Surround Sound

in Broadcasting Stations

Surround Sound Systems

By Karl M. Slavik

Radio Guangdong in China, XM-HD Satellite Radio in the US or Germany's ARD Radio Network - they all belong to the slowly but ever increasing number of radio broadcasters transmitting in 5.1 surround. Although 5.1 might be less of an option for the average AM or FM pop, rock, news or talk station, it definitely is an enhancement for jazz, classical and other "culture" stations on DAB, HD-Radio (formerly IBOC) or the many forms of DVB - not to forget surround sound streaming via the World Wide Web.

Even radio broadcasters currently not on air in 5.1 are starting to produce and archive a growing part of their content as multichannel audio - e.g. live recordings of concerts for later release on DVD, DVD-Audio or SACD. Concerts, radio dramas, cabaret and even stand-up comedy and sports programmes are some of the content that definitely profit from surround sound - on air as well as in the DVD or at the download site. But before getting into surround sound broadcasting, one should have a good understanding of the basic principles and a thorough grip on how to solve some very special surround sound issues.

Mono, Stereo, Surround

Every spatial cue perceived by the human auditory system is based on sound pressure and arrival time differences between our two ears. As monaural production and transmission uses only one audio channel, no spatial information can be conveyed. Stereophonic production and transmission requires a minimum of two separate audio channels. If carefully recorded and reproduced it provides a good amount of spatial information which allows for precise stereo imaging between the left and right loudspeaker. Due to its very nature there is little or no spatial information outside the stereo panorama, leaving the listener somewhat "outside" of the performance. Current surround sound formats try to immerse the

listener in a two-dimensional, horizontal field of sound. The term "multichannel stereo" acknowledges the fact that every possible multichannel configuration is finally perceived by our two ears - the human stereophonic auditory system. Therefore, all surround sound systems and channel configurations from 3.0 to 7.1 or even higher can be termed as "multichannel stereo".

Channel Configuration or Format

The channel configuration or channel format describes the number of audio channels involved and their arrangement in the listening room. The number ahead of the point (e.g. in 2.0 or 5.1) indicates the number of full range audio channels; the number after the point stands for the number of LFE channels (LFE … Low Frequency Effects). Therefore 5.1 stands for five full range (20 – 20 kHz) audio channels and one band limited (20 - 120 Hz) low frequency effects channel. By the way: Coming from the world of cinema with its rumbling bomb explosions and dinosaur stomps, the LFE was never intended to be used for music - just for low frequency sound effects. The highly informative document "5.1-Channel Music Production Guidelines" from Dolby (available as a free download) clearly states that the LFE should not be used for music and if only with utmost caution. One reason for that is clear: When down mixing 5.1 surround to stereo (2.0) or mono (1.0), most surround receivers (or notebooks, playing a DVD) simply leave out the LFE channel to protect the mostly smaller speakers of such systems.

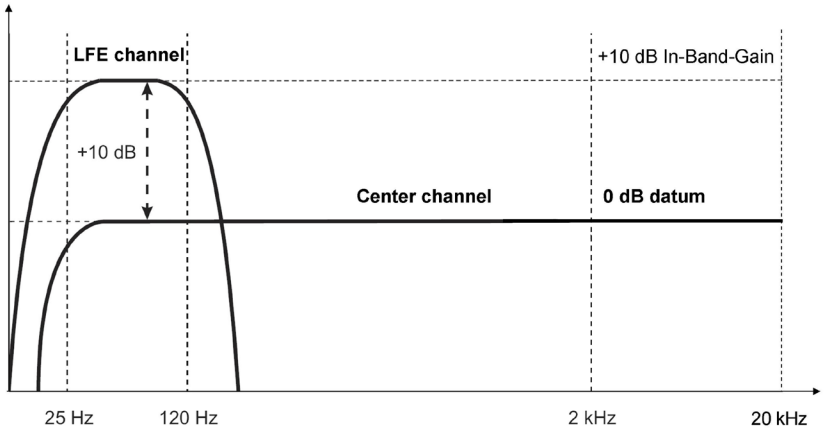
Speaker Setup and Alignment

Studio monitors are for the audio engineer what accurate navigation tools are for a pilot. Without fairly accurate monitoring, surround sound mixing is like flying blind. Document ITU-R BS.775-1 (and others) describes the setup of 5.0 or 5.1 surround speaker systems. All speakers should be arranged along the perimeter of a circle, aiming at the engineers seating position in the centre (sweet spot, see drawing). If one or more speakers are at different distances from the centre, delays have to be used to time align the setup. Ideally, all speakers should be full range systems of the same type. If not full range, a subwoofer (which is mostly available to reproduce the LFE channel) can be used to reproduce the low frequencies of the other speakers (by using a built-in multichannel crossover & summing device). By using this type of "bass management", the subwoofer is fed with the LFE signal AND with the summed low frequency content of the other channels.

If fed with a test signal (e.g. pink noise) all five speakers should

reproduce almost identical sound pressure levels, at least within a frequency range of 100 – 10 kHz, with deviations of less than ± 0.5 dB between each other. The playback level of the LFE speaker should be 10 dB higher within the frequency range of 20 to 120 Hz. This "In-Band-Gain" (see drawing) is employed in home theatre receivers as well and enables the single LFE to reproduce the same acoustic energy as the other five full range channels together.

XL2 with M4260 measurement microphone



Level alignment of LFE and surround sound speakers

The sound pressure level of the full range speakers should be measured with a C-weighted SPL-meter, whereas the In-Band-Gain of 10 dB should be measured with an RTA (Real Time Analyzer) by comparing the individual response of the five main speakers to the response of the LFE. Handheld digital instruments, such as the XL2 from NTi Audio, are perfectly suited for the job. This one not only supports SPL-measurements with different weighting characteristics (Flat, A, C, X-Curve, RLB) and Leq, but also functions as a RTA- and FFT-analyzer with the possibility to store all data for later processing on an external computer. As a signal generator, the NTi Audio Minirator MR-PRO could be used. It allows for the generation of sine waves with freely selectable frequencies, sweep signals with any frequency interval of up to 1/12 octave, white and pink Noise, as well as polarity and delay test signals. Furthermore it can play back audio files (WAV) from its internal RAM.

Coding Formats - Pro and Consumer

Digirator DR2

The term coding format or surround sound delivery system describes the method that is used to encode and decode the multichannel audio - hence the term "codec". Codecs range from analogue or digitally implemented matrix based systems (Dolby ProLogic, ProLogic II and SRS Circle Surround) to "quasi-discrete" codecs (e.g. MPEG Surround, Neural Surround and DTS-AAC) on to fully discrete systems such as Dolby Digital, Dolby Digital Plus and DTS. Practically all consumer formats or "transmission bit streams" mentioned above are optimized for high coding efficiency which translates into low bandwidth requirements during transmission or little space requirements when stored. Nevertheless, the inherent audio quality of these formats or "bit streams" can be truly excellent, as long as no multiple coding/decoding takes place.

Modern audio production and distribution techniques sometimes require a large number of encoding and decoding cycles to cope with the multiple steps of recording and post production, often performed in different locations. To keep audio free from coding related artefacts only mild data reduction (~ ratio of 4:1 or 5:1) and relatively high bitrates should be used for multichannel audio distribution in the professional domain. The same is true for archiving and - maybe to a lesser degree - for playback purposes at broadcast stations. The bit streams used for professional applications - e.g. when connecting an OB-van to the studio - are therefore called "production and distribution bit streams". This indicates their ability to withstand multiple coding/decoding cycles without as much signal degradation. Often they support any given bit depth of the transport path (audio resolution, 16 to 24 bit) and every possible channel configuration (mostly Mono 1.0, Stereo 2.0 up to Surround 5.1 + Stereo Downmix 2.0 simultaneously). Professional coding formats such as Dolby E or APT WorldNet Oslo (apt-x) use approximately five times the data rate of Dolby Digital (384 kbit/s) and about 20 times the data rate of MPEG Surround (96 kbit/s). This allows for at least 10 to 15 coding/decoding cycles (opposed to a maximum of three with Dolby Digital or only one with MPEG Surround).

Digirator DR2 from NTi Audio supports linear PCM, Dolby E, Dolby Digital, Dolby ProLogic II and DTS. As all nonlinear, encoded signals are stored within WAV-files, it can be easily upgraded for future multichannel formats that have to be transported over AES/EBU connections. As the name implies it is a mainly a digital generator but with some highly practical analyzer functions as well. An internal low jitter clock generator can be synchronized to AES3, DARS, Word Clock and Video signals. Measurements of channel transparency, channel propagation delay and sample frequency are supported. Outputs are AES3 (XLR, 110 Ohms), AES3-

id (Cinch, 75 Ohms) and TOS-Link which doubles as an 8-channel ADAT-Interface. Input is AES3 (XLR, 110 Ohms) with an adapter for AES3-id (75 Ohms) provided.

Transparency and Latency

Each time that audio signals must be sent across a given, but rather unknown channel infrastructure (such as a digital network or a satellite link) it makes good sense to check for latencies and bit transparency. This is especially important for all coded multi-channel signals.

Due to their very nature, all digital encoding and decoding processes introduce latencies (delay) which largely depend on the bit rate efficiency of the coder. As a rule of thumb one should remember: The lower the resulting bit rate (band width), the higher the latency. The latency is typically in a range from approximately 4 to 600 Milliseconds (e.g. 40 ms with "high bit rate" Dolby E for professional applications, 179 ms with "low bit rate" Dolby Digital for transmission). These latencies are sometimes further increased by the propagation delays of digital networks and satellite links.

Matrix encoded surround signals such as ProLogic II and SRS Circle Surround can be transported via nearly any two channel connection. Nevertheless, a high degree of phase stability between the two channels is mandatory. Phase distortion is not only introduced by "old" analogue paths (e.g. multipath distortion in FM broadcasting) but also by most data reduction algorithms running at low data rates (mostly less than 192 kbit/s, e.g. with ISO-MPEG 1/2 Layer II or III). As a stable phase relation between the channels is crucial for correct matrix decoding, data reduction often leads to heavy phase distortion that produces highly audible artefacts after surround decoding (chirp or twitter effects). The same is true for many "sound processors" that deliberately use phase differences between channels to achieve a "wider" stereo image or other effects.

Most digitally encoded surround sound signals such as Dolby Digital, Dolby E or DTS, can be transported via standard AES/EBU connections, so long as the AES-subframe containing the audio payload is not changed (see drawing). Every system that passes a digital signal without any alteration of its code can be considered as "bit transparent". To measure the latency of any encoder-decoder-path as well as bit transparency, an advanced instrument like the DR2 from NTi Audio can be used. Besides a comprehensive set of audio test signals, the DR2 also supports surround sound test sequences for verification and adjustment of professional Dolby Digital, Dolby ProLogic II, Dolby E and DTS installations. These test sequences are stored within the DR2 memory. Transparency

and latency measurements are automated sequences that allow for easy and quick quality checks of audio links.

Metadata

Metadata is "data about other data" and accompanies the multichannel audio signal. Practically all digital coding formats - from MP3 to Dolby E, DTS and MPEG-2 TS (Transport Stream) contain some form of metadata. In Dolby Digital and Dolby E metadata is used to optimize home playback under the most diverse conditions by e.g. keeping a constant volume level when switching between programmes or getting the right down mix and the best possible dynamic range.

Informational or descriptive Metadata describe the bit stream, but do not influence it. Examples are the name of a song, a programme or a TV channel, the given bit rate per second, the channel configuration, etc. Control Metadata goes one step beyond by controlling encoders and decoders along the signal path. They contain parameters such as values for Dialogue Normalisation (DialNorm), Dynamic Range Control (DynRange, DRC), the Downmix Coefficients, activation of filters and phase shifts, and much more.

It has to be noted that the relative complexity of Dolby metadata is a big stumbling block for some audio engineers. Numerous productions are on air every week or sold as DVDs that contain totally incorrect metadata or only "default" metadata, leading to a less than perfect playback performance. All Dolby and DTS sequences stored in the NTi Audio DR2 have correct metadata, enabling engineers to compare their own work with these given data.

Recommended reading and web links:

- Spatial Audio. Francis Rumsey (Focal Press / Elsevier)
- The Art of Digital Audio. John Watkinson (Focal Press / Elsevier)
- 5.1-Channel Music Production Guidelines (Dolby)

- www.aptx.com
- www.axiaaudio.com/surround/
- www.dolby.com
- www.dts.com
- www.mpegsurround.com/
- www.neuralaudio.com/
- www.srslabs.com/

The author:

Karl M. Slavik is senior consultant / owner at ARTECAST, an independent consulting, training and media project operation based in Vienna, Austria. He's a regular lecturer at the ARD-ZDF Media Academy in Nuremberg (Germany) and at the University of Applied Sciences St. Poelten (Austria). He is a Dolby-approved trainer and consultant and can be reached via mail at k.m.slavik@artecast.com.