Multichannel Audio Technologies: Lecture 7

Mastering for Surround Sound

Surround sound mastering is a highly specialized skill which differs significantly from stereo mastering. Although the purpose of mastering in surround is the same as that in stereo — to maximize the musical content inherent in the mix — it requires a somewhat different approach, as well as specialized equipment such as multichannel dynamics processors with flexible channel linking. Mastering engineers are also sometimes called upon to add reverb or other effects to a mix. In those cases, it is important to have access to equipment that can deliver both correlated and decorrelated multichannel processing. In addition, the monitoring setup in a surround mastering facility is considerably more demanding, since the mastering engineer has to listen even more closely for phase anomalies.

Another important difference between surround sound mastering and stereo mastering — especially those stereo releases targeted for radio play — is that it is not necessary to highly compress the mixes in order to have them stand out. In the world of surround sound, compression should only be used to make the mix more exciting, not to raise the overall level. Because there are six or more speakers doing the work instead of just two, maximum dynamic range can and should be preserved. Similarly, mastering engineers also report that, in general, they need to apply less equalization to surround material than stereo material.

Even on projects destined for release on SA-CD or DVD-Audio, the surround mastering engineer should filter the LFE channel, despite the fact that the specification does not require such filtering. Limiting the frequency range of the LFE channel to only the bottom few octaves allows for more consistent playback in different environments. However, this does raise the spectre of potential latency/phase issues, since the main channels remain unfiltered. If the LFE channel is delayed even slightly, it can completely cancel redirected bass content from the main channels when monitored on a bass-managed system. Therefore, the surround mastering engineer must listen on a bass-managed system to check the results of combining the low frequency signals. In some cases, it may be necessary to time-shift the main channels in order to ensure that all channels remain in absolute perfect phase after the filtering process.

Time code issues can also present problems when mastering surround sound that has to be locked to picture (such as music concerts presented on DVD Video). For those projects, the mastering facility may have to serve almost as a postproduction house, with good quality video monitors and a source of extremely stable, jitter-free sync. Lastly, the job of encoding multichannel audio in Dolby Digital, DTS, MLP, and/or DST formats is often done by the surround sound mastering engineer, requiring additional critical listening and the presence of both encoding and decoding equipment (for playback monitoring purposes).
Surround Mastering in Cubase 4:

Cubase 4 offers a good environment for multichannel mastering due to its flexible routing options, multichannel plugins and high resolution support. To set up for mastering:

1. Add in a 5.1 surround sound bus in the VST connections window, and route the channels out your audio card accordingly.
2. Import the 6 audio files to six separate tracks and route them out directly to their associated busses.
3. On the 5.1 bus, add in, but don’t enable yet, a multichannel equalization tool, a dynamics processing tool, an image enhancer tool, a stereo mixdown tool, a bit quantizer and a commercial playback emulation tool.

Equalisation:

As stated previously, equalization is not performed in surround mixing to the same extent as it is in stereo mixing since the sources are no longer ‘battling for space’ as they were previously. Nevertheless, in a mastering situation, it may be relevant to change the overall tonal balance of the mix due to poor mix-room acoustics. In this case, special attention should also be paid to the low frequencies and the cross-over to the subwoofer. The eq’ing stage is a good opportunity to put a low pass filter on the LFE channel. This filter should be set somewhere in the region of 80-120Hz, and should have as linear phase as possible. The Waves LFE 360 is a good choice for this.

Unless there is a particular problem with one channel’s tonal balance, the equalizer should be on the main surround bus. As with conventional eq in stereo mastering, a good tonal balance across all frequencies is desirable. Take note whether the rear channels sound too dull or too bright, and always remember to A/B between mastering eq on and off. The Studio Eq in Cubase 4 is a parametric eq that supports surround sound.
Dynamic Range Compression:

Unlike conventional stereo, there isn’t the same requirement to compress the audio to the point of over saturation. The loudness wars have different rules in the surround world. Application of compression should be subtle and used to help maintain good level balance, without affecting the dynamic range. Remember also, the more compression you apply, the more you’ll affect the surround imaging. The Waves C360 is a good choice of surround compressor, as it allows us to apply the same compression settings to multiple channels.

Image Enhancement:

At this stage, we may notice that the surround imaging has deteriorated somewhat due to compression effects. It is then justifiable to use some form of surround image enhancement. S360° Imager can be used here as it lets you set the Rotation and Width for a surround mix.
Dithering

As always, when doing anything non-linear or when decreasing the bit-depth, dithering is crucial. Here we can use the IDR 360 Bit Quantizer on the surround bus, if we have to cater for 16 bit audio.
**Stereo Compatibility:**

Downmixing (sometimes called “fold-down”) refers to the process of automatically reducing a multichannel mix to two channels — i.e., 5.1 playback to stereo. A downmix is almost always inferior to an original stereo mix; however, because almost all consumer receivers include the feature, it is a necessary part of surround sound production. Both Dolby Digital and MLP encoding, as well as the WMA9 delivery format, provide for metadata settings known as *downmix coefficients* to describe how the 5.1 mix will be converted to stereo. Note that if no downmixing coefficients are specified, MLP will create a downmix consisting of only the L and R channels! With DTS encoding and the SA-CD format, no downmixing is possible; indeed, SA-CDs always include a dedicated stereo mix.

When encoding for Dolby Digital, MLP, or WMA9, the Center channel downmix coefficient should be set to -3dB in relationship to the Left and Right front channels. However, the rear channel coefficients are largely program dependent. Many of today's surround productions are mixed with an "ensemble" perspective. In those instances, the rear channels must be at the same level as the front left and right channels. However, the rear channels *may* be lowered -3dB if they contain only ambience or audience material. The stereo sum of the 5.1 mix will usually need to be lowered in level in order to avoid creating signals so loud that they create digital "overs" (distortion).

Typical settings for downmix coefficients are:

- **Left Front** = -6dB,
- **Right Front** = -6dB,
- **Center** = -9dB,
- **LFE** usually OFF or to taste (never usually more than -9dB),
- **Left Surround** = -6dB,
- **Right Surround** = -6dB.

Surround mixes should always be checked in this typical downmix configuration. The mix engineer should indicate to the mastering engineer whether these coefficients are suitable or whether any changes are necessary.

If surround sound material is Dolby Digital- or MLP-encoded, there is no way to completely avoid the possibility that a consumer might listen to a downmix instead of a true stereo mix. However, there are ways to reduce the chances of that occurring. One is to ensure that a separate dedicated stereo mix is included on the disk; in the case of legacy material being repurposed, this should preferably be the original stereo mix. The second is to specify to the mastering engineer that the disk be authored with the stereo mix as the default audio stream. This will, of course, require that the consumer proactively select "5.1" if they want to hear the surround sound mix. However, if this is not done - if the disk is instead authored with the 5.1 mix as the default - then consumers who have only two speakers attached to their receiver will hear the downmix and not the stereo mix, even if one exists.
We can check our mixes stereo compatibility using the **Mix 6 to 2 plugin**:
Mixdown:

As before, we should mixdown 6 independent audio tracks that correspond to the six surround channels. However, we now have to consider the final bit depth and resolution before encoding. The chart below is useful in determining this:

<table>
<thead>
<tr>
<th>Distribution format</th>
<th>Encoding format</th>
<th>Sample rate / Word length / Bit rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CD</strong> (Traditional “Red Book”)</td>
<td>2-channel discrete PCM (unencoded)</td>
<td>44.1kHz 16-bit Constant-rate, 1.411Mbps stream</td>
</tr>
<tr>
<td><strong>Single stream</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DVD-V</strong> Multiple streams:</td>
<td>2-channel discrete PCM (unencoded)</td>
<td>48 or 96kHz 16, 20, 24-bit Constant-rate, 6.144Mbps max stream</td>
</tr>
<tr>
<td><strong>Dolby Digital (AC-3)</strong></td>
<td></td>
<td>32, 44.1, 48kHz 16, 20-bit Constant-rate, 448kbps max stream</td>
</tr>
<tr>
<td><strong>DVD-A</strong> Multiple streams:</td>
<td>2-channel through 6- (5.1) channel discrete PCM (unencoded) or MLP</td>
<td>44.1kHz – 192kHz 16, 20, 24-bit, Variable-rate (MLP) or constant-rate (PCM), 9.6Mbps max stream</td>
</tr>
<tr>
<td><strong>Dolby Digital (AC-3)</strong></td>
<td></td>
<td>32, 44.1, 48kHz 16, 20-bit Constant-rate, 448kbps max stream</td>
</tr>
<tr>
<td>(contained in video area only)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DTS</strong></td>
<td></td>
<td>48 or 96kHz 16, 20, 24-bit Constant-rate, 1.5Mbps max stream</td>
</tr>
<tr>
<td>(contained in video area only)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SA-CD</strong> Single or multiple streams:</td>
<td>2-channel and 6-channel, single-bit, high-rate encoding.</td>
<td>Release format (64FS = 2.8224Mbps bit rate DSD format)</td>
</tr>
<tr>
<td><strong>Windows Media Audio 9,</strong></td>
<td>Various bit-rate reduction methods</td>
<td>From very low to reasonably high rates</td>
</tr>
</tbody>
</table>
Discrete and matrixed surround sound:

The next step in the mastering process is to choose the codec (if any) to use. Here, we first need to decide if we are mastering for ‘Discrete’ or ‘Matrixed’ surround sound.

In a matrixed surround sound system the discrete audio channels are matrixed down to fit onto 2 channels. For example, in a 4-2-4 system, four channels of discrete audio are matrixed into 2 channels using a matrix encoder, and then ‘expanded’ back to 4 channels using an appropriate decoder. Matrixed surround sound was a huge part of the Quadraphonic era, and many feel that the confusion of the different types of decoder led to its demise. However, matrixed surround still exists today, in particular in the most popular matrixed format: Dolby Surround Sound.

Matrixed Surround Sound Case Study: Dolby Surround

Dolby Surround was developed in the mid 70s amidst the confusion of the Quadraphonic era. Dolby took the idea of matrixed surround already common on vinyl and four track and applied it to the cinema. However, instead of taking the matrixed four channels, and translating two to the front and two to the back, they decided to have a layout consisting of front LCR and 1 rear channel (a 3-1 setup). In cinemas the mono rear channel would often be fed to numerous rear loudspeakers. It was not until the release of Star Wars in 1977 that the big push came for cinemas to carry Dolby Surround.

In Dolby Surround the four signals are encoded (matrixed) to two and then decoded back to four. i.e.

\[
\begin{align*}
    L & \Rightarrow L_T & \Rightarrow & L_R \\
    R & \Rightarrow R_T & \Rightarrow & R_R \\
    C & \Rightarrow & \Rightarrow & C_R \\
    S & \Rightarrow & \Rightarrow & S_R
\end{align*}
\]
At the heart of Dolby Surround is a very simple mathematical concept:

The left and right total signals $L_T$ and $R_T$ are given by:

$$L_T = L + \frac{C}{\sqrt{2}} + j \frac{S}{\sqrt{2}}$$

$$R_T = R + \frac{C}{\sqrt{2}} - j \frac{S}{\sqrt{2}}$$

What do these equations mean? Essentially they tell us to take the Centre and surround channels, reduce them by 0.707 (or 3dB) and add them to the left and right channels. They also tell us to phase shift the surround channel by +90 degrees and add it the L channel and phase shift by -90 degrees and add it to the R channel (Giving a phase separation of 180 deg). This means that in order to extract four signals from the two signals the decoder must do the following:

$L_R = L_T$

$R_R = R_T$

$C_R = 0.707(L_T + R_T)$. Note that in adding the two signals, the surround channel cancels.

$S_R = 0.707(L_T - R_T)$. Note that in subtracting the two signals the centre channel cancels.

So upon decoding, we are left with four different audio streams:

- Stream A (left channel)
- Stream B (right channel)
- Stream C (the data that's identical in streams A and B)
- Stream D (the difference between the data in streams A and B)

Note that reduction of the surround and centre channels by 3dB is a common feature on surround encoders (More later!).

The problem with matrix encoding is that you can never reconstruct the original four channels perfectly: something is always lost through the process -- normally separation or isolation between the four channels! Or, put another way, matrixing introduces a lot of crosstalk. For example, we still have a certain portion of the left and right channels appearing in the surround channel. To combat this we can use a psychoacoustic trick known as the ‘law of the first wavefront,’ implemented by delaying the rear channel. This delay ensures that crosstalk from the front channels into the rear arrives at the listener long after (mSecs wise!) the front-channel sounds, and that it is therefore not perceived. Furthermore the rear channel is band-limited from 100Hz – 7kHz.
**Discrete Surround Sound** on the other hand is a system where we have $n$ discrete channels of audio usually for $n$ loudspeakers. An important consideration in discrete systems is the sheer amount of data necessary for all the channels. We may need to compress the audio in order to fit it onto a given storage medium. Or we may not need to compress at all if we have enough storage space and a player that can read the high bit-rates (e.g. Blu-ray or HD-DVD).

A good example of one of the most popular discrete surround systems is DTS Digital Surround (usually referred to simply as "DTS"). DTS is a scalable lossy encoding scheme for multichannel audio that operates on up to 24-bit words at sampling rates of up to 96kHz. When employed on DVD-Video, two bit rates are offered: 1.509Mbs and a half-rate of 754kpbs; the latter yields reduced audio quality but increased storage capacity. When employed on encoded CD, bit rate is 1.235Mbps, for a compression ratio of approximately 3:1 (for 16-bit data) or 4:1 (for 24-bit data). DTS is considered an "authorized optional" encoding scheme for both DVD-Video and DVD-Audio. It uses the *Coherent Acoustics* perceptual encoder.
Simple Preparation of your surround audio for home listening:

As a ‘quickstep’ to taking your 5.1 mixes/masters home for listening, I’m going to outline here how to create a DTS CD (yes a CD!!) for playback on most home DVD players:

This can be accomplished since standard compact discs (CDs) utilize a fixed bit rate of 1.411Mbps (as defined in the “Red Book” specification). We know that while this is sufficient to carry two channels of 16-bit PCM audio at 44.1kHz, it is insufficient to carry un-encoded surround sound audio. However, the application of DTS encoding allows the delivery of 5.1 channels of 24-bit, 44.1kHz audio at this same bit rate.

These "encoded CDs" are playable on standard DVD players; however, unless the player has a built-in DTS decoder (or unless it provides a S/PDIF digital output for connection to an external DTS decoder), only high-level noise will be heard. Up to 74 minutes of 5.1 audio can be stored on a DTS-encoded CD — exactly the same amount of stereo signal that can be stored on an unencoded audio CD.

To encode the audio for DTS we can use the SurCode DVD Professional program.

SurCode allows you to choose between two sample rates. The 44.1 sample rate is used for making audio DTS CDs, and the 48.0 sample rate is used for DVD-Video.

Encoding procedure

1) Tell SurCode where to find the 6 sound files of your Surround Sound master.

2) Tell SurCode where to place the encoded sound file.

3) Select sample rate option. (44.1K)

4) Hit the "Encode" button.
This will save the file as a DTS WAV file. This file can then be burnt to CD using any quality CD burning software. Once it is inserted into the DVD player (and assuming the DVD player has a DTS decoder), the DVD player should recognize the DTS stream and decode it to 5.1 automatically.

**Documentation**

Complete, clear, and accurate documentation should always accompany the master. This information is important not only when the master is in use but also as a reference, once it is archived. Dolby has created *Mix Data* and *Mastering Information Sheets* to facilitate proper documentation or to use as a guide for creating similar documents. The *Mix Data Sheet* provides concise information about the source media to all the engineers on a project. Typically, it will include information on sampling frequency, bit resolution, time code, track assignment, titles, and program start and stop times. The *Mastering Information Sheet* provides documentation relevant to the mastering engineer or authoring facility on source media, timing, and encoder settings, as well as general notes.

**- Mix Data Sheet**

The purpose of a *Mix Data Sheet*, is to provide all production engineers with thorough and concise media layout information. The information contained in the *Mix Data Sheet* should be distributed for technical parameters prior to encoding for final delivery, i.e., in production or postproduction PRIOR to the output distribution (DVD or Digital Broadcast). While size of the recording and production media may be on a smaller scale, accurate labeling of the media with *Mix Data Sheet* information will provide additional engineers with the proper insight of how the work is intended to be heard.

**- Mastering Information Sheet**

The purpose of a *Mastering Information Sheet*, is to provide the mastering engineer or the digital authoring specialist, technical information with respect to media layout, timing information, and encoder-specific information. This information is created during authoring/creation of the final delivery medium, e.g., Dolby Digital encoding of an AC-3 stream for a movie on DVD. Additional documentation such as production notes will be invaluable in completing a project. The purpose of *Notes* is to provide engineers with any explanation for key actions with relationship to time, level, error, or artistic consideration.
Mix Data

Date / /

Project

Client

Studio

Project #

Producer

Engineer

Sampling Frequency

☐ 44.1 kHz  ☐ 48 kHz  ☐ ×2  ☐ ×4

Bit Resolution

☐ 16-bit  ☐ 20-bit  ☐ 24-bit

Timecode Format

☐ 25 fps  ☐ 29.97 fps  ☐ 30 fps  ☐ DF  ☐ NDF

Media Format

☐ __________

Surround Level SPL Calibration

☐ Equal to Front  ☐ −3 dB to Front  _____ dB SPL/channel

Tones

☐ 1 kHz @_____dBFS  ☐ 100 Hz @_____dBFS

Downmix data

☐ See Notes  ☐ 100 Hz @_____dBFS

Channel Configuration

<table>
<thead>
<tr>
<th>CH1</th>
<th>CH2</th>
<th>CH3</th>
<th>CH4</th>
<th>CH5</th>
<th>CH6</th>
<th>CH7</th>
<th>CH8</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Timecode</th>
<th>Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>: : :</td>
<td></td>
</tr>
<tr>
<td>: : :</td>
<td></td>
</tr>
<tr>
<td>: : :</td>
<td></td>
</tr>
<tr>
<td>: : :</td>
<td></td>
</tr>
<tr>
<td>: : :</td>
<td></td>
</tr>
<tr>
<td>: : :</td>
<td></td>
</tr>
<tr>
<td>: : :</td>
<td></td>
</tr>
<tr>
<td>: : :</td>
<td></td>
</tr>
<tr>
<td>: : :</td>
<td></td>
</tr>
<tr>
<td>: : :</td>
<td></td>
</tr>
<tr>
<td>: : :</td>
<td></td>
</tr>
</tbody>
</table>

Notes:


# Mastering Information

**Date** / /  
**Project**  
**Client**  
**Studio**  
**Project #**  
**Producer**  
**Engineer**  

**Mastering Status**  
- [ ] Recommendation  
- [ ] Final Master  

**Sampling Frequency (fs)**  
- [ ] 44.1 kHz  
- [ ] 48 kHz  
- [ ] ×2  
- [ ] ×4  

**Bit Resolution**  
- [ ] 16-bit  
- [ ] 20-bit  
- [ ] 24-bit  

**Timecode Format**  
- [ ] 25 fps  
- [ ] 29.97 fps  
- [ ] 30 fps  
- [ ] DF  
- [ ] NDF  

**Media Format**  

**Tones**  
- [ ] 1 kHz @ ____ dBFS  
- [ ] 100 Hz @ ____ dBFS  

**Coding**  
- [ ] MLP Lossless™  
- [ ] Dolby® Digital  

**Timecode Control**  
- [ ] No  
- [ ] Yes  
- [ ] Notes: In/Out/Duration  

**Downmix data**  
- [ ] See Notes  

<table>
<thead>
<tr>
<th>Dolby Digital Encoding Information</th>
<th>Bitstream Information</th>
<th>Processing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Metadata Parameters</strong></td>
<td>Bitstream Mode: ________</td>
<td>[ ] Digital De-Emphasis</td>
</tr>
<tr>
<td>LFE Filter [ ] ON [ ] OFF</td>
<td>RF Overmod Protection: ________</td>
<td>[ ] DC Highpass Filter</td>
</tr>
<tr>
<td>Dialogue Level: ________</td>
<td>Copyright Bit</td>
<td>[ ] Bandwidth Lowpass</td>
</tr>
<tr>
<td>Mix Level: ________</td>
<td>Center Downmix Level: ________</td>
<td>[ ] LFE Lowpass Filter</td>
</tr>
<tr>
<td>Data Rate: ________</td>
<td>Surround Downmix Level: ________</td>
<td></td>
</tr>
<tr>
<td>Channel Mode: ________</td>
<td>[ ] Extended Bitstream</td>
<td></td>
</tr>
<tr>
<td>Encoder Used: ________</td>
<td>[ ] Dolby Surround EX</td>
<td></td>
</tr>
<tr>
<td>Software Version: ________</td>
<td>Preferred Downmix ________</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lf/Rt C Downmix Lvl: ________</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lf/Rt S Downmix Lvl: ________</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lo/Ro C Downmix Lvl: ________</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lo/Ro S Downmix Lvl: ________</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[ ] HDCD converter</td>
<td></td>
</tr>
</tbody>
</table>

**Surround Channel Processing**  
- [ ] 90-Degree Phase Shift  
- [ ] 3 dB Attenuation  

**Line Mode/RF Mode Preferences**  
- [ ] None  
- [ ] Speech  
- [ ] Film:  
  - [ ] Standard  
  - [ ] Light  
- [ ] Music:  
  - [ ] Standard  
  - [ ] Light
## Channel Configuration

<table>
<thead>
<tr>
<th>Mode</th>
<th>CH1</th>
<th>CH2</th>
<th>CH3</th>
<th>CH4</th>
<th>CH5</th>
<th>CH6</th>
<th>CH7</th>
<th>CH8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Timecode</th>
<th>Program</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>