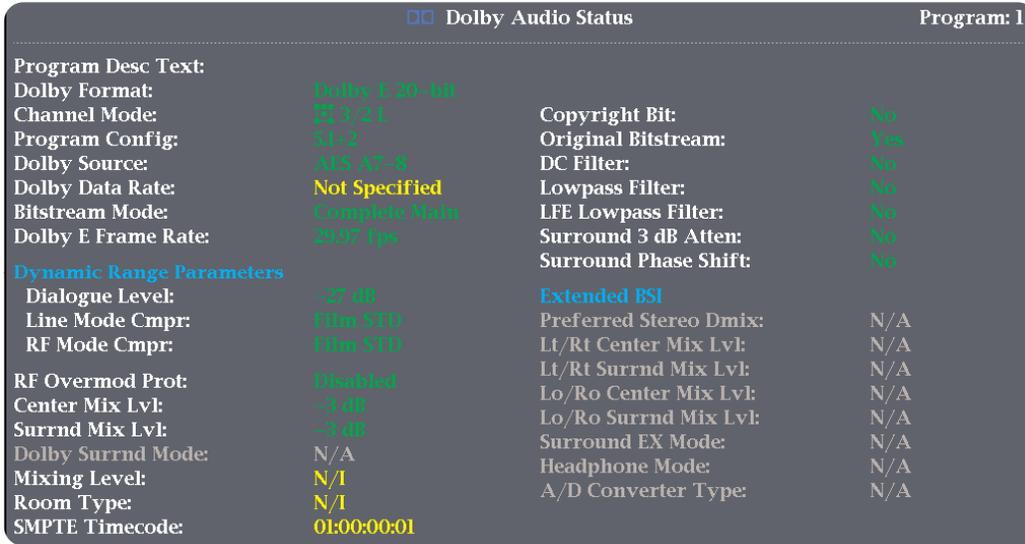
The encoding process has a low latency of 1 Frame and allows frame accurate editing to occur, provided the

Dolby E stream is correctly synchronized with the video signal. Within the WVR series, an alarm can indicate incorrect synchronization between the video signal and the Dolby E stream to ensure correct alignment. Besides the compressed audio data, within the stream, there is a set of metadata parameters which provides information about the signals present within the Dolby E stream. Some of these parameter sets support of embedding Dolby Digital metadata - useful when the signal is to be further converted to a Dolby Digital (AC-3) stream.

The 8 discrete audio channels within Dolby E can be configured in a variety of different ways as shown in Table 2.

Dolby E Metadata parameter	WVR Series Audio Bar Mapping	Number of Programs
5.1 + 2	L, C, R, Ls, Rs, LFE L1, R1	2
5.1 + 2x1	L, C, R, Ls, Rs, LFE M2, M3	3
4 + 4	L1, C1, R1, S L2, R2, C2, S	2
4 + 2 + 2	L1, C1, R1, S L1, R1, L2, R2	3
4 + 2 + 2x1	L1, C1, R1, S L1, R1, M1, M2	4
4 + 4x1	L1, C1, R1, S M2 M3 M4, M5	5
2 + 2 + 2 + 2	L1, R1, L2, R2, L3, R3, L4, R4	4
2 + 2 + 2 + 2x1	L1, R1, L2, R2, L3, R3, M4, M5	6
2 + 2 + 4x1	L1, R1, L2, R2, M3, M4, M5, M6	6
2 + 6x1	L1, R1, M2, M3, M4, M5, M6, M7	7
8x1 1+1+1+1+1+1+1+1	M1, M2, M3, M4, M5, M6, M7, M8	8
5.1	L, C, R, Ls, Rs, LFE	1
4 + 2	L1, C1, R1, S, L2, R2	2
4 + 2x1	L1, C1, R1, S, M2, M3	3
2 + 2 + 2	L1, R1, L2, R2, L3, R3	3
2 + 2 + 2x1	L1, R1, L2, R2, M3, M4	4
2 + 4x1	L1, R1, M2, M3, M4, M5	5
6x1	M1, M2, M3, M4, M5, M6	6
4	L1, C1, R1, S	1
2 + 2x1	L1, R1, M2, M3	3
4x1	M1, M2, M3, M4	4
7.1	L, C, R, Ls, Rs, LFE, Lb, Rb	1
7.1 Screen	L, C, R, Ls, Rs, LFE, Le, Re	1

► **Table 2.** Dolby E Metadata Program Configuration parameters for Channel Mode.



► **Figure 25.** *Dolby E metadata for program 1 ("5.1").*

Metadata

Dolby E specific metadata includes:

- Program Configuration
- Frame Rate Codes
- SMPTE Timecode
- Descriptive Text

Program Configuration

As shown in Table 2, this metadata set provides information on the number of programs present within the Dolby E stream and the format of those audio channels. The program configuration is shown as the “Program Config” within the Dolby Audio Status display of the WVR series, as shown in Figure 25.

Dolby E Frame Rate

A Dolby E signal must be locked to a video reference in order to ensure correct synchronization of the video and audio. The frame rate metadata parameter indicates the current frame rate used. Table 3 shows the possible

Frame Rate Code	Frame Rate
1	23.98fps
2	24 fps
3	25fps
4	29.97fps
5	30fps
0, 6-15	Reserved

► **Table 3.** *Frame Rate Codes*

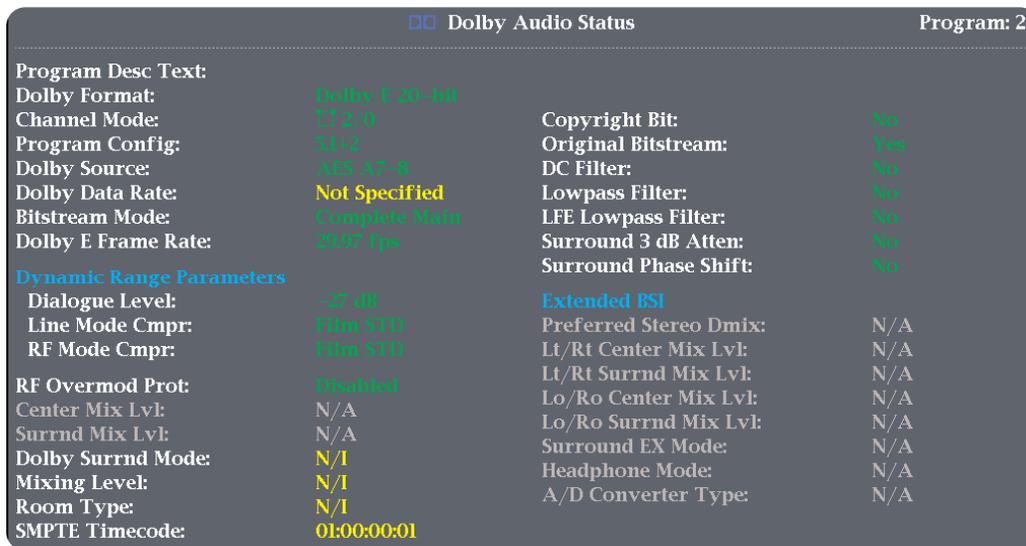
values. The WVR series will indicate this value in red if the Dolby stream frame rate does not match the frame rate of the video being monitored. This condition will also be logged.

SMPTE Timecode

This value indicates the SMPTE timecode associated with the Dolby E frame and is typically green when active data is present. When the data is inactive and not incrementing, the Dolby status display will be yellow to indicate a stale value.

Monitoring Surround-Sound Audio

► Application Note



► Figure 26. Dolby E metadata for program 2 ("2").

Descriptive Text

This metadata set can contain a user description of information about the Dolby E program.

The following list details the Dolby Digital (AC-3) metadata parameters which may either be carried by the Dolby Digital signal itself or within the Dolby E stream associated with material intended to be coded by a Dolby Digital encoder.

Dolby Digital Metadata

- Bitstream Mode
- Channel Coding Mode
- Center Mix Level
- Surround Mix Level
- Dolby Surround Mode
- LFE On
- Dialog Normalization
- Language Code Exists
- Language Code
- Audio Production info exists
- Mix Level
- Room Type
- Copyright
- Original Bitstream
- Extended Bitstream Information Exists
- Preferred stereo Downmix code
- Lt/Rt Center Mix Level

- Lt/Rt Surround Mix Level
- Lo/Ro Center Mix Level
- Lo/Ro Surround Mix Level
- Extended Bitstream 2 Information Exists
- Dolby Surround EX mode code
- Dolby Headphone Mode
- A/D converter type
- High Pass Filter on
- Bandwidth Low pass filter
- LFE Low pass filter on
- Surround 90 degree phase shift on
- Surround 3dB attenuator on
- RF pre-emphasis on
- Compression profile
- Compression
- Dynamic Range

The 3D's of Metadata

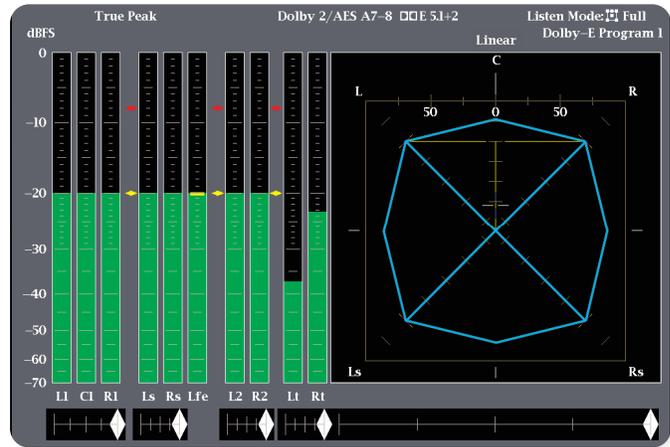
There is a considerable amount of information that is carried within the metadata for Dolby Digital (AC-3) to help the decoder process the signal appropriately. Of these, there are three key features which use this metadata and allow the user to customize his listening environment.

1. Dialogue Normalization (DialNorm)
2. Dynamic Range
3. Downmixing

Dialogue Normalization

This is referred to as dialnorm, is intended to keep the dialogue level at a constant level within the listening environment of the user and could be thought of as an automatic volume control. It works in conjunction with the volume setting of the receiver within the listening environment and the metadata parameter, to adjust the level so that the dialogue remains uniform throughout different kinds of program material. The dialogue normalization value should be obtained (set) during the post production process. Dialog normalization is an equivalent loudness method, standardized by IEC 60804, using an A-weighting integrated measurement over time known as Leq(A). As was mentioned earlier, the A-weighting filter adjusts for the frequency response of the human auditory system and yields an audio level value that better approximates the perceived loudness of the audio signal. The value of Leq(A) for speech measured over the time of the program material should be used as the dialog normalization value for the material. Dolby specifies a value of -31dBFS Leq(A) for a decoder operating in line mode and a value of -20dBFS Leq(A) for the RF mode. Different types of material will produce different dialnorm values. For instance, a news program may have a value of -20dBFS Leq(A) and a film may have a value of -27dBFS Leq(A). [Note that these are examples of the values that maybe obtained from this type of program material. Each program should be measured to obtain the Leq(A) value].

In order to obtain an -31dBFS Leq(A) at the decoder, an attenuation of -11dB must be applied for the news program and -4dB for the film, so that the dialogue level remains constant in conjunction with the volume level set by the user's receiver for the listening environment. Within the WVR series audio configuration menus, it is possible for the user to select the dialnorm value to be used for the levels of the audio bars display. This allows the user to see a representation of the effect of dialnorm on the audio level within the listening environment. Figure 26 shows the Dolby parameter for Dialogue Level of -27dB, which is typical for a film. Note that -27dB is the default value set up for a Dolby encoder and could possibly mean that the program provider has just used the default value that shipped with the unit rather than monitoring the signal and selecting the appropriate



► **Figure 27.** WVR series audio display of a Dolby E 5.1+2 signal with a Lt/Rt downmix of the 5.1 program.

value. If this value is red in the WVR series Dolby status screen, this indicates the value is invalid; the valid range is between -1 to -31.

Dynamic Range Control (DRC)

This is referred to as dynrng which allows for the program provider to set up specific dynamic range reduction for certain listening environments, while maintaining the full original dynamic range for other audiences. Within a movie, there can be many quiet scenes interspersed with loud, action packed, explosions and music. A quiet scene could be 40dB quieter than the normal average dialogue level, while the loud sound could be 20dB louder than the average dialogue level. Depending on the listening environment, these loud sounds could become objectionable if the impact of an explosion causes the floor to vibrate in your neighbor's apartment. Therefore, these sounds need to be compressed in level. If your receiver was set at a lower volume, then the quiet sounds could be inaudible and must be increased in level in order to be heard. Previously, material was compressed and expanded to meet the general audience requirements, but the original dynamic range of the program was lost. By use of the dynamic range parameters, the full dynamic range of the audio signal can be sent to the user and the receiver can change the settings of dynamic range to suit the user's environment. One can think of this as the good neighbor mode, allowing the user to listening in the late evening without disturbing anyone, yet still enjoying the material with a suitable compressed dynamic range within the listening environment.

Monitoring Surround-Sound Audio

► Application Note

There are two types of compression modes:

Line Mode, used on decoders with six channel outputs or two channel outputs. The user may be able to set the type of compression between Off, Light Compression and Heavy Compression, depending on the type of Dolby Receiver being used.

RF Mode, used on decoders which provide a RF re-modulated output to a television set or video recorder. In some cases this mode may also be used for personal computers which have small speakers with limited frequency range. This mode uses a full compression profile to prevent over modulation of the RF signal applied to the television or other device.

Six preset modes are available for the Dynamic Range Control profiles.

- Film Light
- Film Standard
- Music Light
- Music Standard
- Speech
- None

Within the WVR series, the specific profile being used is indicated in the Dolby Status display as shown in Figure 26. In this case, the Film Standard is being used.

These profiles apply specific cuts and boosts to certain regions of the input level in order to produce the desired compressed output.

Film Light

Max Boost: 6 dB (below -53 dB)

Boost Range: -53 to -41 dB (2:1 ratio)

Null Band Width: 20 dB (-41 to -21 dB)

Early Cut Range: -26 to -11 dB (2:1 ratio)

Cut Range: -11 to +4 dB (20:1 ratio)

Film Standard

Max Boost: 6 dB (below -43 dB)

Boost Range: -43 to -31 dB (2:1 ratio)

Null Band Width: 5 dB (-31 to -26 dB)

Early Cut Range: -26 to -16 dB (2:1 ratio)

Cut Range: -16 to +4 dB (20:1 ratio)

Music Light (No early cut range)

Max Boost: 12 dB (below -65 dB)

Boost Range: -65 to -41 dB (2:1 ratio)

Null Band Width: 20 dB (-41 to -21)

Cut Range: -21 to +9 dB (2:1 ratio)

Music Standard

Max Boost: 12 dB (below -55 dB)

Boost Range: -55 to -31 dB (2:1 ratio)

Null Band Width: 5 dB (-31 to -26 dB)

Early Cut Range: -26 to -16 dB (2:1 ratio)

Cut Range: -16 to +4 dB (20:1 ratio)

Speech

Max Boost: 15 dB (below -50 dB)

Boost Range: -50 to -31 dB (5:1 ratio)

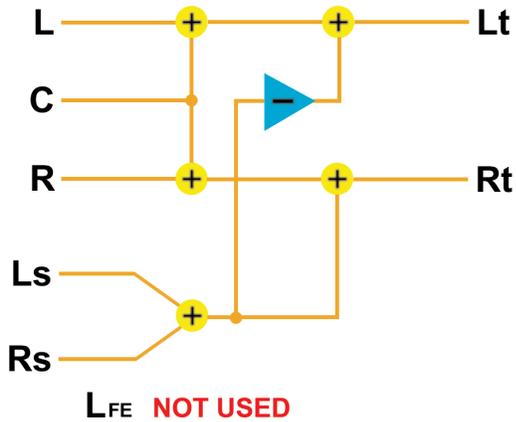
Null Band Width: 5 dB (-31 to -26 dB)

Early Cut Range: -26 to -16 dB (2:1 ratio)

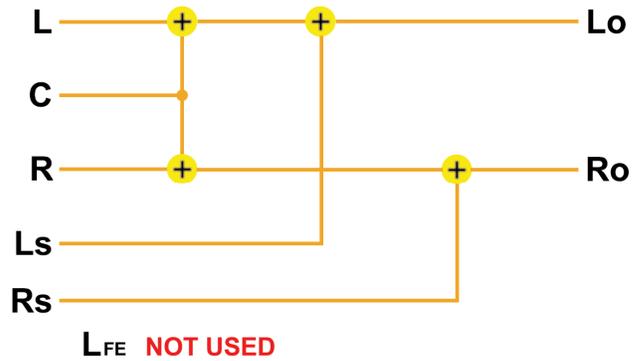
Cut Range: -16 to +4 dB (20:1 ratio)

None No Dynamic Range control profile is applied.

It is important the dialogue normalization is set accurately in order to guarantee the performance of the dynamic range controls. Within the WVR series, in the configuration menu, it is possible to apply both the dialnorm and specific dynamic range parameters of Line or RF to the audio bar levels. This allows the user to monitor the impact of the Dolby parameters on audio signal levels. If the dialogue normalization is not set correctly, this can have an effect on the dynamic range parameters, since dynamic range is dependent on a correct value of dialnorm. In some cases, this can cause the Dolby system to apply overload protection in order to prevent clipping and introduce distortion into the output signal. By monitoring the audio bars with dialnorm and dynamic range applied the user can also observe and log clipping which may be present in the program.



▶ Figure 28. Configuration of Lt/Rt Downmix.

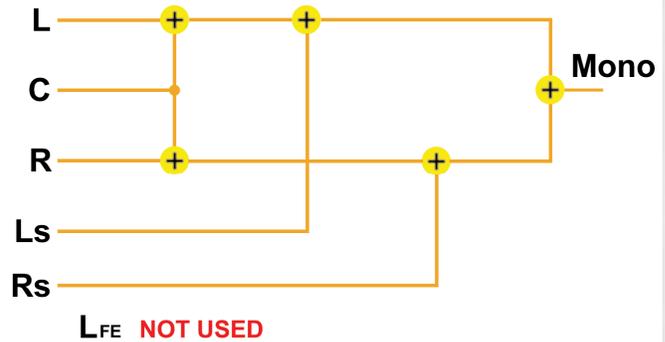


▶ Figure 29. Configuration of Lo/Ro Downmix.

Downmixing

This allows the multiple separate channels to be combined into either a mono or stereo mix and provides compatibility with users who do not have a digital surround sound system. They can still enjoy the program in mono or stereo. The stereo downmix can take two forms - a Lo/Ro (Left only/Right only) which is a standard stereo signal or Lt/Rt (Left-total/Right-total) which is a Dolby Pro-Logic stereo mix. This also allows for compatibility with older analog surround sound systems.

It is important to ensure routing of the multiple channel downmix into the appropriate output. The WVR series allows the user to configure the Downmix output as Lt/Rt or Lo/Ro or mono. The appropriate output is then displayed as a bar display in the audio measurement tile. In the WVR series, it is also possible to output the analog audio Downmix to a pair of speakers within the listening environment. Figure 28, 29 and 30 show how the separate channels are combined in the various Downmix options. Within the Dolby metadata, several parameters are used by the decoder to determine attenuation factors that will apply to the various separate multiple channels that are then combined in



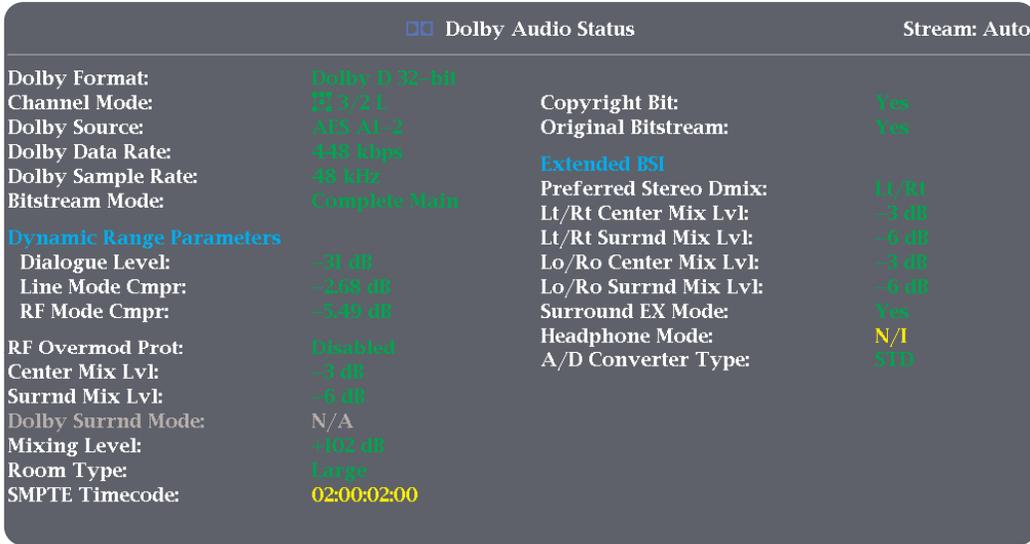
▶ Figure 30. Configuration of Mono Downmix.

the Downmix output. The WVR series employs a 180 degree phase shift to create a simplified compatible Dolby Pro Logic output that is suitable for decoding by a Dolby Pro-Logic receiver for monitoring purposes.

However, if a Dolby Pro Logic mix is required for broadcast purposes, then the discrete decode channels should be combined by a Dolby Pro Logic encoder that produces an output signal with a 90 degree phase shift between the surround channels and the left channel.

Monitoring Surround-Sound Audio

► Application Note



► **Figure 31.** Dolby Digital (AC-3) status display from WVR series.

Dolby Digital Status Display

Figure 31 shows the Dolby Digital (AC-3) status display from the WVR series. The extended Bitstream information show that the preferred Downmix is Lt/Rt and with this output the Center channel should be attenuated by -3dB. The surround channels should be attenuated by -6dB to produce the desired Downmix. In this example, if the Downmix was a Lo/Ro it would use the same attenuation factors for the center and surround channels. There is a wide array of informational parameters within Dolby metadata as shown in Figure 31.

Values in green indicate active allowed values for the metadata parameters. A yellow value indicates that the parameter is inactive or not indicated within the

metadata display. A grayed-out value indicates that this information is not present and Not Available (N/A) within the data stream.

Channel Mode

This provides information on the active channels, and how they should be handled by the encoder and decoder. It is typically displayed as two values separated by “/”. The first value indicates the number of front channels (Left, Center and Right) and the second number indicates the number of rear channels (S, Ls and Rs). Table 4 shows the possible configurations allowed and the appropriate data rates used to convey the specific number of channels.

Channel Mode	Data Rate	Definition
1+1		Dual mono (not valid for DTV broadcast or DVD production)
1/0	From 56kbps usually 96kbps	Mono
2/0	From 96kbps usually 192kbps	Stereo
3/0	From 256kbps	(L, C, R)
2/1	From 256kbps	(L, R, S)
3/1	From 320kbps	(L, C, R, S)
2/2	From 320kbps	(L, R, Ls, Rs)
3/2	From 384kbps usually 448kbps	(L,C, R, Ls Rs)

► **Table 4.** Channel Modes

LFE Channel

This indicates whether or not the LFE channel is present within the bitstream. Within the WVR series, this parameter is indicated by “L” within the channel mode parameter. The in-bar display for the LFE channel will show “DISABLED” if this channel is missing.

Dolby Source indicates which audio channels (embedded or AES/EBU input) are being used to decode the Dolby stream. This is not a specific metadata parameter, but is used as an identification of the channels being used by the WVR series.

Dolby Data Rate indicates the current data rate of the Dolby encoded stream. This can range between 56kbps – 620kbps for Dolby Digital (AC-3).

Dolby Sample Rate provides information on the sample rate of the audio signal. For broadcasting, this is typically 48 kHz as shown in Figure 31.

Bitstream Mode

This describes the audio services contained within the Dolby stream.

Bitstream Mode	Definition
Complete Main (CM)	All audio channels are present to form a Complete Main audio program. This is the most common setting for this bitstream mode. It can contain from one to 5.1 channels.
Main Music & Effects (ME)	This mode contains no dialogue channels but all Main Music and Effects are present. This can be used to support multiple languages where the dialogue channels can be carried as associated services with a single ME service.
Associated Visual Impaired (VI)	This is a single channel program which carries a descriptive narrative of the picture content for the Visually Impaired. The VI service maybe a complete mix of all audio channels present.
Associated Hearing Impaired (HI)	This is a single channel program which carries audio that allows the Hearing Impaired to more clearly understand the program material. The HI service maybe a complete mix of all audio channels present.
Associated Dialogue (D)	This is a single channel program which is intended to provide the dialogue channel for an associated Music and Effects service.
Associated Commentary (C)	This is a single channel program, which is intended to provide additional commentary that can be optionally decoded along with the main audio service. The C service maybe a complete mix of all audio channels present.
Associated Emergency (E)	This is a single channel service that is given priority for an Emergency announcement. In this case the main service is muted.
Associated Voice Over (VO)	This is a single channel service intended to be decoded and mixed to the Center channel. It requires a special decoder.
Main Service Karaoke (K)	This is a special service for Karaoke playback. In this case the Left and Right channels contain the music and the Center channel has a melody track. The Left surround and Right surround may provide an optional backing track.

► **Table 5.** *Bitstream modes.*

RF Over-modulation Protection

This is used to provide over-modulation protection when the Dolby stream is RF modulated. Typically this parameter is disabled as is shown in Figure 31. In a case where it is enabled, the Dolby encoder applies pre-emphasis in its RF mode compression.

Center Mix Level

This parameter provides information on the amount of attenuation to apply to the center channel in order to provide suitable mix for a left and right speaker system. In some cases if there is no extended bitstream information (BSI) then the Dolby decoder will use this parameter to produce the Lo/Ro Downmix.

Surround Mix Level

This parameter provides information on the amount of attenuation to apply to the surround channels in order to provide a suitable mix with respect to the front channels for a left and right speaker system. In some cases if there is no extended bitstream information (BSI) then the Dolby decoder will use this parameter to produce the Lo/Ro Downmix.

Dolby Surround Mode

This parameter tells the decoder that a two channel program stream contains a Lt/Rt Dolby Pro Logic coded signal. In this case, the decoder could choose to switch on the Dolby Pro Logic decoder to provide a surround sound output. In Figure 31 this value shows a grey-out value “N/A”, which tells the user that this parameter is not available within the metadata stream.

Audio Production Information

This value indicates the presence of Mixing Level and Room Type information. This is not shown directly on the WVR series to simplify the Dolby Status display. If present the Mixing Level and Room Type information will be displayed.

Mixing Level

This parameter indicates the peak sound pressure level (SPL) used during the final mixing session in the production process. This allows the Dolby receiver to set its volume control such that the mixing room and the listening environment are matched. This value can range from +80dB to +111dB.

Room Type

This parameter describes the equalization used during the final mixing session in the production process. There are three allowed values - Not Indicated, Large and Small. A large room is a dubbing stage with industry standard X-curve equalization. A small room is an environment which has flat equalization. This allows the Dolby receiver to set its equalization such that the mixing room and the listening environment are matched.

Copyright Bit

In this case, the value indicates the Dolby Bitstream is copyright protected. This value has no effect on the Dolby decoder.

Original Bitstream

This value indicates whether this is the original bitstream or a copy. This value has no effect on the Dolby decoder.

Extended Bitstream Information Parameters

This information was added after the original definition of ATSC A/52 to allow for more choice in the Downmix to a stereo program. Older Dolby receivers may not use these parameters but both older and newer systems are still compatible with either bitstream.

Preferred Stereo Downmix Mode

This parameter allows the operator to select a preferred Downmix selection which can automatically be used by the Dolby receiver. However consumers can still override this selection in their listening environment. There are three allowed selections for this parameter - Not Indicated, Lt/Rt Preferred, Lo/Ro Preferred.

Lt/Rt Center Downmix Level

This parameter indicates the attenuation to the center channel when adding this signal to the Left and Right channels to produce a suitable stereo Lt/Rt Downmix.

Lt/Rt Surround Downmix Level

This parameter indicates the attenuation to the surround channels when adding this signal to the Left and Right channels to produce a suitable stereo Lt/Rt Downmix.

Lo/Ro Center Downmix Level

This parameter indicates the attenuation to the center channel when adding this signal to the Left and Right channels to produce a suitable stereo Lo/Ro Downmix.

Lo/Ro Surround Downmix Level

This parameter indicates the attenuation to the surround channels when adding this signal to the Left and Right channels to produce a suitable stereo Lo/Ro Downmix.

Surround EX mode

This parameter indicates that the Dolby stream has been audio encoded in Surround EX™. An additional audio channel has been embedded in the Ls and Rs channels to provide more back surround channels for 6.1 or 7.1 systems. Within the WVR series the Dolby decoder mode can be selected as EX. This will allow two more audio bars to be displayed (Lb and Rb), which will show this additional coded audio data if EX coding is present.

Headphone Mode

This parameter indicates that the Dolby stream is capable of producing a Dolby Headphone™ surround from the two channels.

A/D Converter Type

This parameter allows audio that has passed through an HDCD (High Definition Compatible Digital) Analog to Digital conversion process to be indicated so that a

complementary decode through a suitable Digital to Analog converter can be used.

The following information values are not displayed within the WVR series Status display.

DC Filter – Indicates whether or not a DC-blocking 3Hz filter is applied

Lowpass Filter – Indicates whether or not low pass filter above 20kHz is applied

LFE Lowpass Filter – Indicates whether or not a 120Hz low pass filter is applied

Surround 3dB Attenuation – Indicates whether or not the surround channels have been attenuated by 3dB before encoding

Surround Phase Shift - Indicates whether or not a 90 degree phase shift has been applied to the surround channels. This allows the creation of a simple Lt/Rt Downmix.

Dolby Digital (AC-3) vs. Dolby E

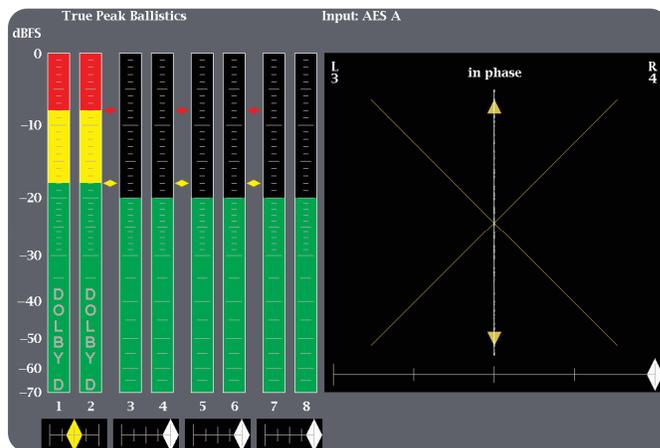
The Dolby Digital (AC-3) format has been designed for transmission and consumer applications and is not suitable for post-production applications. Dolby Digital (AC-3) is heavily compressed, therefore making it suitable for only one decoder by the consumer. Multiple encode and decode process can lead to distortions within the audio signals. Dolby E is designed for light compression and can be encoded and decoded multiple times. The frame size of a Dolby Digital (AC-3) stream is 32ms while the frame size of a video frame at 29.97fps is 33.34ms. This means that frame accurate editing is not practical with a Dolby Digital (AC-3). There is also a large latency of six frames required for decoding of a Dolby Digital (AC-3) signal which further complicates the post production issues. With a low latency of one frame and video frame-accurate packets sizes for Dolby E, means that it is more suitable for post-production than Dolby Digital (AC-3).

Configuring Audio Bars

Within the production processes there are currently no standards on how the AES/EBU and Dolby data streams should be assigned to the various available audio channels. Each post production facility and broadcasters can define their own requirements for which channels carry which data streams. This can lead to confusion when you are editing a program or when you want to configure your audio monitor to quickly troubleshoot a problem. Therefore, there are a few simple steps to remember in finding out how your audio channels are configured.

On the WVR series, select the audio measurements for one of the four tiles within the instrument display. Then press and hold either the audio button or the appropriate tile button to bring up the audio menu. Scroll through the menu to Audio Input item and select the appropriate audio input - AES A, AES B or Embedded. The audio display should now show all eight audio channels. If not, some audio channels may have been de-selected in the audio configuration menu and may need to be re-enabled so that you can see all present audio channels as shown in Figure 32. If any Dolby stream is present the audio level bars will be at full scale and the in-bar indicators will alert the user to the presence of either Dolby E or Dolby D data streams within the audio programs. If the Lissajous display is selected for the Dolby input it will show a clipped phase display, similar to Figure 10. You may wish to save this set-up as a preset within the WVR series instrument so that you can quickly access this configuration in the future. To do this just push and hold one of the preset buttons until the display indicates that the preset has been saved.

Now that the user has determined the presence of a Dolby stream it is possible to decode this data in the Dolby options of the WVR series. In this case Figure 32 shows that Dolby D is present on AES input 1-2. The user needs to select one of the Dolby configurations (1-4) in order to allow decoding of the data stream within the audio display. To setup each of these Dolby configurations enter the “Config” menu and scroll



► **Figure 32.** Audio Display showing AES/EBU and Dolby D signal present.

through to select the Audio Input Output menu item. Within this menu item are the selections for each of the Dolby configurations. Each Dolby configuration allows for selection of the specific audio input source to be used for decoding of the Dolby stream. In this example the user selected Dolby 1 and configured this menu to use AES1-2 for the Dolby Source as shown in Figure 33. Up to ten audio bars can be displayed depending on the Dolby format, for instance a Dolby E 5.1 + 2 program will show the eight audio bars plus, if configured, the downmix of the 5.1 program to a mono or stereo signal giving a total of ten bars.

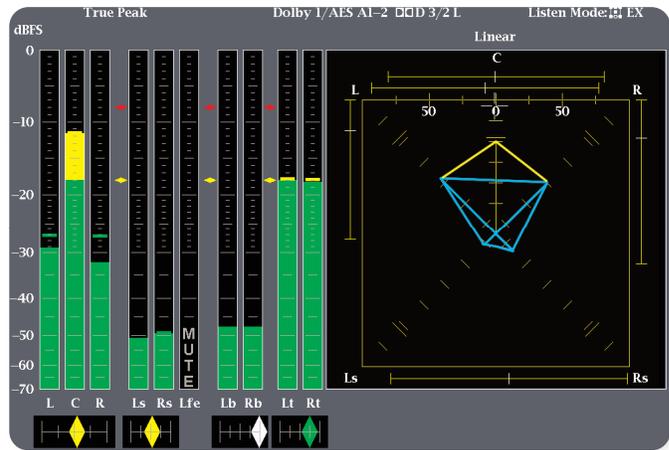
For Dolby Digital (AC-3) data streams there are a variety of decode modes available in the WVR series. A data stream which has been EX coded during the post-production process, allows support for 6.1 or 7.1 surround systems. If the EX decoding mode is selected then the decoder will produce a 7.1 channel output (L, C, R, LFE, Ls, Rs, Lb, Rb). If the Dolby stream contains only a 5.1 channel mix then the full decode selection will display a standard six audio bar configuration (L, C, R, LFE, Ls, Rs). Optionally for each of these decode modes a downmix can also be done to produce a bar display of Lt/Rt, Lo/Ro or mono. Therefore a maximum of ten audio level bars can be displayed from a Dolby EX coded stream as shown in Figure 33.

Configuring Audio Outputs

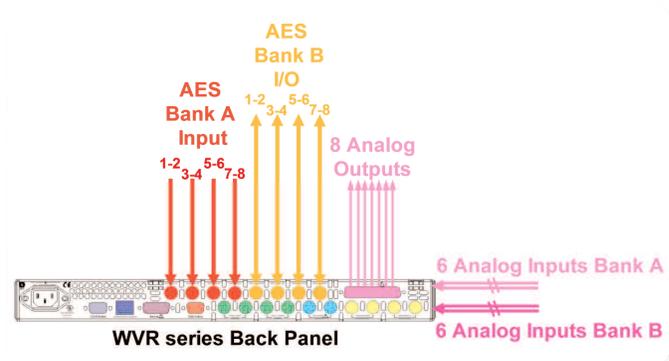
Within the WVR series option DDE, there are a wide variety of audio configuration options available to suit the requirements for decoding or passing through the audio data stream. Figure 34 shows the input and output connections available on the WVR series. AES Bank A is the primary digital audio input supporting up to 4 inputs (8 channel pairs). AES Bank B can be configured as an input or output supporting up to 4 inputs or outputs (8 channel pairs). When AES Bank B is configured as an output it can act as an active loop through of the AES A audio signals or can de-embedded the audio from an SDI video input. A simple cross point matrix in the audio input/output configuration menu allows for mapping of the various inputs and outputs of the audio signals.

When a Dolby signal is decoded in the WVR series DDE option, a full decode of the audio channels can be routed to the AES Bank B output or to the eight analog audio outputs. Mapping of the various audio channels can be done in each of the Dolby configurations (1-4). This allows the user to configure the system for multiple different configurations. Each of these configurations can be saved within the presets of the WVR series, allowing a possible twenty different setups for the device. A limited decode of the Dolby stream to two channels is only possible in the WVR series DD option. The user can select any audio pair to be output in this case, for example Lt/Rt.

For those facilities which do not have a full surround sound listening environment the user can also choose to just output a stereo channel pair of either the down-mix or from the selection of the channels used by the Lissajous display. The operator can also use flexible Lissajous display to customize the individual channels for comparison.



► Figure 33. Dolby Digital (AC-3) display with ten audio bars selected.



► Figure 34. WVR Series back panel audio inputs and outputs.

There are many ways in which the audio inputs and outputs can be configured for your facility. Once you have defined your appropriate settings it is important to ensure these configurations are saved as presets within the instrument. These presets can also be saved onto a PC for archive purposes by using a Java application available at the Tektronix website (www.tektronix.com). Detailed instruction of how to save these presets can be found within the latest WVR series manual.

Monitoring Surround-Sound Audio

► Application Note

Audio Session										
Audio Input:	Dolby 1				Data Collect:			Running		
Signal Loss:					Runtime:			0 d, 00:19:29		
Analog Output	1&2: L, R		3&4: C, Lfe		5&6: Ls, Rs		7&8: Lb, Rb			
AES B Output	1-2: L, R		3-4: C, Lfe		5-6: Ls, Rs		7-8: Lb, Rb			
Channel	L	R	C	Lfe	Ls	Rs	Lb	Rb	Lt	Rt
Clip(s)	0	0	0	0	0	0	0	0	0	0
Over(s)	26	23	59	0	0	0	0	0	0	0
Mute(s)	0	0	0	4278	0	0	0	0	0	0
Silence(s)	0	0	0	34	16	20	11	11	0	0
Peak (db)	0.18	0.09	0.00	-5.69	-7.09	-8.77	-12.89	-12.89	-5.39	-6.37
High (db)	0.24	0.10	0.00	-5.73	-7.15	-8.79	-12.91	-12.91	-5.40	-6.43
Active bits	24	24	24	24	24	24	24	24	24	24

► Figure 35. Audio Session display of WVR series.

Error Status Log				Page 2 of 48
Error Status	Timecode	Date	Time	
✓ Audio Silence (----4-----)	00:01:42:05:1		14:47:02	
✓ Audio Silence	00:01:45:06:2		14:47:05	
✓ Audio Silence (----4-----)	00:01:46:06:2		14:47:06	
✓ Audio Mute (----4-----)	00:01:50:24:1		14:47:10	
✓ Audio Silence	00:01:51:06:2		14:47:11	
✓ Audio Over Level (---3-----)	00:01:53:08:1		14:47:13	
✓ Audio Over Level	00:01:55:08:1		14:47:15	
✓ Audio Mute	00:01:55:24:1		14:47:15	
✓ Audio Over Level (---3-----)	00:01:58:23:2		14:47:18	
✓ Audio Over Level	00:02:00:23:2		14:47:20	
✓ Audio Mute (----4-----)	00:02:01:15:2		14:47:21	
✓ Audio Mute (----4-----)	00:02:01:23:1		14:47:21	
✓ Audio Silence (----4-----)	00:02:03:08:2		14:47:23	
✓ Audio Mute	00:02:04:23:1		14:47:24	
✓ Audio Over Level (---3-----)	00:02:05:08:2		14:47:25	
✓ Audio Silence	00:02:05:08:2		14:47:25	
✓ Audio Over Level (123-----)	00:02:06:08:2		14:47:26	
✓ Audio Over Level	00:02:07:08:2		14:47:27	
✓ Audio Mute (----4-----)	00:02:07:12:1		14:47:27	
✓ Audio Mute (----4-----)	00:02:07:14:2		14:47:27	
✓ Audio Silence (----4-----)	00:02:08:20:1		14:47:28	
✓ Audio Mute	00:02:18:13:1		14:47:38	
✓ Audio Silence	00:02:18:20:1		14:47:38	
✓ AES Frame Sync Error (123-367890)	00:02:19:00:1		14:47:38	

Arrow Left – Previous, Right – Next, Up – First, Down – Last.

► Figure 36. Error log of WVR series.

Audio Session Display and Error Log

After the final program has been assembled or during the ingest of material, it is important to check quality to ensure there are no audible errors. With multiple audio channels to monitor it becomes challenging to ensure that no audio errors are present within the program material. The WVR series and WFM700 instruments have several parameters which can be automatically monitored for errors such as Clips, Mutes, Silence and

Overs. These errors are reported in the Audio Session display and the event can be recorded in the error log. They can be linked to the timecode of the program material allowing the operator to quickly go back through the material, to a specific event where the error occurred. In this way the operator can investigate the problems within the program and determine if these errors need to be fixed within the material.

Event	Event Description	SMPTE Timecode	TimeOfDay
AlarmStart	Audio Silence (---4-----)	00:01:42:05:1	14:47:02
AlarmEnd	Audio Silence	00:01:45:06:2	14:47:05
AlarmStart	Audio Silence (---4-----)	00:01:46:06:2	14:47:06
AlarmStart	Audio Mute (---4-----)	00:01:50:24:1	14:47:10
AlarmEnd	Audio Silence	00:01:51:06:2	14:47:11
AlarmStart	Audio Over Level (--3-----)	00:01:53:08:1	14:47:13
AlarmEnd	Audio Over Level	00:01:55:08:1	14:47:15
AlarmEnd	Audio Mute	00:01:55:24:1	14:47:15
AlarmStart	Audio Over Level (--3-----)	00:01:58:23:2	14:47:18
AlarmEnd	Audio Over Level	00:02:00:23:2	14:47:20
Alarm	Audio Mute (---4-----)	00:02:01:15:2	14:47:21
AlarmStart	Audio Mute (---4-----)	00:02:01:23:1	14:47:21
AlarmStart	Audio Silence (---4-----)	00:02:03:08:2	14:47:23
AlarmEnd	Audio Mute	00:02:04:23:1	14:47:24
AlarmStart	Audio Over Level (--3-----)	00:02:05:08:2	14:47:25
AlarmEnd	Audio Silence	00:02:05:08:2	14:47:25
AlarmStart	Audio Over Level (123-----)	00:02:06:08:2	14:47:26
AlarmEnd	Audio Over Level	00:02:07:08:2	14:47:27
Alarm	Audio Mute (---4-----)	00:02:07:12:1	14:47:27
AlarmStart	Audio Mute (---4-----)	00:02:07:14:2	14:47:27
AlarmStart	Audio Silence (---4-----)	00:02:08:20:1	14:47:28
AlarmEnd	Audio Mute	00:02:18:13:1	14:47:38
AlarmEnd	Audio Silence	00:02:18:20:1	14:47:38
AlarmStart	AES Frame Sync Error (123-567890)	00:02:19:00:1	14:47:38

► **Table 6.** Event Log View imported into spreadsheet application.

Figure 35 shows the parameters monitored in the audio session and a summary of the number of occurrences of any errors. In order to determine where these errors occur within the material the user can display the Error Log from the status measurements menu as shown in Figure 36.

The Audio Session not only provides summary information of Clips, Mutes, Silences and Overs, but information on the peak and high audio levels that have been applied to each input during the monitoring of these signals. It also provides summary information on how the audio inputs and outputs are currently configured.

The error log shown in Figure 36 provides more detailed information on when the various errors occurred within the program material and at what timecode value and

internal time of the instrument that these errors occurred. Red indicates the start of an error condition and green indicates the end of an error condition. This error log can be downloaded from the instrument over an IP/Ethernet connection to the instrument. Simply enter the IP address of the instrument from a web browser application and then select View Event Log from the menu. This file can then be saved as a text or html file and imported to a spreadsheet program or printed and attached to a work document for further post-production. Table 6 shows the Event Log downloaded from the instrument and imported into a document. This provides the same information that is present from the error log on the display of the instrument as shown in Figure 36.

Conclusion

The emergence of surround sound audio in film, television and home entertainment has created the need for monitoring tools that help audio engineers visualize the critical level and phase relationships need to create a high quality surround sound experience. A complete monitoring tool set will include the traditional level bar and phase displays, displays that show the metadata content of digital signals and specialized displays that show the combined effect of the different channels in multi-channel audio content. These displays help audio production and post-production engineers accurately and efficiently record, mix and master multi-channel audio content. For broadcast operations and engineering staff, these solutions help them quickly detect and isolate problems in multi-channel audio signals and reduce the time and effort needed to diagnose and fix these problems.

Contact Tektronix:

ASEAN / Australasia / Pakistan (65) 6356 3900
Austria +41 52 675 3777
Balkan, Israel, South Africa and other ISE Countries +41 52 675 3777
Belgium 07 81 60166
Brazil & South America 55 (11) 3741-8360
Canada 1 (800) 661-5625
Central East Europe, Ukraine and Baltics +41 52 675 3777
Central Europe & Greece +41 52 675 3777
Denmark +45 80 88 1401
Finland +41 52 675 3777
France & North Africa +33 (0) 1 69 86 81 81
Germany +49 (221) 94 77 400
Hong Kong (852) 2585-6688
India (91) 80-22275577
Italy +39 (02) 25086 1
Japan 81 (3) 6714-3010
Luxembourg +44 (0) 1344 392400
Mexico, Central America & Caribbean 52 (55) 56666-333
Middle East, Asia and North Africa +41 52 675 3777
The Netherlands 090 02 021797
Norway 800 16098
People's Republic of China 86 (10) 6235 1230
Poland +41 52 675 3777
Portugal 80 08 12370
Republic of Korea 82 (2) 528-5299
Russia & CIS 7 095 775 1064
South Africa +27 11 254 8360
Spain (+34) 901 988 054
Sweden 020 08 80371
Switzerland +41 52 675 3777
Taiwan 886 (2) 2722-9622
United Kingdom & Eire +44 (0) 1344 392400
USA 1 (800) 426-2200

For other areas contact Tektronix, Inc. at: 1 (503) 627-7111

Last Updated June 15 2005

For Further Information

Tektronix maintains a comprehensive, constantly expanding collection of application notes, technical briefs and other resources to help engineers working on the cutting edge of technology. Please visit www.tektronix.com



Copyright © 2005, Tektronix, Inc. All rights reserved. Tektronix products are covered by U.S. and foreign patents, issued and pending. Information in this publication supersedes that in all previously published material. Specification and price change privileges reserved. TEKTRONIX and TEK are registered trademarks of Tektronix, Inc. All other trade names referenced are the service marks, trademarks or registered trademarks of their respective companies.

7/05 FLG/WOW

25W-17981-1

Tektronix
Enabling Innovation