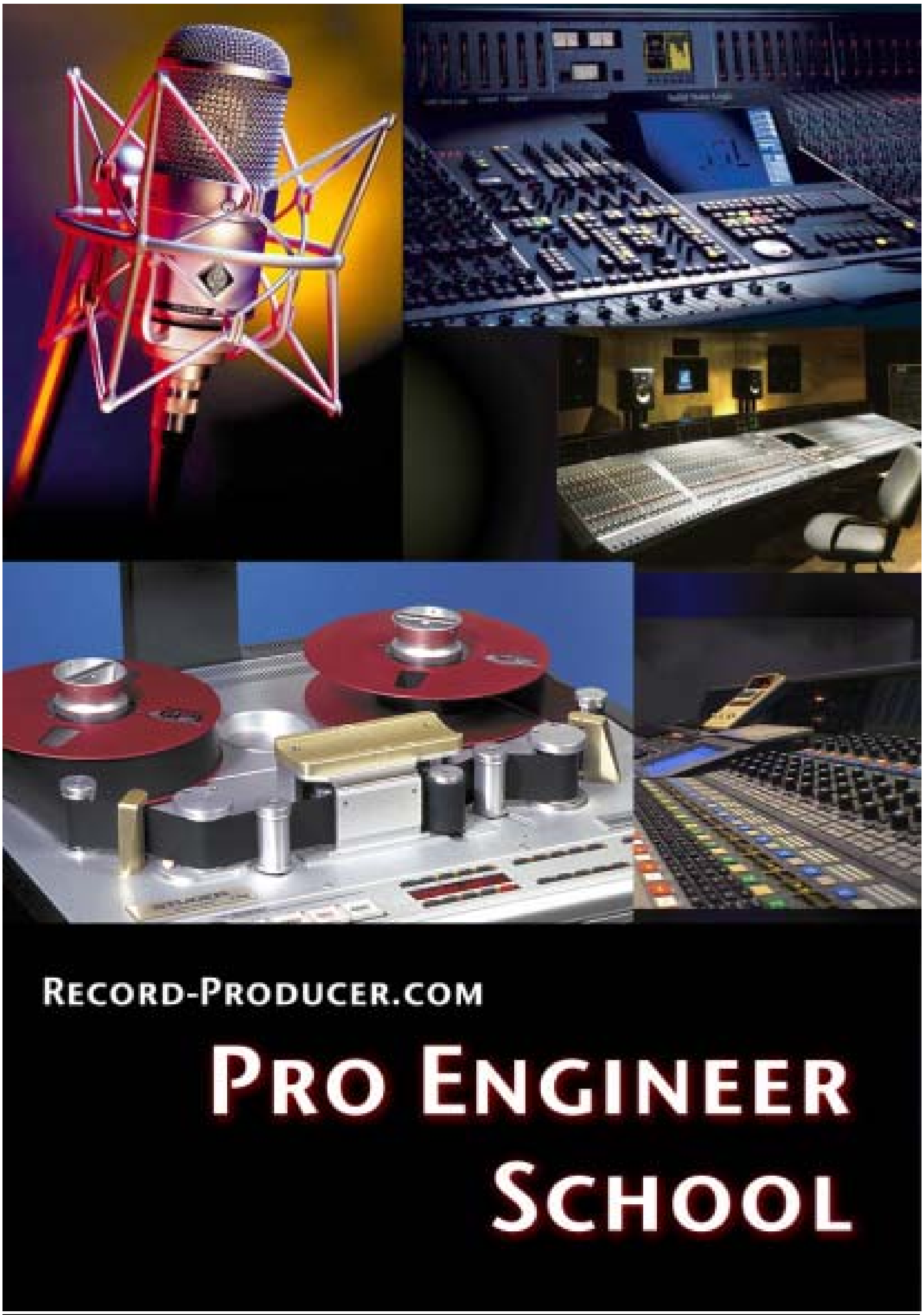


Volume 2

Contents

Copyright Notice

This work is copyright © Record-Producer.com

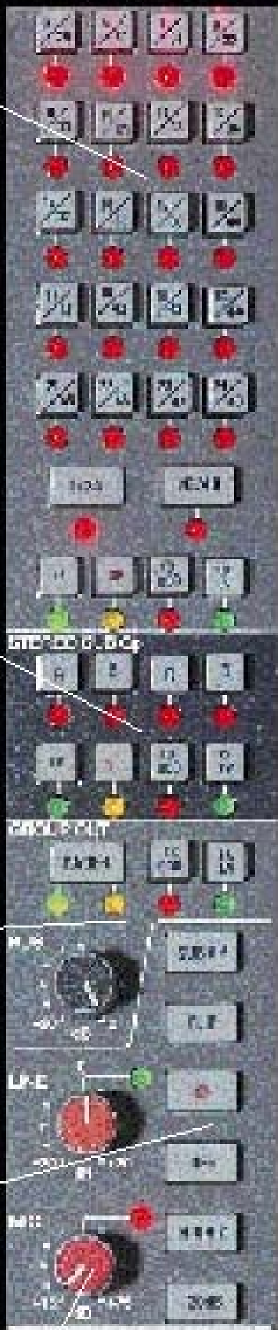
You are licensed to make as many copies as you reasonably require for your own personal use.

Chapter 1: Mixing Consoles (1)

The mixing console is the centerpiece of the recording studio, operationally and visually. The choice of mixing console defines a commercial studio - we talk of an 'AMS-Neve' studio (often simply 'Neve'), or an 'SSL studio'. There are other mixing consoles, but these are definitely the top two. Neve has a long tradition in recording dating back to the 1960s. Many Neve consoles manufactured from the early 1970s onward are still in use and are respected for their sound quality. SSL is a younger company, but they single-handedly defined the modern mixing console as the center of studio operations including control over tape machines, automation and recall. Whereas Neve have had a number of rethinks over the years on how a mixing console should work, SSL have been very consistent and there are many engineers who won't work on anything else, largely because they would have a tough learning period to go through.

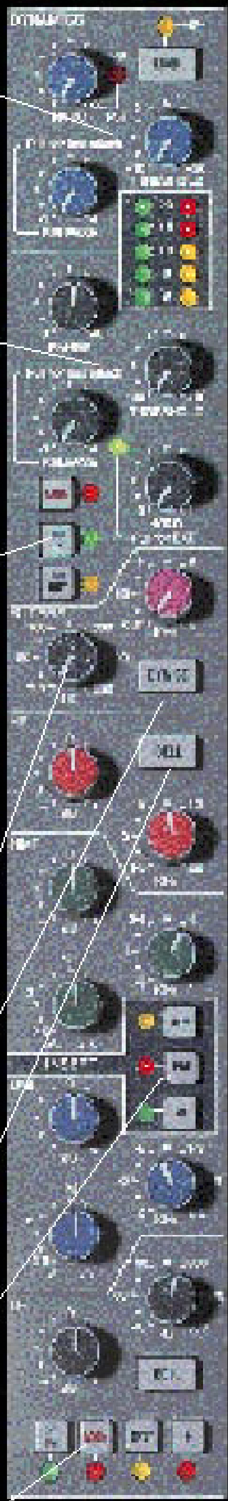
The first thing that a newcomer to recording has to realize is that we are not in home studio territory any more. These consoles are expensive - $300,000 or more. They are expensive because they are designed to do the job properly without compromise, allow efficient use of studio time, and attract business to the studio. As a learning music recording engineer, it should be ones ambition eventually to work in studios on Neve or SSL consoles. Anything else would be second best.

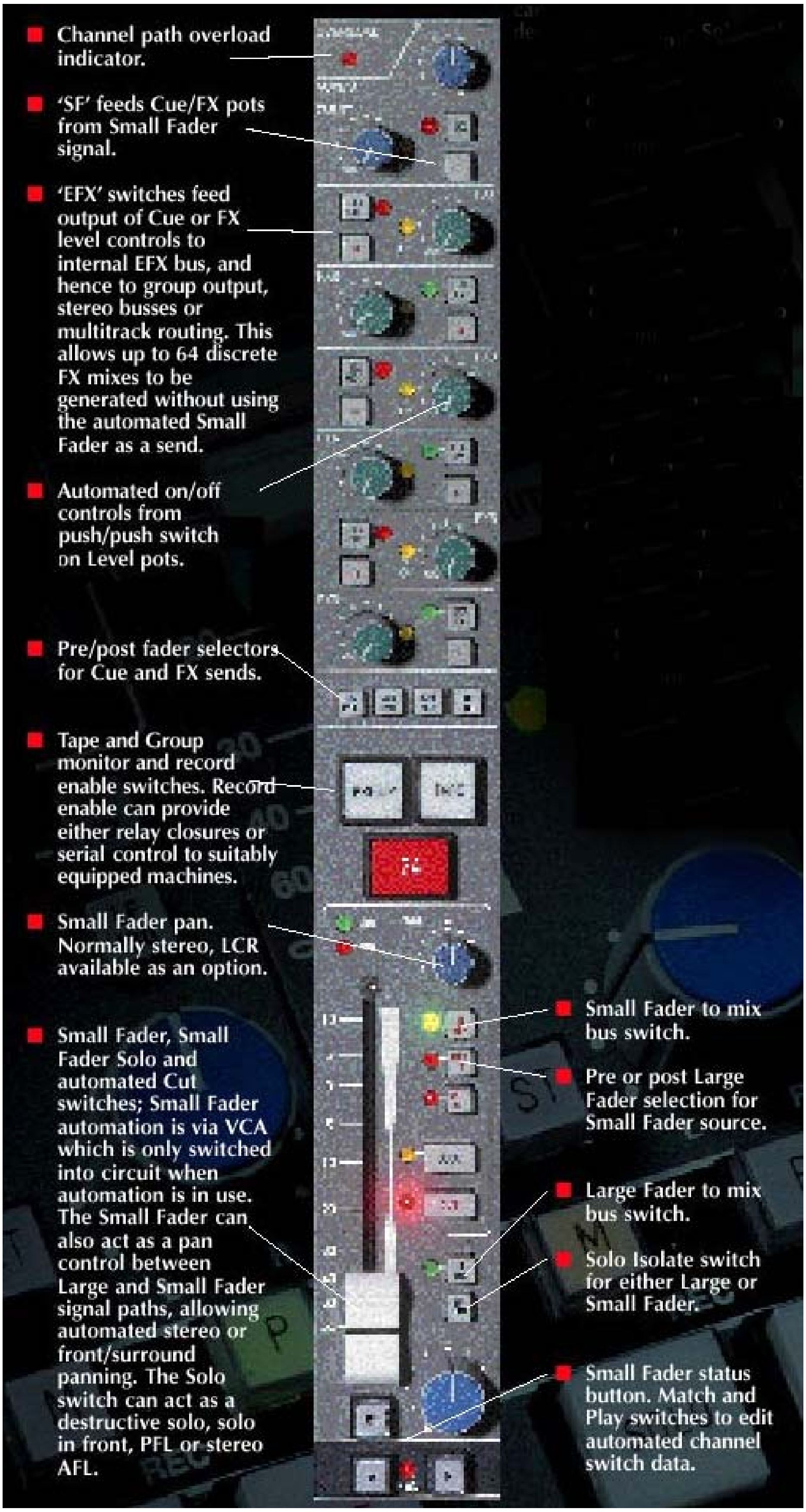
The next four pictures show the channel module of an SSL SL9000J console:



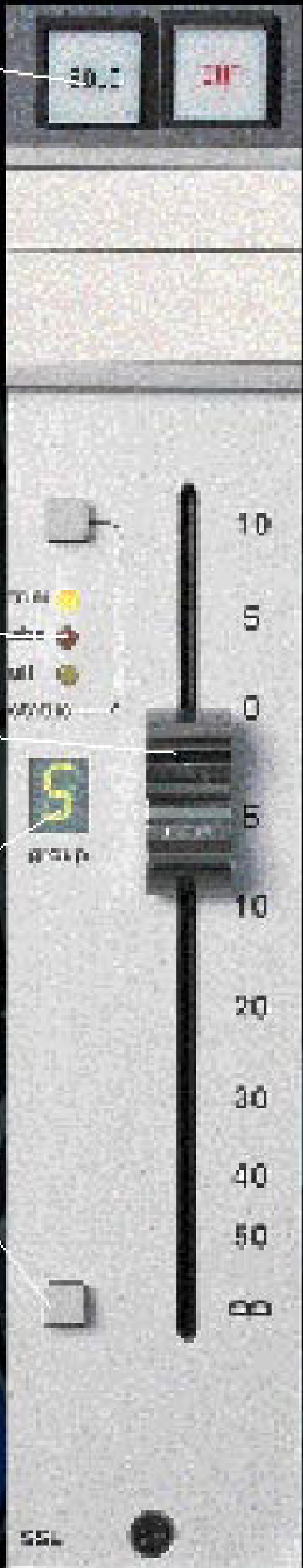
- 48 bus routing.
- 4 additional stereo busses with source select from Large Fader, Small Fader or EFX busses. The stereo busses can be used as stereo sub-groups or additional effects sends.
- Group output source –defaults to multitrack mix bus. 'Fader' selects post fade channel signal. Trim control for multitrack mix bus 30dB to 0dB.
- Input source select – Mic, Line or Sub Gp (multitrack mix amp). Automatically defaults to Mic in record and replay, and Line in mix status.
- Continuous mic gain control +15 to +75dB, plus 20dB pad and High Z switch changing impedance to suit line inputs.

- SSL's standard RMS compressor, switchable to peak sensing, hard knee characteristic via pull switch on Ratio control.
- Expander/Gate – Range, Threshold, Release and Hold controls. Expand selected via pull switch on Hold control.
- Dynamics switching – to monitor path, to channel path pre EQ, and to channel path post EQ. Selecting Mon and Ch In or Ch Out places the dynamics gain element in the channel path and feeds the side chain from the monitor path, allowing patch free keying of signals.
- Hi and lo pass filters – hi pass 18dB per octave, lo pass 12dB per octave.
- 'Dyn SC' – places filters in dynamics side chain.
- Bell switch – active in all EQ modes.
- Insert can be switched pre or post EQ. Insert return can be used as Key input to dynamics side chain. Automated Insert In switch.
- Automated EQ In switch. Filter Split and EQ to Monitor switches. EQ Type switch changes EQ curves from 'G' to 'E' type.




Channel path overload indicator.
'SF' feeds Cue/FX pots from Small Fader signal.
'EFX' switches feed output of Cue or FX level controls to internal EFX bus, and hence to group output, stereo busses or multitrack routing. This allows up to 64 discrete FX mixes to be generated without using the automated Small Fader as a send.
Automated on/off controls from push/push switch on Level pots.
Pre/post fader selectors for Cue and FX sends.
Tape and Group monitor and record enable switches. Record enable can provide either relay closures or serial control to suitably equipped machines.
Small Fader pan. Normally stereo, LCR available as an option.
Small Fader, Small Fader Solo and automated Cut switches; Small Fader automation is via VCA which is only switched into circuit when automation is in use. The Small Fader can also act as a pan control between Large and Small Fader signal paths, allowing automated stereo or front/surround panning. The Solo switch can act as a destructive solo, solo in front, PFL or stereo AFL.
Small Fader to mix bus switch.
Pre or post Large Fader selection for Small Fader source.
Large Fader to mix bus switch.
Solo Isolate switch for either Large or Small Fader.
Small Fader status button. Match and Play switches to edit automated channel switch data.

- Large Fader Solo and automated Cut switches. Solo switch can act as a destructive solo, solo in front, PFL or stereo AFL.
- Status switch and status LEDs for Large Fader.
- Motorised fader. Audio switchable between fader track and VCA.
- 7 segment display shows current group assignment.
- Group select switch-cycles through all 8 group positions. Group assignments stored and reset as part of project data.



Firstly, let's consider the functions of a multitrack music recording console:

- Record from many microphones and line input sources simultaneously.
- Record to multitrack, or mix live sound sources into stereo.
- Allow previously recorded tracks to be monitored while overdubs are made.
- Mix a multitrack recording into stereo.

With these points in mind, let's run through the console. Each channel module contains the following:

- Microphone input
- Line level input
- Multitrack monitor input
- Insert point
- Equalizer
- Auxiliary sends
- Compressor
- Noise gate
- Small fader & pan
- Large fader & pan
- Solo/PFL
- Automation controls
- Microphone and Line Input

A microphone delivers a signal of 1 mV up to around 1 V (in extreme conditions). The microphone preamplifier is specially designed to amplify this as necessary up to around 1 V on peaks and deliver a low noise signal to the following circuitry. The amount of gain is set using the Gain control. The microphone input will supply 48 V phantom power to the microphone.

The line input accepts signals of around 100 mV to 1 V and provides amplification as necessary. On some consoles, there is no gain control on the line input.

The phase button inverts the signal.

The filter cuts low frequencies below around 100 Hz.

The pad button attenuates the signal, usually by 20 dB.

Insert Point

The insert point allows access to the signal in the channel so that it can be routed through a compressor, noise gate or equalizer. Each channel has its own independent insert point. The insert send sends signal to the external device. The insert return accepts the output from the external device. The whole of the signal is processed and is not mixed with the input signal. The insert point may be positioned before the EQ or after the EQ.

The outputs of the mixing console have insert points too. For example the master stereo output has insert points into which a compressor or EQ could be inserted to process the whole mix, and the level of the mix could still be controlled by the master faders.

Equalizer

The equalizer has the following sections:

- High pass filter (cuts low frequencies)
- Low pass filter (cuts high frequencies)
- High frequency EQ with controls for frequency and bell/shelf
- Two mid frequency EQ sections with controls for frequency, gain and Q
- Low frequency EQ with controls for frequency and bell/shelf
- EQ in/out switch

Auxiliary Sends

Auxiliary sends are used for two main purposes:

- Sending foldback signal to musicians.
- Sending signal to effects units such as digital reverb.

Auxiliary sends can be either 'pre-fade' or 'post-fade'. Pre-fade means that the signal is taken from a point before the fader. This is suitable for foldback since a foldback mix can be constructed that is completely independent of the positions of the faders. Post-fade is suitable for reverb since when you fade the signal out, you generally want the reverb to fade out as well, and be in proportion at all other fader positions.

Dynamics

The compressor and noise gate are similar to outboard compressors and noise gates, which will be covered in a later chapter.

Small Fader

Since this is an inline console, each channel module has two signal paths:

- The input signal, which is the signal from the mic that is being recorded to multitrack
- The monitor signal, which is the output of a single track of the multitrack recorder, the track number of which corresponds to the channel number (usually).

On some consoles, the small fader is normally set to control the level of the monitor signal. So the large faders are used to set recording levels to multitrack, and a temporary monitor mix is set up on the small faders. On other consoles, this - as a normal condition - is reversed. All inline consoles allow the input and monitor signal paths to be 'flipped', i.e. reversed.

At this point, it is worth saying that the other facilities of the channel can be allocated to either the input or monitor signal paths, or shared. So for example, you could place the EQ in the input signal path if you wanted to EQ the signal before it went down to tape. On the other hand if you wanted to record the signal flat, you could put the EQ in the monitor path and use it to temporarily sweeten the monitor mix. This applies to the dynamics section and auxiliary sends too.

Solo/PFL

Mixing consoles always have the ability to 'solo' any channel so that you hear than channel by itself through the monitors, generally without affecting the signal being recorded (or amplified or broadcast). This can operate in a variety of ways. A large-scale console will give you options, on a smaller console it will probably be fixed.

PFL: Pre Fade Listen means that you hear the signal of the selected channel alone, picked off from a point before the fader, so that the fader level has no influence on the PFL level. On some consoles, the PFL signal is routed equally to left and right speakers, on others the position of the pan control is retained. PFL is often used for setting the gain control: on each channel, one at a time, press the PFL button and set the gain so that the main meter reads a good strong level without going into the red.

Solo or AFL: Solo, or After Fade Listen, is as above, but the position of the fader is taken into account.

Solo-In-Place ("SIP", "Cut Solo" or "Kill Solo") mutes all other channels. Exceptionally, this does affect the other outputs of the console, which is why you will never find it on a live broadcast console.

Solo Safe: Often it is useful to protect some channels so that even when other channels are soloed, they still operate. This is of greatest benefit with Solo-In-Place, where the channels used as effects returns can be placed in Solo Safe mode. Now when you solo a channel, you will hear it in isolation, at its correct level, correct pan position, and with its reverb intact.

Large Fader

The large fader is motorized and incorporates a VCA (Voltage Controlled Amplifier). It has the following facilities:

Can be grouped to a 'master' fader so that the master controls all the movements of the slaves.

Can be automated (other automatable functions include mutes and some other switched functions).

On the SL 9000 and some other SSL consoles, the 'Ultimation' automation system allows audio to pass through the fader track (so-called 'moving fader' automation), or it can be routed through the VCA ('VCA automation').

VCA Grouping

Many large consoles offer a function known as VCA grouping.

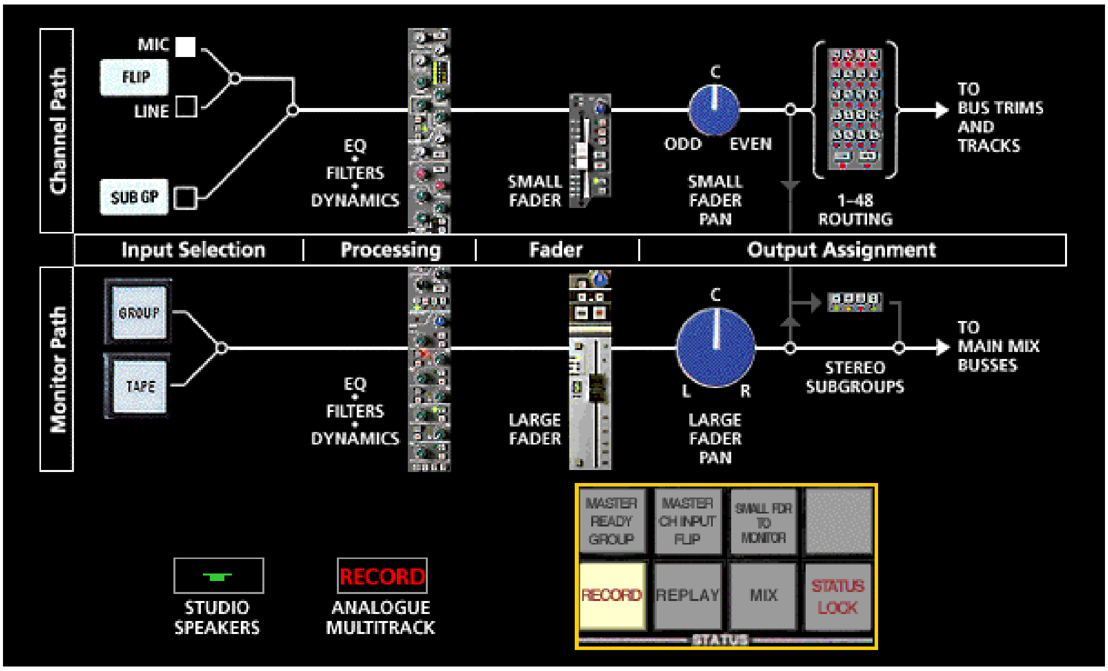
Subgrouping is a technique by which several channels can be routed through groups, which are then in turn routed to the main outputs. For examples, the eight or so channels of a drum kit are balanced to give a good mix, then they are routed to two group faders (used as subgroups) for convenient overall level control of the kit, without disturbing the internal balance.

On a console with VCA grouping, all the channel faders have a small switch that can select one of (usually) eight VCA masters. When a VCA master is selected, that controls the overall level of all channels set to that group. So, the level of the entire kit is controlled by one fader.

With conventional subgrouping, the audio is mixed through the group fader(s). With VCA grouping, the VCA fader simply sends out a control voltage to the channel faders.

Signal Paths

The path the signal will take, whether through small fader or large fader, depends on the point you are at in the recording procedure - record, replay, overdub or mix. A large-scale console will offer master switching so that all channels can be set instantly to the appropriate mode, and individual channels can be flipped as necessary. See the following SSL SL9000J signal flow diagrams.

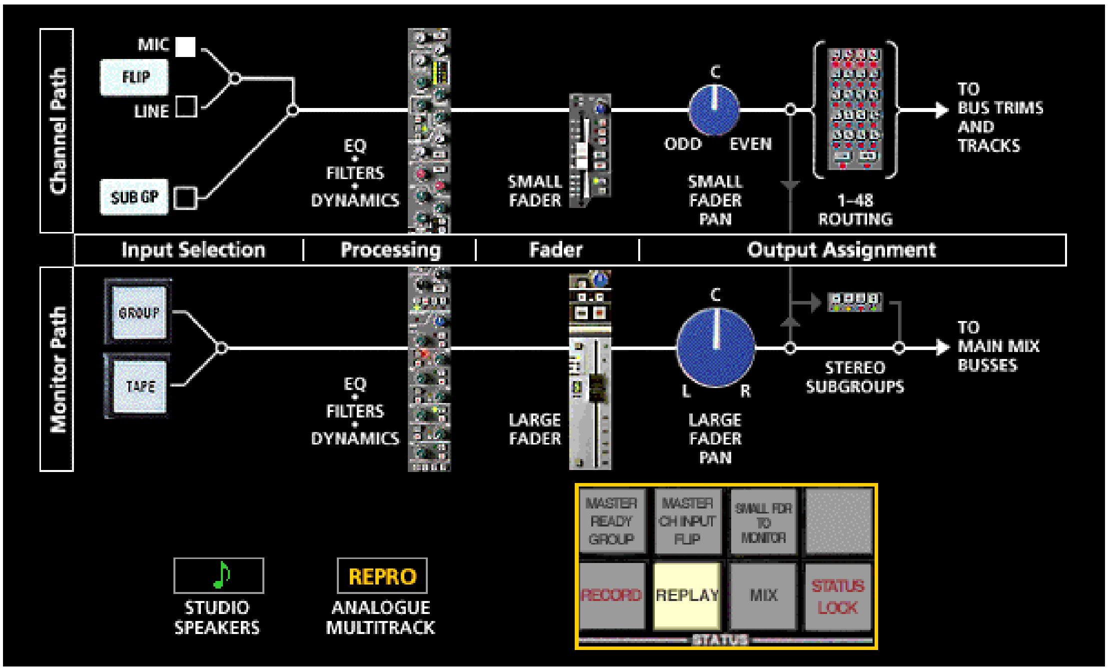


Record Status

Recording basic tracks onto a blank multitrack tape is the starting point! In the record mode, with the RECORD status button selected, the various elements in the module signal paths are connected as shown. This will be the preferred recording mode for most engineers.

The Large Faders are much more useful if used as monitor faders during recording, as they can, if required, be automated for end-of-the-day monitor mixes.

The upper section of the diagram shows the ‘Channel’ signal path whilst the lower part shows the ‘Monitor’ signal path. The Channel signal path is that path which originates from the Channel Input section of the I/O module. The Monitor signal is derived from the Monitor Input section.



Replay Status

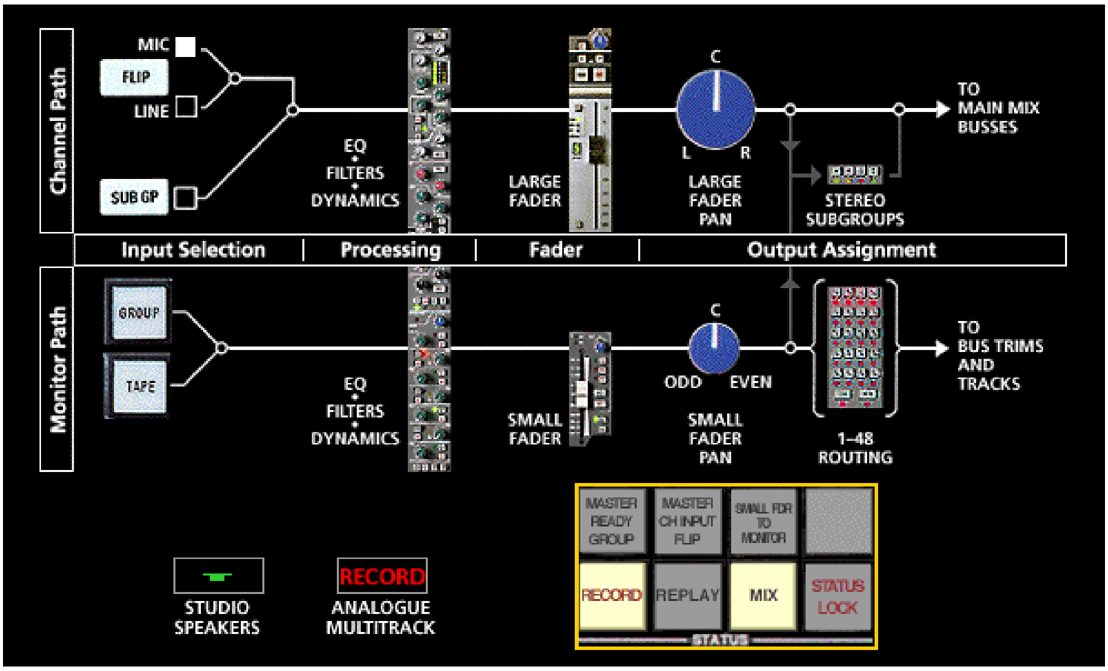
This mode is used when working in RECORD, or RECORD + SMALL FADER TO MON status. The current console status is put on ‘standby’ and the tape returns are automatically routed to the Monitor faders. This allows a quick replay of the tape without disrupting the console setup. If the Sync/Replay option is wired, then an analogue multitrack machine will be switched to normal Replay.

This status is useful during track laying. For example, when operating in RECORD status, the time will come when a quick monitor mix is required. This can be accomplished in RECORD status by deselecting any GROUP buttons, switching the multitrack machine to Replay manually, and mixing down the monitor inputs via the main output busses onto a stereo ATR. REPLAY status does all this with one button.

Any GROUP selections are temporarily disabled and the monitor inputs pick up multitrack returns from the Replay head.

Reselecting RECORD status will reinstate all the previous GROUP and TAPE button selections, and an analogue multitrack will switch back to Sync, ready for more recording. REPLAY status is also useful for playback over the Studio Loudspeakers, as RECORD status prevents the

SLS outputs from receiving signal.



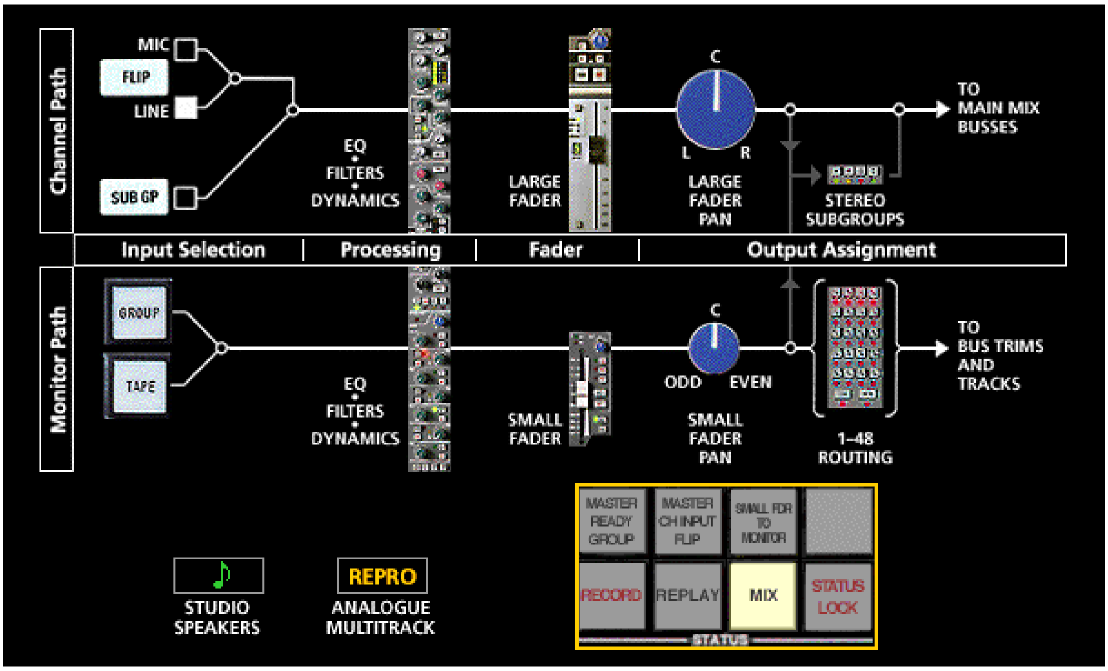
Record + Mix Status (Overdub Mode)

This mode was designed for use in overdubbing but many engineers will use this status when laying basic tracks. Select this combined status by pressing the RECORD and MIX status buttons simultaneously.

The desk is basically in MIX status but an individual module may be put into the RECORD status, in order to record onto that track, if either the TAPE or the GROUP button is selected. The advantage of this mode is that the majority of modules will be in MIX status and you can mix with the Large Faders as if you were doing a final mix.

In other words, the modules are not split into source signal paths and monitor signal paths unless you are recording from that module. You can work towards the final mix as you are tracking, using the mix capabilities to their full extent but with the ability to record onto the necessary tracks.

It is quite usual for the desk to be split for this way of working. The first 24 or 48 modules are dedicated to the multitrack, and modules upwards from 25 or 49 act as source channels, although this is not essential.



Mix Status

Line inputs are selected on the channels, sent via the Large Faders and Large Fader Pans to the main Mix bus and then, via the Master Fader, out to the mastering machine. The multitrack machine is usually normalled to the Line inputs, so this single status button will instantly set you up for a mixdown.

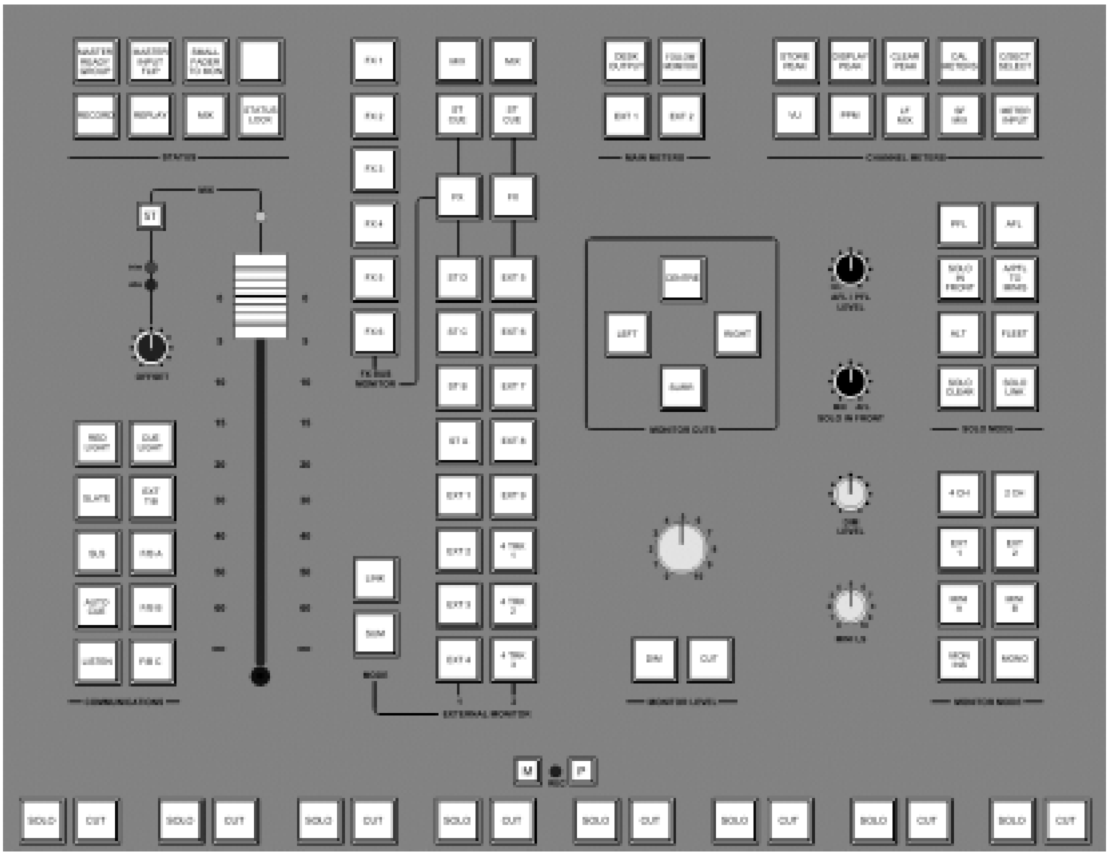
The Small Faders can be used for a variety of different purposes in MIX status. You will see from the drawings above and opposite that the default Monitor path in MIX status feeds the Tape Monitor Inputs via the Small Fader and Pan to the Multitrack Routing Matrix.

Check Questions

- What is the function of the gain control?
- What other switched functions are available close to the gain control?
- What is an insert point?
- What type of processing would an insert point be used for?
- What is meant by 'pre-fade auxiliary'?
- What is meant by 'post-fade auxiliary'?
- Describe the two signal paths in the channel module of an inline console?
- If the small fader is used to control the signal level sent to the multitrack recorder, what is the large fader used for?
- What is the meaning of 'flip'?
- Comment on the sharing of facilities between input and monitor signal paths.
- Describe two uses of PFL.
- Comment on the danger of solo-in-place.
- What is 'solo safe'?

Chapter 2: Mixing Consoles (2)

As well as channel modules, all mixing consoles have a center section with controls that affect all channels, the stereo output, and stereo monitoring. The center section of the SSL SL9000J is shown here.



Status Buttons

The status buttons are the important master controls that set signal routing for all channel modules.

Record: used when recording basic tracks

- Record from microphone inputs to the multitrack recorder.
- The multitrack recorder is switched to Sync output so that the output signals are synchronized with the signal being recorded. (In an analog recorder, the output is taken from the record head, not the playback head).

- The channel inputs are routed to the small faders via the multitrack routing matrix to the multitrack.

- The large fader carries the group output or the multitrack return signal as selected locally on the channel module. (A 'group' is a mix of signals, or single signal, sent to one track of the multitrack. There are as many groups as channel modules. The first 48 are accessed via the routing matrix on each channel).

Replay: used for playback of basic tracks

- As Record status except that the multitrack is switched to the normal playback output rather than sync, which offers slightly improved quality in an analog recorder.

- The studio loudspeakers are switched on so that the musicians may hear the playback.

- All large faders are switched to Tape, overriding any local Group selections.

Mix: used for mixing to stereo

- When recording is complete, the console is switched to Mix status.

- All the channel modules are switched to line input, to which the outputs of the multitrack recorder are normally connected.

- The signals are routed via the large faders to the stereo mix.

- The small faders are now fed from the Group/Tape selection buttons, and send signal via the multitrack routing matrix. The small faders can be used to feed additional signals to the mix, or as additional auxiliary sends.

Record + Mix (Overdub): used when adding overdubs to basic tracks

- When the Record and Mix status buttons are pressed simultaneously, the multitrack switches to sync output, if available.

- All channel modules go into Mix status, unless the module's Group or Tape switch is pressed, in which case that channel goes into Record status.

Additional Buttons

Master Input Flip

- Flips all channel inputs between mic and line.

Small Fader to Monitor

- In Record and Replay status, swaps the operation of the small and large faders so that the small faders feed the monitor mix and the large faders feed the routing matrix.

Master Ready Group

- Sets all channel modules to Group so that the console is instantly set up to monitor group outputs.

Status Lock

- When the console is used for live mixing or for broadcast, the operation of certain functions would affect audio heard by the public. The following functions are disabled:
    - Status button changes - Record. Mix, Replay, Small Fader to Monitor
    - Master Input Flip
    - Oscillator On
    - Slate Talkback
    - Listen Mic to Tape
    - Autocue
    - Solo-In-Place

- Studio Loudspeaker Output
- Also, AFL (After Fade Listen) is selected as the solo mode of operation and the studio's red light is switched on, if available.

Main Outputs

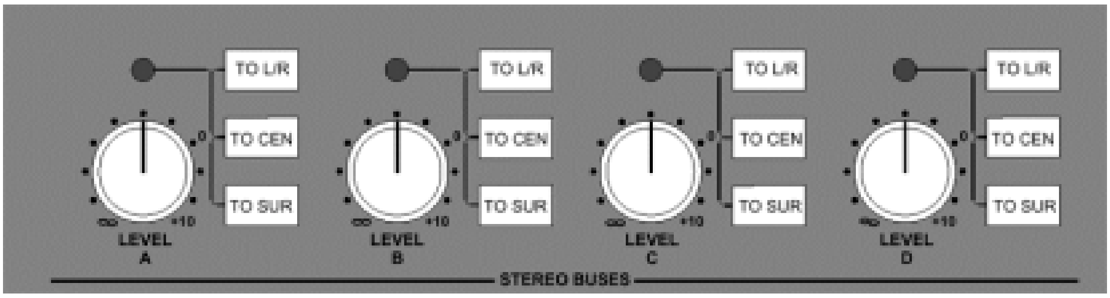
Master Fader

- The Master Fader controls the output level of the main stereo mix (in fact the surround sound mix too, if this is required).

Offset Control

- Raises or lowers the main output level by +/- 20 dB. This allows the master fader to be left at maximum, which is usually a convenient setting, and also has the advantage that the output VCAs are bypassed for optimum sound quality.

Stereo Bus Master Controls



- The SL9000J has four stereo buses, which can be routed to the left/right stereo output or to the center or surround channels of a surround mix. Each stereo bus has an insert point on the patchbay so that a stereo signal can be processed by an external compressor or EQ before continuing to the main outputs.

Control Room Monitoring System

Monitor Selection and Control

The large monitor level rotary control sets the level of the main control room monitor loudspeakers. The output to the control room loudspeakers is normally derived from the console's stereo output, but may be switched to external inputs, EXxt1 or Ext 2. Ext 1 and 2 can each take their signal from any one of eleven sources. The large monitor level rotary control sets the level of the main control room monitor loudspeakers. The output to the control room loudspeakers is normally derived from the console's stereo output, but may be switched to external inputs, EXT 1 or Ext 2. Ext 1 and 2 can each take their signal from any one of 11 sources.

The Dim button reduces the monitor level to a level set by the Dim Level control. The Dim function is also automatically activated when any of these buttons are pressed: Listen Mic; Foldback A, B or C; SLS (Studio Loudspeakers); Oscillator to Mix; ABCD or Busses 1-48.

The Cut button cuts the control room monitors and is also automatically activated by the Slate button.

'2 Ch' selects normal two-channel stereo monitoring.

'4 Ch' selects surround sound monitoring.

Mono sums left and right channels to the monitor, allowing the degree of mono compatibility of the mix to be assessed.

'Mini A' and 'Mini B' allow either of two sets of nearfield monitors to be used instead of the main monitors. The mini loudspeaker output has its own level control relative to the main monitor level.

Solo

The default solo mode is Solo-in-Place, which destructively cuts all other channels to all outputs. There are two separate solo buses for the small and large faders, which can be linked using Solo Link.

- AFL: solos the post-pan, post-fader signal.
- PFL: solos the pre-pan, pre-fader signal.
- A/PFL to Minis: routes the solo signal to the nearfield monitors without affecting the signal to the main monitors.

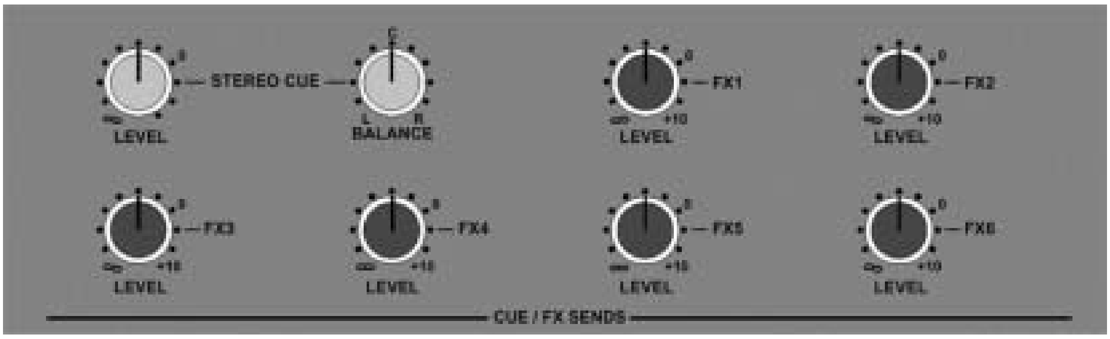
- Solo-in-Front: dims the main output in relation to the soloed channel, rather than muting it completely.

- Alt: With this selected, each new solo button pressed cancels currently selected solos.

- Fleet: makes all solo buttons momentary action rather than latching.

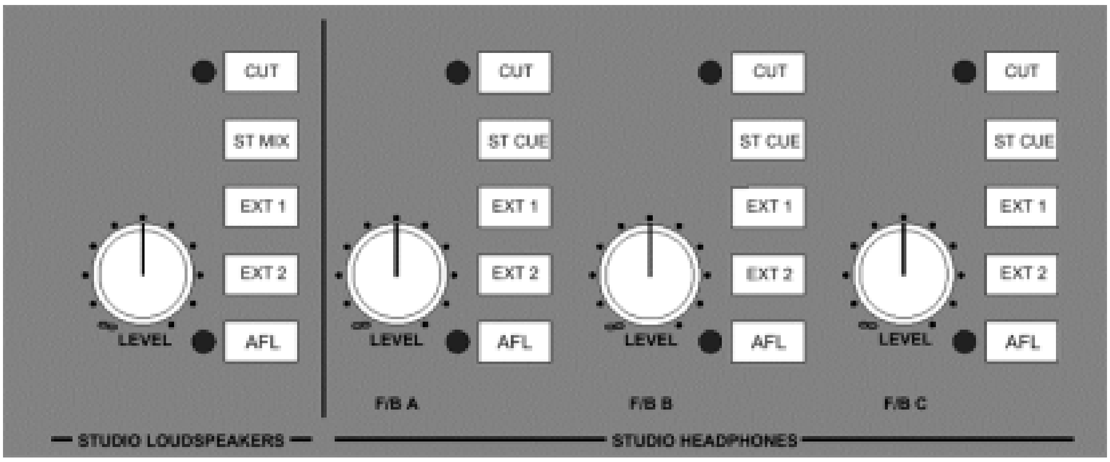
- Solo Clear: clears all solos other than center section AFLs, which latch mechanically.

Cue-FX Sends



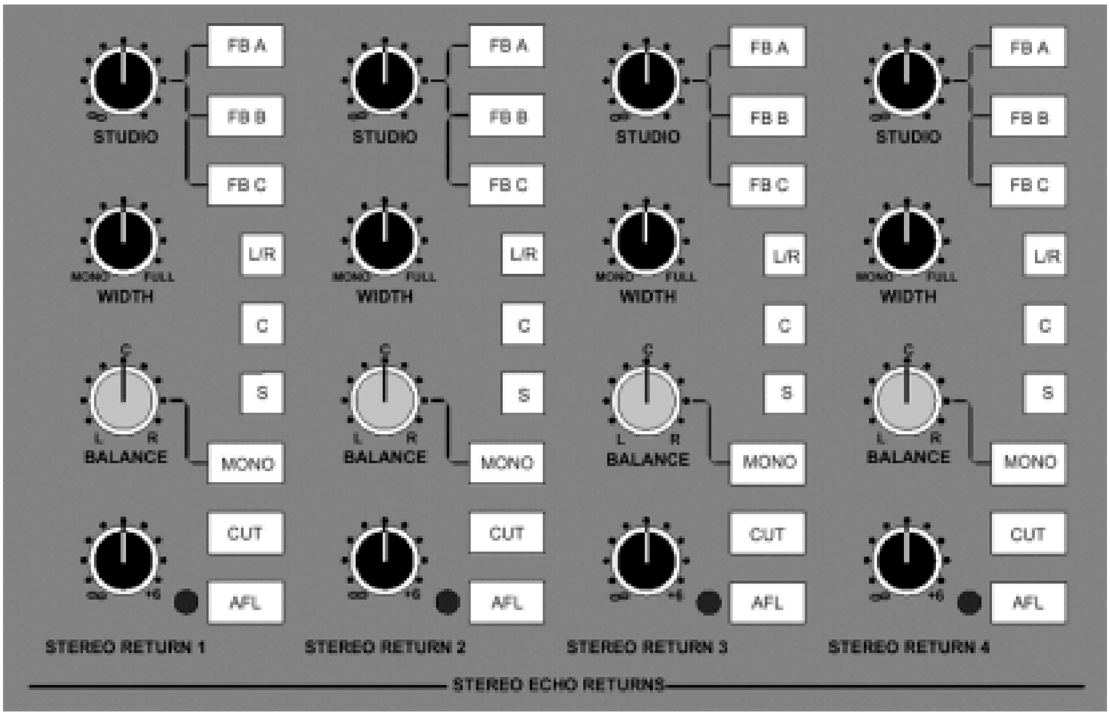
These are the master controls for the auxiliary sends in each channel module.

Studio Loudspeakers and Foldback Sends



The studio loudspeakers are for musicians to hear work in progress in the studio, rather than have them come into the control room. The console provides three channels of foldback.

Stereo Echo Returns



These are normally used for the return signals from reverberation units. Reverb can be mixed into the main outputs or into the foldback signal.

Meters

The main meters can be switched to read:

- The main console outputs, or;
- Whatever signal is being fed to the main monitors, or;
- Ext 1 or Ext 2 external sources.

The individual channel module meters have a number of switched options:

- Store Peak: retains the highest level reading reached in the channel, until cleared by the Clear Peak button.

- Display peak: shows the highest level reached as a single segment at the top of the meter column.

- Cal Meters: allows technical staff to calibrate the meters.

- C/Sect Select: selects the source for additional meters located above the computer monitor.

- VU: gives the meters the ballistic characteristics of an analog VU meter. The traditional VU meter does not show peaks accurately but corresponds quite well to the subjective loudness of a signal.

- PPM: gives the meters peak program meter characteristics. Peaks are measured accurately.

- LF Mix: displays the VCA control voltage of the large fader. This is useful to see changes in fader levels during replay of an automated mix with the fader motors switched off.

- SF Mix: as above, but for the small fader.

- Meter Input: with this selected, all meters display channel input signal levels.

Group Faders

Below the center section there are eight VCA group faders. These can control the levels of any channel assigned to them using the switch in each channel fader module. The Cut control mutes all channels routed to that group. The Solo control cuts all groups not soloed.

Communications

There is a built-in talkback microphone with functions as follows:

- TB to Foldback sends talkback signal to the foldback outputs.

- TB to SLS sends talkback to the studio loudspeakers.

- Slate cuts all loudspeakers and sends talkback to multitrack outputs 1-48, main outputs, foldback sends and studio loudspeakers. A 30 Hz tone is mixed in to allow rapid finding of slate announcements recorded on analog tape machines.

In addition:

- Two 'listen mics' can be installed in the studio to monitor audio activity other than via the recording microphones.

- Red Light allows a studio red light to be switched on automatically when the multitrack is put in record mode.

- There is an additional cue light with button activation.

- Ext T/B provides an additional talkback destination, plugable from the patchbay.

- Auto Cue allows the Talkback and Listen Mic switches to be latched on for continuous communication between the control and studio, but only when the multitrack is in Stop or Wind mode.

- The Oscillator sends tone to the multitrack and main outputs. The main use of this is to record tones of a range of frequencies at the same level to an analog multitrack. When the tape is played back, the playback machine can be aligned for a flat frequency response. Digital recordings also use tone, but just to set a reference level. SSL recommend switching the oscillator off when not in use to prevent tone leaking into other signal paths (!).

In conclusion, large-scale mixing consoles are complex because they incorporate a function to correspond to every need the recording engineer might have. There are no superfluous functions, although any one engineer might only use (and understand) perhaps 80% of the console's capabilities.

Check Questions

- Describe the use of Record status.
- Describe the use of Replay status.
- Describe the use of Mix status.
- Describe the use of Record + Mix (Overdub) status.
- Why is 'status lock' important for some types of use of the console?
- Comment on the options provided by the Monitor Selection and Control section.
- What is 'Solo-in-Front'?
- Describe the relationship between the FX Send 1 control in the channel module and the FX 1 Level control in the Cue/FX sends section.
- Describe the main metering options in the SL9000J console.
- Describe the main communication features options in thee SL9000J console.

Chapter 3: Equalization

EQ is a very powerful and effective audio tool, but it is always best to ensure that you get as good a sound as possible from the microphone, synth or sampler coming into the mixing console. If you start off with good sounds, then a good result is almost inevitable. It is becoming increasingly popular to use microphones for recording, even when DI (direct injection) is possible, because of the wider variation of tonal qualities available. Even small variations in microphone position make vast differences to the sound picked up. It is a sign of an expert recording engineer that he or she will listen carefully to the sound from the mic and adjust its position and angle, and even try out several microphones, rather than pretend that it is always possible to get it right first time.

There are certain 'rules of EQ':

- You should always aim to use EQ to improve an already wonderful sound.
- If the sound isn't intrinsically good without EQ, then you will never end up with anything but second best.
- The only time you should ever use EQ to 'save' a sound is when you have been given a tape to work on that was recorded by a lazy engineer.

Just as there is an art to creating a brilliant sound, there is an art to bringing that sound to perfection, and also blending several sounds together to make the perfect mix. Van Gogh didn't learn to paint overnight, and no-one is born with the inbuilt ability to EQ. It's a skill that is learned by experience and a good deal of careful listening.

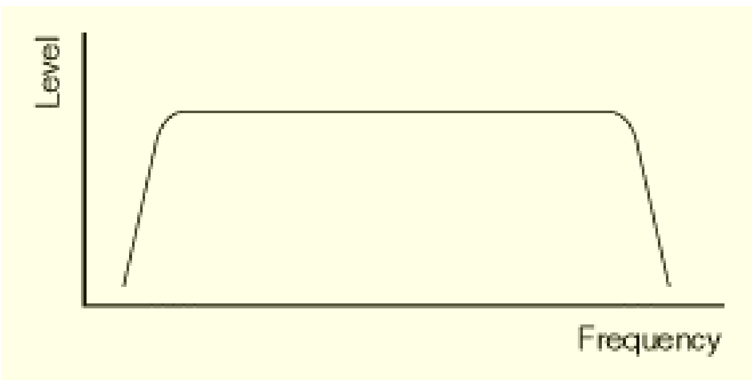


Fig. 1

Figure 1 shows one of the parameters you would expect any item of sound equipment to aspire to - a flat frequency response. This, or at least a very close approximation, will be the frequency response of your mixing console with the EQ controls set to their centre positions, or with the EQ buttons switched off. Here, the balance of frequencies of the original signal is preserved in correct proportion at the output. In other words it is just as trebly, tinny, harsh, nasal, honky, bassy or boomy as it was when it left the microphone; or just as perfect perhaps.

Notice that the frequency response indicates what the EQ does to the sound. A cymbal will naturally have strong high frequencies, for example, and that emphasis towards HF will be preserved by a flat EQ setting. Likewise, a flat EQ will reproduce perfectly the boomy bottom end of an undamped bass drum.

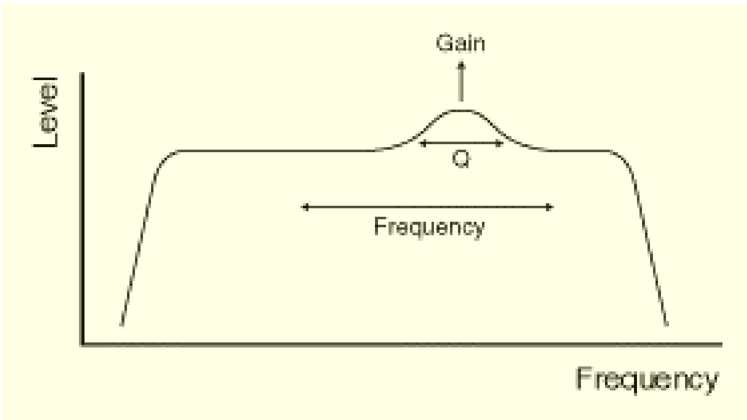


Fig.2

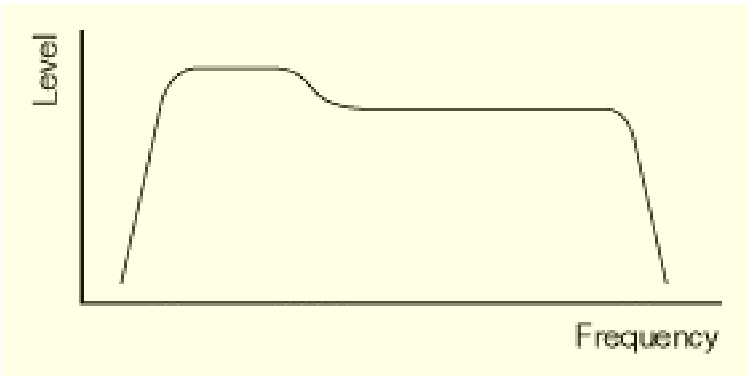


Fig.3

If Figure 1 shows a flat response between about 20Hz and 20kHz, Figure 2 and Figure 3 show two of the curves you might expect to get from a mixing console EQ. Oddly enough, measuring the EQ and plotting the curve is something that only 0.001% of recording and sound engineers ever get around to doing at any stage in their creative careers, and only 0.0001% have their own equipment to do it to any reasonable accuracy. Even if it's hardly ever done, except on the test bench, it's a useful concept that you can carry around in your head without ever bringing to the forefront of your mind. So if a producer ever says to you, "Let's have a little more presence in the vocal", your subconscious mind will retrieve the bell-shaped curve of Figure 2 from your memory while your conscious mind adjusts the controls and judges the sound.

In Figure 2 we are adding an EQ boost, and there are three parameters that we would like to be able to control (if the EQ has knobs for all three). First and foremost is the frequency: this boost could be centred on any frequency according to the instrument and according to which characteristics you want to accentuate. Second is the gain, which is the degree of boost and can be measured in decibels (dB) at the centre frequency. Some mixing consoles have this control calibrated in dB, up to 12 or 15 dB at maximum. Gain can also be negative, producing an EQ cut, which would be written as a gain of -6 dB (or whatever) at the center frequency, so the curve would dip downwards. EQ cut, by the way, is a vastly under used resource on many consoles, but more on this later.

The third parameter is Q. Q is a measure of the 'bandwidth' of the bell-shaped curve. A low Q - 0.3 is low - will allow the EQ to cover a wide range of frequencies, while a higher Q - 5 is high - will allow you to home in on a particular feature of the sound.

The bell-shaped curve of Figure 2 is often referred to as 'peaking' EQ, and applies to all mid frequency range EQ sections and a good proportion of high and low frequency EQ sections too. Figure 3 shows a 'shelving' EQ, where the boost (or cut) extends from the chosen EQ frequency all the way to the extreme end of the range. I have shown a low frequency shelving EQ in boost mode, but it could have been a high frequency cut with a similarly shaped but differently orientated curve. It isn't possible to say which type of curve is better, for it depends on what you want to achieve, but some consoles have a button to allow you to choose.

Filters

A filter is the simplest form of EQ section possessing only the parameters of cut-off frequency and slope.

The cut-off frequency is the frequency at which the response has already dropped by -3 dB.

The slope is the rate at which the level drops above the cut-off frequency, measured in dB/octave. It is easy to design filters with slopes of 6, 12, 18 and 24 dB/octave.

A low-pass filter allows low frequencies through and reduces the level of high frequencies.

A high-pass reduces the level of low frequencies.

A band-pass filter reduces the level of low and high frequencies.

A notch filter filters out a narrow band of frequencies.

Outboard EQ

No matter how good the EQ on your mixing console, there will come a time when you need to use an external or 'outboard' unit. This might be because you need a facility not available from your console EQ, or you might prefer to use an EQ unit for some subtle characteristic sheen it gives to the overall sound.

Outboard EQs come in two basic flavours: graphic and parametric. A good graphic equalizer typically has 30 or so slider controls for frequency bands nominally covering a third of an octave each. You would use two for stereo. The basic idea of a graphic is that as you set the slider controls to achieve the sound you want, the levels of the sliders 'draw' the EQ curve, as if you had measured and plotted it the long way. Unfortunately, graphic equalizers are somewhat economical with the truth and only give a rough idea of the actual curve. This is because each band does not cover only a third of an octave; its effects are felt most there but the slider will actually affect frequencies belonging to two or three bands either side of it to a distinctly noticeable extent.

Whatever deficiencies graphic equalizers may have in displaying their response curve, they are still very useful tools to have around. Mixing consoles can handle basic EQ tasks better and more quickly, but there are certain applications where graphic EQs have the edge. More on this shortly.

The alternative to a graphic outboard EQ unit is the parametric EQ. This is so called because it offers control over all three EQ parameters mentioned earlier - frequency, Q, and gain. A good parametric EQ unit may offer five bands, which cover the entire frequency range, or you might find three fully parametric bands with dedicated low and high frequency bands too.

Using EQ

Successful equalization requires good equipment and a thoughtful approach from the engineer. Experienced engineers EQ by instinct and their fingers operate the controls as fluently as a jazz pianist tickles the ivories. But this fluency doesn't come automatically, it can only be won by experience. Anyone can grab the low frequency knob and wind up the bass to the maximum, but if you are serious about your recording then you will realize that it isn't just yourself you have to please; you have to consider what other listeners like and what systems they may be playing the recording on.

There is also a good technical reason why you should think before adding a lot of bass: for a given level of input, any small or medium size loudspeaker will produce much more sound at mid frequencies than at low, and if you boost the low frequencies too much then the overall level the speaker can achieve without significant distortion is less - sometimes much less. It's a matter of compromise: the more bass you add, the lower the overall level can be. This also applies to other frequencies in the mixing console itself.

Adding EQ boost adds level, and it is very easy to boost the signal so much in the EQ section of the console that you run into clipping and distortion. Since the fader comes after the EQ, lowering the fader will do nothing to solve this. The answer is to reduce the gain, to allow the signal a little more headroom if necessary. One further technical point: changing the EQ of a signal nearly always changes the level, so each time you adjust the EQ you will have to consider moving the fader to compensate. It's something that will come automatically after a time, but newcomers to recording often concentrate more on the change in the sound itself and don't notice that it has suddenly become more or less prominent in the mix.

EQ Hints & Tips

If your mix sounds 'muddy', boost the main frequency range of each of the principal instruments. Boost 'decorative' sounds even more and pull the faders right down.

If you can't get your tracks to blend together in the mix, cut the main frequency range of the principal instruments.

To make vocals stand out in the mix, boost at around 3kHz.

For extra clarity, cut the bass element of instruments which are not meant to be bass instruments.

Adding EQ boost often adds noise. Listen carefully to arrive at the best compromise.

Changing the EQ changes the level. Always consider re-adjusting the level after you EQ.

If you add a lot of EQ boost, you may run into clipping and distortion. Reduce the channel's gain to eliminate this.

If you use EQ to reduce feedback in live work, take care not to take too much level out over too wide a range of important frequencies, particularly the vocal 'presence' range around 3kHz.

If your mixing console has an EQ Off button, use it frequently to check that you really are improving the sound.

Enough of the technical stuff, recording is an artistic occupation so let's consider the subjective facets of EQ. If we consider individual sounds first, let's assume that the signal coming from the microphone is already as perfect as can be, being the result of careful positioning and angling. Each instrument has certain bands of frequencies that are strong and some that are weaker. The human voice, for example, is strong around the 3 to 4 kHz region, no matter whether male or female, or what note is being sung. When using EQ, you will be considering which characteristics of the sound you want to accentuate, or which you want to reduce. One way to consider this might be to imagine an instrument which was an 'average' of all real instruments, where the characteristics of normal instruments were smoothed out into something that had a neutral sound. When EQing a real instrument, you will either want to exaggerate its individual characteristics and make it more distinctive, or reduce its individuality and make it more like this hypothetical 'average' instrument.

This is quite simple to do, and we can make use of the standard sweep mid range control that is found on most mixing consoles, with controls for frequency and gain. A fully parametric equalizer with a Q control can offer even more precision.

First set the gain control to a medium amount of boost - the three o'clock position of the knob is usually OK. Now sweep the frequency control up and down to the limits of its range and listen for the frequencies at which the effect is strongest. These are the frequencies in which the instrument is rich. Boosting the instrument's strong frequencies will enhance its individual characteristics and, for example, make a clarinet even more dissimilar to an oboe or any other instrument. In effect, you are making the clarinet even more clarinet-like.

When you have found the instrument's strongest frequency band, set the amount of boost according to taste and always compare what you are doing with the flat setting. If you have EQ sections to spare, you may be able to cut down on frequencies which don't enhance the sound of the instrument. Some instruments which are not known as bassy instruments nevertheless have a high low frequency content; cymbals for instance. On many occasions it will be well worth cutting down on frequencies which you don't consider to be any use to the instrument, freeing up a space in the frequency spectrum for another instrument to use.

Enhancing the sounds of individual instruments in this way is useful, but watch out when mixing that you are not boosting the same frequencies on each instrument. It is a trap for the unwary to boost every instrument at around 3 kHz to help it cut through at a frequency where the ears are very sensitive. This will produce a mix that is very tiring to listen to.

The opposite of the enhancement technique is where you lessen the individuality of each instrument and make it more like our hypothetical 'average' instrument. To do this, find the instrument's strong frequencies with the mid EQ set to boost as before, but then cut these frequencies, by as much as you feel appropriate. This won't make the instrument sound better in isolation, but it will help it blend in with the other instruments in the mix.

Many aspiring engineers do not appreciate how useful EQ cut can be, but the expert will skilfully share the frequency spectrum among all the instruments so that each has its own space and doesn't have to fight with the others for attention. Using EQ in this way can result in a powerful and full sound from a small number of tracks.

EQ With Boldness

When adjusting the amount of EQ to apply (ie. the EQ gain), it's tempting to adjust it very carefully and change the setting in small increments. The problems with this method are: (a) that if the EQ setting isn't right then it is wrong and thus needs total reconsideration; (b) that the ear quickly grows used to changes in the frequency balance of a sound.

It may not always be appropriate, but the next time you want to change the EQ level of a sound, grab the control firmly, twist it all the way up and all the way down and quickly settle on a new position which will hopefully be just right.

Mixing consoles differ in the usefulness of their high and low frequency EQs, and it is often necessary to bring in an outboard EQ that can do the job better. I would say that it is the purpose of the low frequency control to add 'weight' to the sound without making it 'boomy'. These are subjective terms, but we can all appreciate the difference between a sound which is firm and solid in the bottom end, and one which has plenty of bass but gives the impression of being out of control. In the other direction, the low frequency control should cut low frequencies that are not contributing anything useful to the sound, while retaining the depth and body of the low mid. At the high frequency end, you should be able to cut any 'fizz' from the sound while still leaving it clear and incisive, and you should be able to make the sound brighter without the extreme top becoming aggressive. If you can't achieve all this with your console's EQ, you may have to spend a considerable amount of money on an outboard unit that can.

When you have explored all the possibilities your console's EQ can afford and you have visited your local hire company for outboard units that perform the same function only better, you'll be keen to get your hands on a graphic equalizer. This is a rather different animal which appears at first to offer the ultimate in flexibility: just raise or lower the frequency bands you are interested in for quick and precise control. Unfortunately, you will find that precision is lacking because each individual band alters frequencies over quite a wide range on either side of its nominal centre frequency.

This doesn't mean that graphics are useless far from it. Graphics are great for EQing an entire mix so that you can shape the sound as a whole, even after you have processed the individual elements. If you know your way around, you can do this by taking a couple of outputs from the mixing console back into two channels and using the console's EQ again, but you'll only be applying more of the same, and doing it the graphic way really is much more satisfying. Graphics are also great for adding bite to a sound: just raise one or two sliders somewhere in the upper frequency region and you will make the sound more cutting without lifting the whole of the high frequency range. Experiment at your leisure.

Problem Solving

If you are working on a tape made by another engineer who isn't quite as fastidious as you, then you may find yourself faced with problems that EQ can help rectify. Unwanted sounds have a knack of finding their way onto recordings, particularly live recordings. If you have a 50 Hz or 60 Hz (depending on which country you live in) mains hum, for example, then a graphic will be able to help at only a little loss to the musical sounds on the recording. You can also use a parametric equalizer set to a high Q to home in on the unwanted frequency. Some equalizers have special notch filters to cope with precisely these situations. 50/60 Hz hum may be removed to a reasonable extent, but the buzz caused by lighting dimmers may be impossible to get rid of. If the buzz isn't too harsh then you can try cutting the 50/60 Hz fundamental and its harmonics at 100/120 Hz, 150/180 Hz, 200/240 Hz etc. No promises, but it may make the recording just listenable.

Apart from hum or dimmer noise, if a recording is too noisy then very often the noise is most noticeable at high frequencies. Here you can use your EQ to strike the best compromise between cutting as much of the offending component of the noise as possible while still retaining some brightness in the sound. You may be able to apply a little boost at high mid frequencies, although the result will remain a compromise.

Even if the recording has no hum, buzz or noise, it may previously have been over-EQ'd. It is quite difficult to ameliorate the results of over-zealous EQing, particularly if some frequencies have been cut to a large extent. Trying to boost these frequencies back up again may result in an unacceptable amount of noise becoming apparent. Once again, compromise is necessary, although if you were dealing with one instrument from a multitrack mix you may be able to patch in a noise gate to help in this instance.

EQ Terminology

CUTOFF FREQUENCY: The frequency at which a high or low frequency EQ section starts to take effect. Also referred to as turnover frequency.

SLOPE: The rate at which a high or low frequency EQ section reduces the level above or below the cutoff frequency. Usually 6, 12, 18 or 24 dB/octave.

PASS BAND: The frequency range that is allowed through.

STOP BAND: The frequency range that is attenuated.

FILTER: An EQ section of the following types:

HIGH PASS FILTER: A filter section that reduces low frequencies.

LOW PASS FILTER: A filter section that reduces high frequencies.

BAND PASS FILTER: A filter section that reduces both high and low frequencies.

NOTCH FILTER: A filter that cuts out a very narrow range of frequencies.

GAIN: The amount of boost or cut applied by the equalizer.

Q: How broad or narrow the range of frequencies that is affected.

SWEEP MID: A middle frequency EQ section with controls for frequency and gain.

PARAMETRIC EQ: An EQ section with controls for frequency, gain and Q.

GRAPHIC EQ: An equalizer with a number of slider controls set on octave or third octave frequency centres.

BELL: An EQ with a peak in its response.

SHELF: A high or low frequency EQ where the response extends from the set or selected frequency to the highest or lowest frequency in the audio range.

HF: High frequencies

LF: Low frequencies

MID: Midrange frequencies

TREBLE: Hi-fi enthusiasts' word for HF.

EQ OFF BUTTON: The sign of a good mixing console!

Check Questions

- Summarize the 'three rules' of EQ.
- List ten words that are used to describe the EQ characteristics of a signal in subjective terms.
- What is 'cut-off frequency'?
- What is 'slope'?
- What is a 'notch filter'?
- What are the three controls of a parametric EQ section?
- What is the meaning of 'Q'?
- What is the meaning of 'bandwidth'?
- What is the meaning of 'shelf'?
- What is the meaning of 'bell'?
- What is the fallacy of graphic EQ?
- Why is it necessary to consider adjusting the level after EQing?
- Why might it be necessary to adjust the gain after EQing?
- Why is it not necessarily desirable to add extreme LF boost, even if you like lots of bass?

Chapter 4: Compression

The function of the compressor is to reduce dynamic range. That is, to reduce the different in level between loud parts of the signal and quiet parts. It does this by reducing the level of the loud sections.

The natural sounds of life have an extremely wide dynamic range, from the rustle of a falling leaf to the roar of a jet engine on take off. The human ear has an automatic gain control which enables it to accommodate all of these sounds from the threshold of hearing to close to the threshold of pain, a dynamic range of approximately 120 decibels.

Even the most modern audio equipment is incapable of handling the full range that the ear can cope with. Analog tape without noise reduction can manage almost 70 decibels dynamic range between its noise floor and the 3% distortion point. 16 bit digital audio equipment can achieve over 90 dB. Still almost 30 dB less than the ear's range.

Even with a theoretical dynamic range of 144 dB (which would be possible in 24-bit digital equipment, given perfect analog to digital converters), would it be desirable - and useful? A listener in a domestic setting might enjoy the exhilarating effects of levels up to 100 dB SPL (Sound Pressure Level) and more, but what annoyance or distress might that be causing to his neighbor? At the other end of the dynamic scale, a typical ambient noise level of at least 40 dB SPL precludes the use of very quiet levels in recorded or broadcast sound media.

Almost always, it is necessary to compress the dynamic range of natural sounds to fit them into a window suitable for comfortable listening.

Focusrite Red 3 compressor

Use of Compression

One of the principal uses of compression is the control of level in vocals. Many singers train for years to achieve the degree of breath control necessary for an even tone and expressive performance. Other vocalists rely on an instinctive voice production technique, which may need help in the studio to maintain a consistent level, and result in a vocal track which 'sits' correctly in the mix.

The level of a vocal may vary widely, and appear like the unprocessed signal (a) in the diagram:

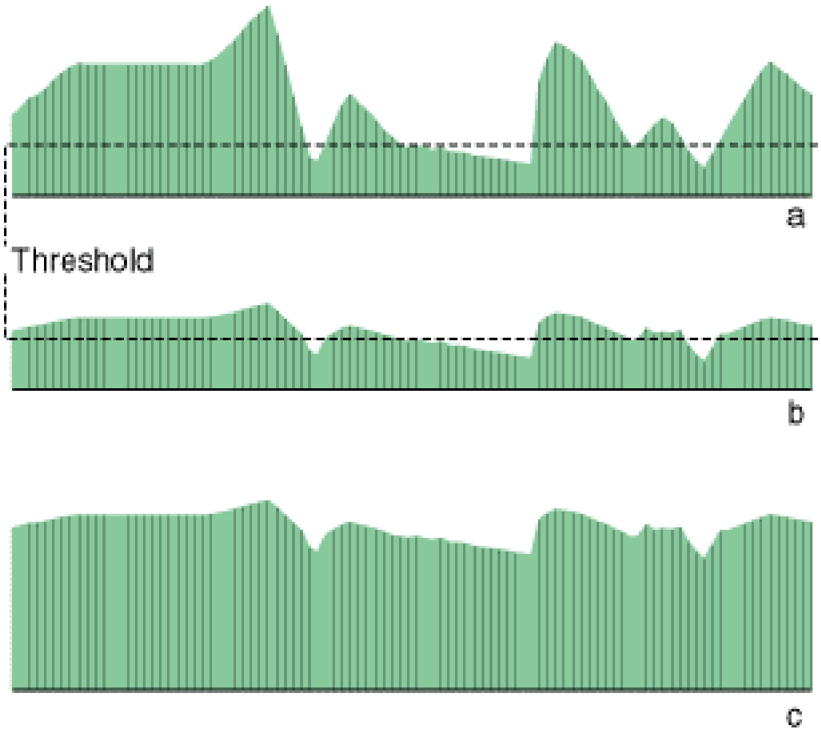
-The unprocessed signal has a large dynamic range between the highest and lowest levels. Applying compression reduces the highest levels, reducing the dynamic range (b). Because the peak level of the signal is now lower, make-up gain is added to restore the original peak level (c). The result is a much more controlled and usable sound.

Interface with the console

Compressors work at line level, therefore the input signal has to be taken from the mixing console, preferably from the channel insert point send. The output from the compressor is brought back to the channel insert return. By connecting the compressor at this position in the signal chain, its operation is unaffected by the use of any of the console controls, except input gain.

An alternative is connection to the group insert point of the console, or the main stereo output's insert point. In either of these situations, a mix of signals is compressed.

Setting the Controls

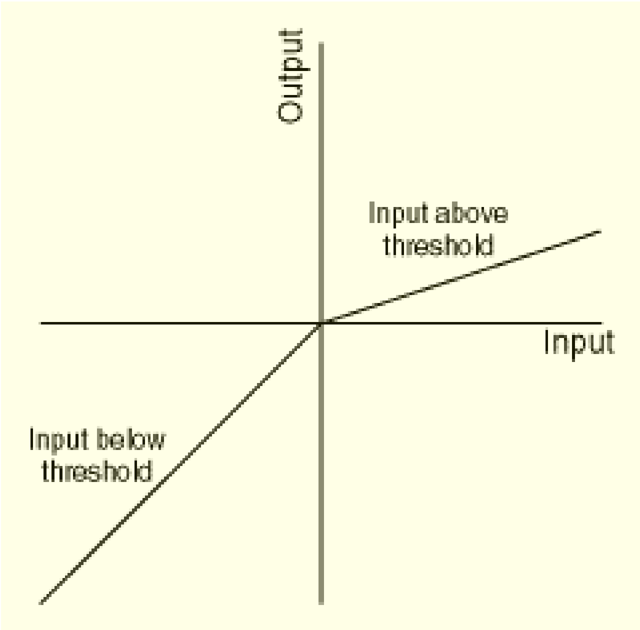


Threshold sets the level above which compression takes place. Signals below the threshold will remain unaltered.

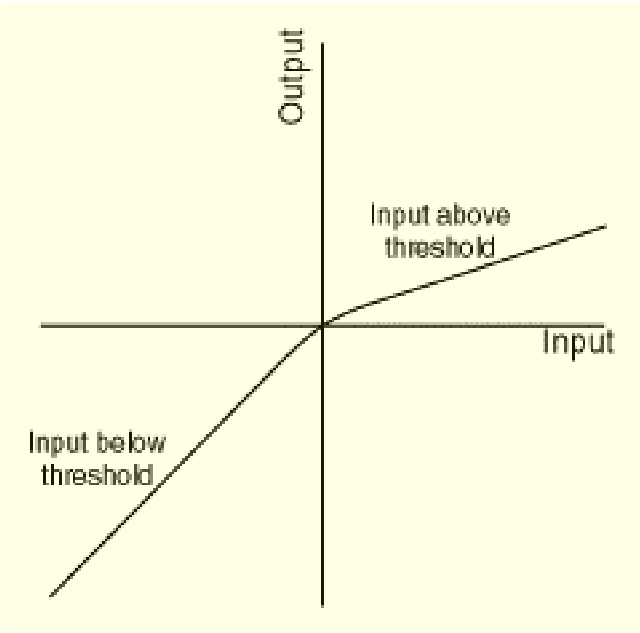
Ratio is the 'strength' of compression above the threshold level. The higher the ratio, the greater the effect. If the ratio is set at 5:1, it means that when the signal is above the threshold level, when the input signal rises by 5 dB, the output signal rises by 1 dB.

At a compression ratio of 2:1, the effect is mild and suitable for the subtle compression of vocals or for a complete mix. At 10:1, compression is much stronger and more noticeable. Ratios between 5:1 and 15:1 are suitable for the 'compressed' sound, used as an effect in its own right. Higher ratios are used for the control of extremely peaky signals. Above 20:1, the compression effect is so pronounced that it is known as 'limiting'. It is possible to buy a dedicated limiter.

The point where the slope of the compressor curve changes is known as the Knee. Some compressors have an adjustable knee, variable between hard (which is normal) and soft:



Hard knee



Soft knee

With a soft knee, signals which only just exceed the threshold level are compressed at a low ratio, the ratio increasing the higher the signal level.

Attack sets the time the compressor takes to respond once the threshold has been exceeded. Attack may be set so that the initial transient of the instrument passes through unaltered, or set to a faster value so that the very start of the sound is compressed. Particularly with drum sounds, careful adjustment of attack time can make the sound more 'punchy' and 'driving'.

Release time plays a very important role in compression. During periods of high signal level, gain is reduced. When the signal level falls below the threshold, the gain will increase at a rate determined by the Release control. If the release time is short, the gain will rise quickly. A long release time will mean that the gain will stay at its reduced level, only recovering gradually:

The setting of the correct release time is a compromise. If the release time is too short, background noise can cause effects often known as 'breathing' and 'pumping'. If the release time is too long, the signal will not be compressed, but simply reduced in level. For effective compression, the release time must be set to as short a value as possible before modulation of the background noise becomes too noticeable. The gain reduction bargraph meter will show how much actual compression is going on. If it stays steady, there is little active compression, just a steady-state reduction in level. The faster the bargraph moves up and down, the harder the compressor is working.

Gain Make-Up restores the level lost in the compression process. Since the compressor works by bringing down peak levels, the level of the output signal would be lower than the input if nothing were done. Sufficient gain make up should be applied so that the peaks of the compressed signal are the same level as the peaks of the inputs signal. The sections of the input signal that were quiet will now be louder.

Stereo Link: When a stereo signal is compressed, the stereo link has to be activated so that both channels provide the same amount of gain reduction. If this is not done, a loud signal in one channel will cause that channel to be lowered in level while the other stays the same. Any signal that is panned center in the mix will swing in the stereo image towards the unaltered channel. With stereo link selected, the stereo image is maintained.

Compression Noise

Compression always has the effect of increasing the noise level. This is because the peaks of the signal are brought down in level, bringing them closer to the noise floor. Then make-up gain is applied to bring the overall signal level back up again, raising the noise floor at the same time. Even if there were such a thing as a perfect compressor, this would still happen.

Side Chain

In addition to the normal signal input, a compressor has a 'side chain' input.

In normal use, the amount of compression or expansion is related to the dynamics of the input signal. The side chain allows the signal passing through the unit to be controlled by the dynamics of another separate signal.

De-Essing

De-essing is an important compression technique using the side chain. Many singers have high level sibilants - 'sss' sounds - which detract from the quality of their performance. Equalizing the signal will reduce the sibilants, but also make the overall vocal sound dull. The sibilants can be selectively removed by compressing only when there is an excessive level of high frequencies.

The microphone channel is routed to a group with the compressor patched into the group insert points. The microphone channel is also paralleled into another channel via the line input. The signal in the second channel is equalized so that high frequencies in the sibilant range are boosted. This channel is fed via an auxiliary output to the compressor side chain input.

Now, the compressor will react whenever there is a sibilant, reducing the gain for the duration of the sibilant and cleaning up the vocal sound.

This technique can also be used to compensate for a 'boomy' bass, or other situations where a band of frequencies is occasionally obtrusive.

Character

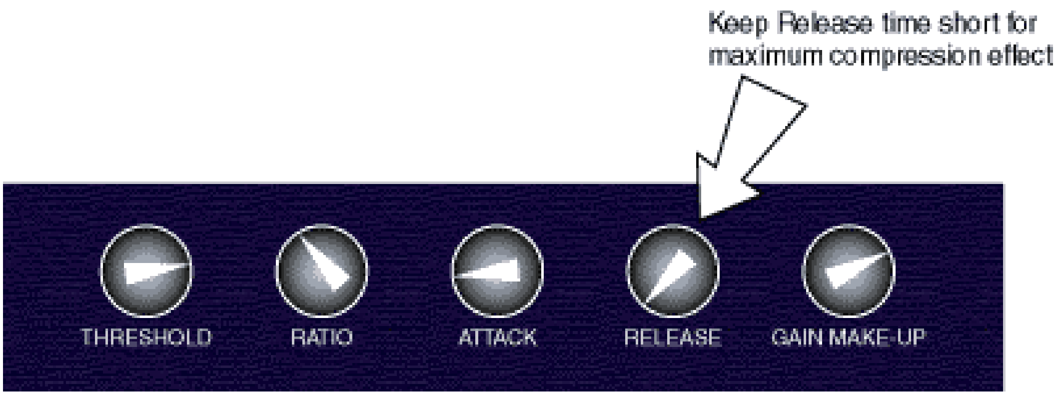
One feature of compressors is that they all seem to have their own individual sonic character, even more so than equalizers. This is due to the 'ballistics' of the attack and release profiles, to any processing applied to the side chain, and to any distortion produced in the gain change element, particularly if tubes (valves) are used in the circuitry.

Advanced Compression

The Hidden Compressor (By David Mellor, originally published in Audio Media)

Every studio has one, every engineer uses one, and every popular music recording - almost - dating back to the 1950s and beyond has benefited from one. Of all the many and varied types of outboard in the processing and effects racks, the compressor is surely the one that is most often used, and one that repays its cost of ownership countless times over during its working life. So I don't need to tell you anything about compressors then? Maybe not - if there does happen to be anything you don't know already then you can easily find it in textbooks and magazine articles that are often aimed more at the beginner than the seasoned pro. However, the compressor is a many faceted instrument, and there are a number of tips, tricks and techniques that are not commonly covered in print. Are these the compressor's secrets, known to the few and hidden from the many? Like the Masked Magician, I intend to reveal these secrets to the world.

Merciful Release

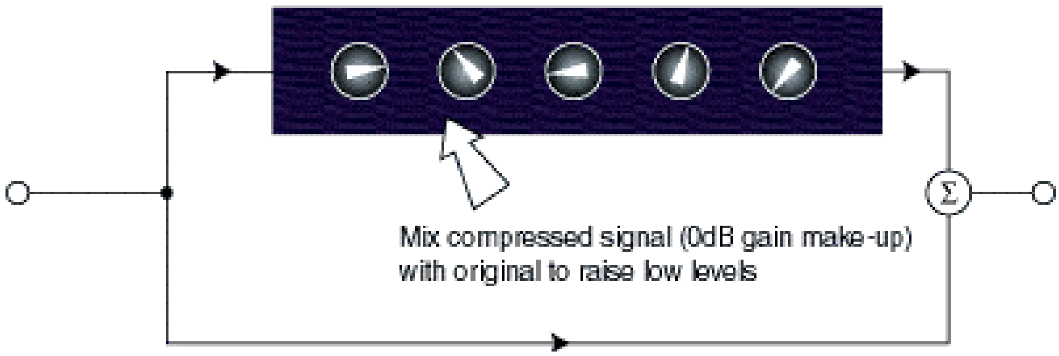


A long time ago when I was a fresh-faced student of sound engineering, I went to a trade show (in the days when you had to blag your way in, if you weren't in the business already) and alighted on the stand of a company who had a new and wonderful compressor to show off. "Listen to this", said the silver-tongued salesman. I listened as he demonstrated his amazing box. "That's 30 dB of compression. Does it sound compressed to you?". I looked at the gain reduction meter, I listened, I looked at the gain reduction meter, I listened. Sure enough, the meter was showing a full 30 decibels of gain reduction and the music I was listening to sounded as fresh as a live performance. I knew something about compressors, and I knew that 30 dB of gain reduction ought to be the sonic equivalent of what an apple looks like after it has been through a cider press. It's a good job I didn't have any money or I might have bought it on the spot. With the benefit of experience I know what happened. I am sure that it was a reasonably good compressor, but not significantly better than any other. What the salesman had done was to turn the release control to maximum. Release, as you know, is the time it takes for gain reduction to return to zero after the signal has passed below the compression threshold. In this case, the signal never passed below the threshold long enough for the level to begin to return to normal, to any significant extent. The result was indeed 30 dB of gain reduction, but not 30 dB of compression. You don't need a compressor to get any amount of gain reduction - just lower the fader. Compression implies a constantly changing amount of gain reduction, and the gain reduction meter must be visibly dancing up and down. If it's not, you're wasting your time. How fast it dances up and down is up to you, but if you want value-for-money compression, a short release time will give you a more audible compression effect. A longer release will lessen the audibility of the compression, but you won't actually get as much real compression.

Over Compression

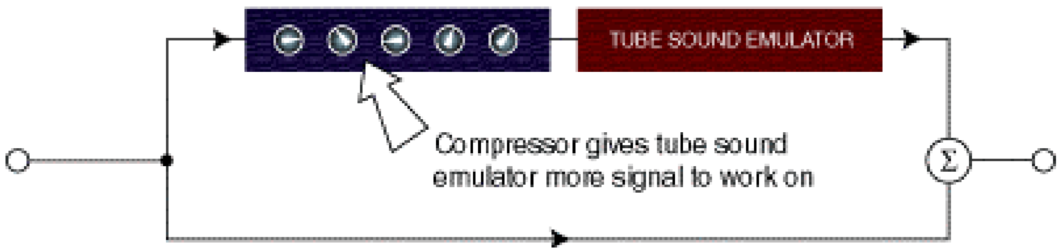
No-one reads the manual for a compressor, and if you did you wouldn't get any warning about the effects of over compression. I don't mean this in the sense of too much compression, your ears will tell you that, but in the sense of setting a lower threshold than you need to get the job done. This will always make the sound worse, with the sole exception of percussive sounds where it might sometimes be a useful effect. Let's assume a scenario where an instrument plays occasionally with silences in between. This is where over compression is most likely to happen. When setting the threshold, many users have an idea of how much gain reduction they want to hear, and see on the meter. The amount of gain reduction is controlled both by the threshold and ratio controls. Suppose these controls are set so that the desired amount of gain reduction, let's say 12 dB for example, is achieved. This should be fine shouldn't it? Look again at the gain reduction meter. While the instrument is playing, does it ever go all the way down to zero? If it doesn't, if it only goes down to 3 dB then you haven't applied 12 dB of gain reduction, you only have 9 dB of compressive gain reduction. The other 3 dB could have been achieved by simply lowering the fader. This in itself isn’t necessarily a problem. The problem is that when the instrument starts to play, the compressor has to go all the way from zero gain reduction to the full 12 dB. The necessity of covering that additional 3 dB will audibly distort the initial transient. Try it, and you will hear it for sure. This leads to rule number one of gain reduction - at some point in the course of the track while the instrument is playing, the gain reduction meter must indicate zero, otherwise the minimum reading obtained shows wasted gain reduction and over compression leading to the distortion of transients that follow silences.

Compression by Stealth



One of the best known uses of compression is to increase the apparent loudness of a mix, or an individual voice or instrument for that matter. Compression, as you know, works by reducing the high signal levels, bringing them closer to the low level passages, and then applying make-up gain. Thus the low level signals are brought up and the whole thing sounds louder. This is fine in theory, the trouble is that the effect of compressing the high level signals is very audible necessitating great care in the set up of the compressor and judicious compromise between getting enough compression and not spoiling the overall sound. Ray Dolby told us this when in the early A-type noise reduction system he left high levels completely alone and modified the gain only of signals below -40 dB. What we need is a compressor that only operates on low level signals. Is there such a thing? Yes there is, and it's in your rack already. You just have to use it in a different way. Since in this situation the object is to bring up the lower levels of the track, what we need is a way of making the quiet sections louder without affecting the loud sections. The answer is to mix the uncompressed signal with a compressed version of the same. At levels below the compressor's threshold the two signals will combine to produce a 6 dB increase in level. Above the threshold the compressed signal will be progressively reduced and add hardly any additional level to the mix. The result is a form of compression where you can get more dynamic range reduction with fewer audible side-effects. I'm not going so far as to say that it is always best way, but it's certainly worth a try. Maybe some enterprising company will bring out a gadget to do just this, in a convenient rack-mounting package. By the way, if you try this with a digital compressor you will get a lesson in the delay involved in digital processing. You will get comb filtering and it will sound dreadful.

Compression vs. Clipping



While I'm on the subject of increasing apparent loudness, I don't know whether it is as widely appreciated as it should be that compression is only half the answer. Compression is a long-term type of gain reduction working at the very least over periods of tens of milliseconds. If you try to achieve very fast acting compression by using very short attack and release times, you may well end up with distortion of low frequencies where the compressor actually changes the shape of the waveform. There comes a point in maximizing apparent loudness where the compressor has given all it has got to give. Clipping on the other hand works on a very short time scale. Transistorized circuitry reacts within microseconds to any level that is too great for the power supply to cope with and cuts it short, creating harsh harmonics, but at the same time extra loudness. The soft clipping of valve and valve-emulating designs rounds rather than clips the peaks but once again operates on a short time scale. The problem with soft clipping if used alone is that it only works on high level signals. Clip-worthy peaks only occur in quantity in high level signals and low level signals, although they may indeed have the occasional clippable peak, are largely unaffected. The answer is to use a compressor and a soft clipper in series. The compressor evens out the general level of the signal, but since it works over a comparatively long time scale, the peaks are not clipped but simply brought to a more uniform level. The clipper then has more material to work on. A useful alternative is to use a series-parallel configuration as shown here. Here, the compressor smooths out the levels, the valve-emulation device soft clips the peaks, and the result of that whole process is added to the uncompressed signal. The result is controllable enhancement over a wide range of levels. If you want to go further then you might add an equalizer after the compressor so that you can choose the frequency range that will be affected to add just the right hint of distortion without going over the top, particularly in the mid range.

MS Compression

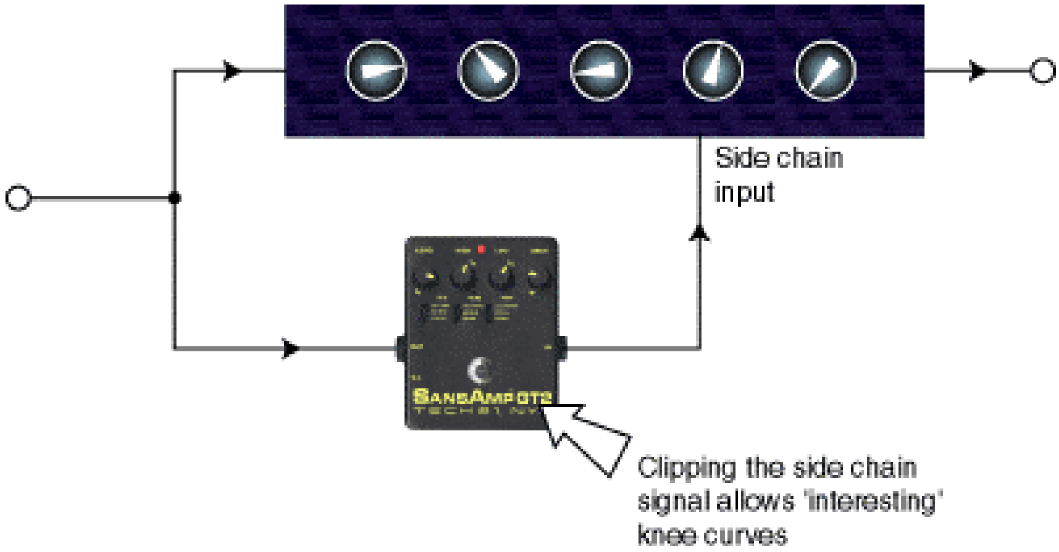
Here's an interesting curiosity. As you know when compressing a stereo signal, a two-channel compressor must have its side chains linked, otherwise heavy compression in one channel will cause an image shift in the stereo sound stage. Both channels must at all times be compressed equally. This of course assumes that you are handling stereo as left and right channels - let's call this LR stereo. Not as popular but certainly very useful is mid-side or MS stereo where the M channel is the mono sum of the whole sound stage and the S channel represents the difference between left and right. MS is a useful microphone technique and is sometimes used at other points in the signal chain for modifying the width of the stereo image. (It's a funny thing that proponents of MS often forget that you can do that to LR stereo signals with the pan controls).

But what about compressing a signal in MS format? Is it possible? Does it have anything new to offer?

Yes it is possible to compress MS signals without converting them to LR. Just pass the M signal through one channel of the compressor and the S signal through the other. Once again, you will need to link the side chains or funny things will happen, but it will all work perfectly. Some might say that it works better than compressing LR stereo, since even when side chains are linked it is not guaranteed that analogue compressors will handle both channels absolutely equally and some image shift may persist. But if you compress in MS domain then any disparity between the channels will result not in an image shift, but a variation in the width of the stereo image, which is arguably less obtrusive. But why not take this a stage further and do something really wacky like compressing the S signal only. What happens now? If you compress the S signal only, then anything panned centre is unaffected and compression only affects signals panned left or right, or signals that are out of phase. Loud signals in these modes will cause a momentary reduction in level of the S channel resulting in a narrowing of image width. I can't say that I recognize any useful function for this myself, but in the hands of more creative people, who knows?

Serious Side Chain

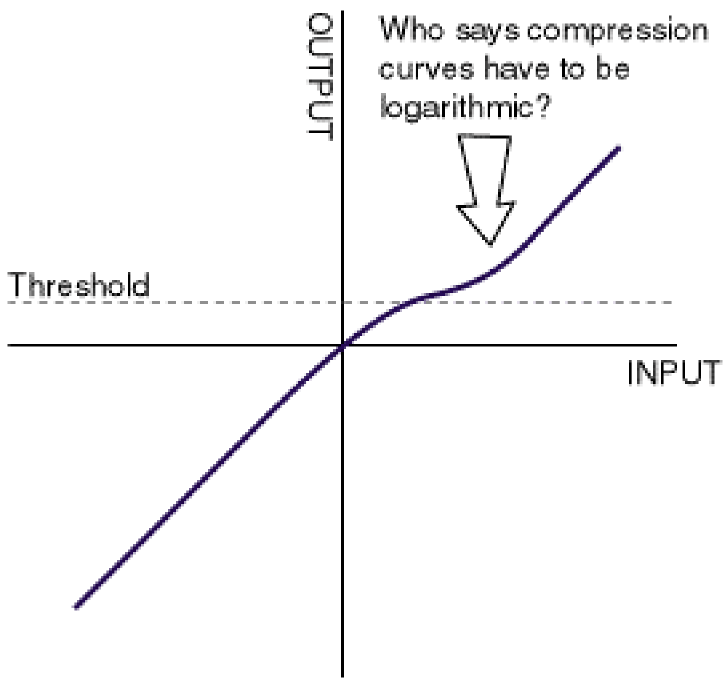
Everyone knows how to direct a high frequency boosted signal to the side chain to perform a crude type of de-essing - now superseded by more sophisticated stand-alone de-essers such as the Drawmer MX50. But what about applying EQ to the side chain in general, rather than this one specific application? If you have never done it, do it now. Parallel a signal so that it enters the normal input of the compressor, and at the same is connected to the side chain input via an equalizer. Now play some signals through this set up. We all know that different compressors have different sounds, but this little trick allows the compressor that's in your rack right now to have an incredible range of sounds going far beyond the normal differences between models, when used in the standard configuration. You will find that the compressor becomes another type of EQ, but instead of simply cutting or boosting different frequencies, you allow different frequency bands to control the amount of compression applied. When you are in search of that elusive ‘phat’ sound and simple EQ and compression are not getting you there, EQing the side chain might just do it for you. In fact I would go so far as to say that all serious compressors should have side chain EQ built in. Once you have really tried it you won't want to do without it.



The side chain can do more. Everyone knows that different compressors sound different, and that soft-knee types are more subtle than hard-knee, which go immediately from uncompressed to compressed at the exact threshold level rather than the gentle blending of the soft-knee type. The precise knee curve of a compressor is an important factor in its sound, but few compressors allow you to modify the knee curve in any way. So can it be done? Well of course it can, otherwise I wouldn't have mentioned it. Here's the deal: set up a side chain configuration as above, but this time instead of an equalizer, insert a distortion processor. A guitar effects unit such as the SansAmp GT2 would be fine. Remember that you are not going to hear any signal coming out of the side chain, unless there is some internal crosstalk within the compressor, so the output signal isn’t going to be distorted. What the GT2, or similar device, will do is apply soft or hard clipping which will bend the shape of the knee curve of your compressor. What effect this has depends on the compressor itself, on such factors as whether peak or RMS detection is used for example. The result will be however that you will feel as though you have a totally different compressor in your rack. In fact, when different settings are used on the distortion box you will feel as though you have installed a whole rack full of different compressors.

Another option for the side chain is to insert an advanced version of the signal to control the level of the signal itself. One of the enduring problems of compressors, and gates for that matter, is that they can only react to whatever information they receive, they can never anticipate what is going to happen and prepare for it. Well now they can. Using a digital tape or hard disk multitrack it is commonly possible to delay individual tracks with respect to the others. Even if it isn't possible to advance a single track, you can always delay the rest, and perhaps make a delayed copy of the track you want to process. Armed with this you can connect the advanced version of the track to the side chain - just 50 to 100 milliseconds should do - and the delayed version to the normal input. Now you will find that the compressor anticipates the amount of gain reduction required and transients in particular are rendered very much more realistically than doing things the normal way. In fact, you can do it the other way round too - delay the side chain so that the compressor takes a moment to react. "Why would you want to do this?", you might ask. The answer is that percussive sounds often benefit from a relatively slow attack, allowing the initial transient to come through unaltered before the 'body' of the sound is compressed. This is just a different way to do it, but this time with a little more control.

Radical Ratios

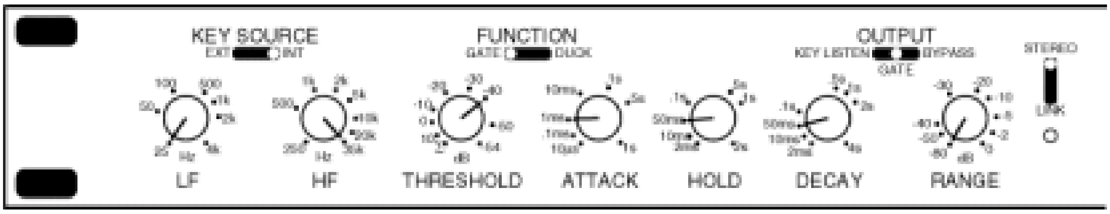


When is a compression ratio not a ratio? I could give you a straight answer but instead I would like to ask another question. Whoever said that it should be a ratio? Some scientist I don't doubt. Virtually every compressor on the market offers logarithmic compression, such that once the knee curve is passed then, for example, at a 2:1 compression ratio a 10 dB increase in level at the input will result in a 5 dB increase in level at the output. This is all very tidy, but I wonder whether this is always going to be the right approach? How about a compressor where once the signal exceeds the threshold it is subjected to a knee curve leading to logarithmic compression, as tradition dictates, but beyond that the compression is lessened and the curve reverts to a straight line, meaning no compression. Here, signals of a certain level are compressed, but louder transients are substantially unaffected. With traditional compression, it is usually the transients that cause the problems, so once you have got the general run of signal sounding pleasant, along comes a transient and the whole thing goes crazy for a second. Why not just let the transient through so it can be on its way, and concentrate on the parts of the signal that will really make a difference. You can always limit the transient later if you need to. There is actually a range of compressors that do depart from the traditional logarithmic curve. I'll give you a clue - they are all bright green in color. But there's a whole world of options waiting to be explored, by users and by designers. Compression the way it is commonly done is boring in comparison with what it could be. Why not have a bit of fun and experiment? Most of the ideas I've outlined here won't cost you a penny, and you may never have to buy another compressor again because you're getting all the fun you need from the compressors you already own!

Check Questions

- Does a compressor act upon loud signals or quiet signals? What does it do to the signals upon which it acts?
- Why is it impractical for a recording to have a wide dynamic range when played in a domestic listening environment?
- What is the function of 'make up gain'?
- How is a compressor usually connected to a mixing console to compress a single signal?
- How is a compressor connected to a mixing console to compress the entire mix?
- Explain 'ratio', in the context of compression.
- Describe the difference between hard knee and soft knee.
- What are 'breathing' and 'pumping'?
- What happens if the release time is too long?
- What does it mean if the gain reduction meter is showing frequent changes?
- Describe the effect of compression on noise level.
- Describe the function of the stereo link switch.
- What is the side chain input?
- Describe a typical use of the side chain input.

Chapter 5: Noise Gates



The noise gate has two functions:

- To assist in managing noise.
- As a creative device.

In its principal role of managing noise, the noise gate can be applied to a single unmixed signal that has periods of silence, during which noise is apparent. This is important enough to state again, that the noise gate can be used thus:

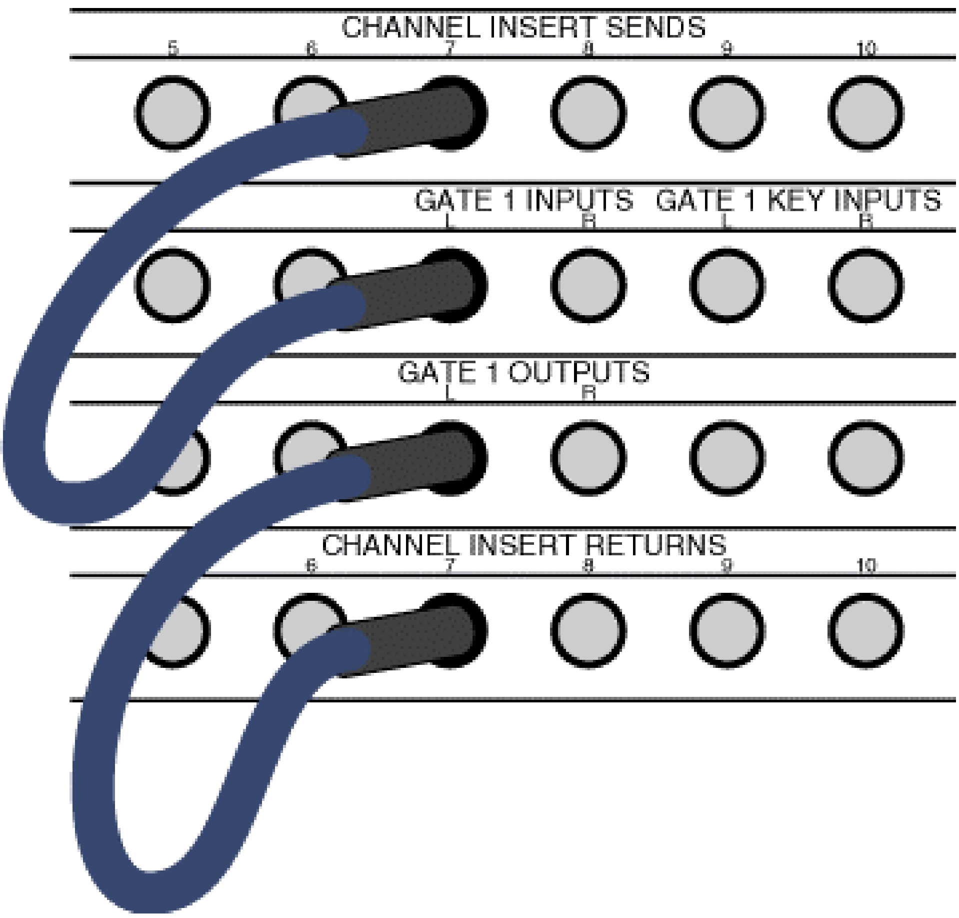
- On a single, unmixed signal.
- The signal has periods of silence (i.e. when the instrument isn't playing), during which noise is apparent.
- It cannot be used on a mix of signals (at least this is very unlikely).
- There is no benefit in using it unless there are sections of silence.

The classical example of where the noise gate would be appropriate is the electric guitar, where the amplifier is likely to be noisy. When the guitarist plays, the sound of the instrument drowns out the noise so there is no problem. When the guitarist isn't playing however, the noise of the amplifier becomes apparent, and irritating. In this situation, what the noise gate does is detect when the signal level is high, when it assumes that the guitar is playing, and opens fully to allow the signal through unimpeded. When the gate detects that the signal level is low, it assumes that the guitarist is not playing and closes completely, blocking off the noise.

Another classic use of the noise gate is on a drum kit, where several might be used. The conventional method of miking up a drum kit demands a mic on each drum, a mic on the hihat and two overhead mics.

The problem with this is that the mics are all so close together that each mic picks up every instrument of the kit to an extent, as well as its own instrument. This inevitably blurs the sound. To make it more focussed, all the mics except the overheads are gated so that each channel is only open when the drum is actually sounding. In this situation it is often subjectively better to set the gate so that it attenuates when closed, by perhaps 10 dB, rather than cut off completely. In live sound, this technique is often extended to virtually every microphone for the entire band.

Since the noise gate can only process one unmixed signal, the place to connect it is in the channel insert point of the mixing console. As with equalizers and compressors, gates are not used via the aux send and return loop. If the console you are working with does not have a patchbay, you will have to make up a special adapter (Y) lead if your console has the usual single stereo jack send/return insert point.



One feature of the noise gate that contrasts with other kinds of outboard units is that there are no presets, and that twiddling the knobs will not produce a satisfactory result by chance (as it sometimes might with an effects unit).

Controls

It helps to have a suitable sound source playing at this point, one with sections that should be silent, but in which noise is irritatingly apparent. We'll go through the controls, one by one. It's good to have a rough starting point, and the example above of the classic Drawmer DS201 is a good one. With another model of gate, simply set similar values for levels and timings.

The Threshold control sets the level at which the gate will decide whether to open or close. When the signal is above the threshold, the gate will be open. When it is below the threshold, the gate will be closed. With the control set at minimum, the signal will always be above the threshold, even when only noise is present. With the control set to maximum, then the signal will never rise above the threshold and the gate will be closed. Somewhere in between there will be an optimum point where the wanted signal gets through, and the unwanted noise is blocked. It is important to make sure that when the musician plays, the first note gets through, right from the initial attack. Also, when the musician stops, the decay of the last note is maintained adequately before the gate closes. There are controls other than the threshold control that influence this, but the threshold must be set as precisely as possible before continuing.

The Range control sets the degree of attenuation when the gate is closed. Noise gates are commonly set to maximum attenuation unless there is a good reason to do otherwise. On a single signal, the gating effect will be obvious, but of course it should not be so in the context of the entire mix. If the opening and closing of the gate is still noticeable, then the range control should be set to achieve the best compromise.

The Hold control sets a time period during which the gate will remain fully open, even though the signal has just dropped below the threshold. If the hold time is set to zero, when the signal crosses the threshold there will be a period of uncertainty when the gate doesn't know for sure whether it is supposed to be open or closed and it will change state rapidly a number of times, causing what is sometimes known as 'jitter' (not to be confused with digital jitter). The hold control is usually set to the minimum value that causes jitter to cease. (Some gates alternatively have a Hysteresis control. This sets a separate threshold for signals that are rising in level than for signals that are falling in level. Having a variable hysteresis control is actually better than having a hold control.

The Attack and Decay (also known as Release) controls are used to shape the envelope of the sound as it comes in and goes away with the object of changing smoothly from silence to signal, then signal back to silence, without cutting off any of the wanted sound, nor letting any noise get through.

Stereo Link is a function only used when applying a twin channel gate to a stereo signal. An example would be an instrument played through a noisy old analog chorus unit (because you just like the sound!). When this is on, both channels are forced to open and close at the same time. If this is left off for a stereo signal, no matter how carefully you set the threshold controls the channels will change state at slightly different times. You won't believe how dreadful it sounds until you try it.

Expander

An expander is a more sophisticated form of gate. Whereas a gate is either on or off, an expander increases the dynamic range when the signal is below the threshold.

When the signal is above the threshold level, there is no change.

When the signal is below the threshold, the quieter the signal gets, the expander will make it even more so.

The expander has a ratio control like a compressor. 1:1 means no expansion; 20:1 means that the expander is working almost like a gate.

The expander isn't as useful a device as one might think. The problem is noise modulation - as the signal level goes up and down, below the threshold, the noise level will also go up and down. This is an unpleasant effect known as 'breathing' or 'pumping'. Breathing and pumping can be evident in the compressor, but since it is high level signals that are being processed, they are better able to mask the noise.

Filters

Basic gating is easy enough when the sound source is well differentiated between signal and noise. But there are times when the noise is so high in level that setting the threshold to a point that fits neatly between the signal and the noise is impossible. A good example would be the drum kit, particularly the snare drum and hihat. The snare and hihat are so close together that the snare mic will pick up as much level from the hihat as it does from its own instrument. It will be impossible to set the threshold for correct gating.

The solution to this problem is found in the filter controls. Just like a compressor, a noise gate has a 'side chain' signal that is used to control the behavior of the processed signal. In a gate, this is commonly called the trigger, or key signal. In normal use with the filters set to their end stop positions - so that they have no effect - the key signal is identical to the input signal. Indeed it is just tapped off from it. To stop the hihat from opening the gate on the snare mic, simply filter out the high frequencies from the key. The snare is sufficiently rich in low frequencies that it will still open the gate, but the lesser LF content of the hihat is not enough to do the same. If there was a mic on the hihat suffering from the same problem, then the low and mid frequencies could be filtered out of the key, and the rich high frequency energy of the hihat would still open the gate.

One important point to note is that the filters have no effect on the frequency balance of the output signal. They only affect the key, which once it has done its job goes nowhere. There is a function, here called 'key listen' that switches the key signal to the output. This can be used for setting up the gate, then switched back to 'gate' for normal operation.

External Key

Suppose you have done all of the above, but the gate still isn't opening and closing reliably on the snare drum, to continue this example. What next? The answer is to use an external key signal to open the gate. Here's the scenario:

- Feed the signal from the regular mic through the gate as normal.
- Tape a contact mic to the shell of the snare drum.
- Feed it via a preamp to the external key input of the gate.
- Switch the gate to external key ('EXT').
- Set the threshold etc. for reliable triggering.

Now, the gate is triggered by a signal that picks up virtually no external sound - only the sound of the drum. What's more, the signal from the contact mic comes maybe half a millisecond earlier than the signal from the regular mic, which allows the gate to open a little in advance so that the transient of the drum is accurately captured. It has to be said that this is probably an over-elaborate technique for most circumstances. But it works very reliably and is worth knowing about for the occasional difficult situation.

Compressor and Gate

The gate is often thought of as a companion to a compressor. Compression always has the effect of increasing the noise level. This is because the peaks of the signal are brought down in level, bringing them closer to the noise floor. Then make-up gain is applied to bring the overall signal level back up again, raising the noise floor at the same time. Even if there were such a thing as a perfect compressor, this would still happen.

The obvious answer is to use a noise gate to remove the noise, providing the signal meets the criteria for gating as outlined above. There are two schools of thought among engineers as to how the two should be connected:

Some think that it is better to put the gate after the compressor, since the compressor will generate some noise itself, in addition to the noise emphasized by the compression process.

Other engineers will say that since the signal has a greater dynamic range before compression, it is easier to set up the gate for reliable triggering if it comes before the compressor.

Gate Effects

The noise gate is also capable of a variety of envelope shaping effects, and is a highly creative tool as well as a problem solver. The classic trick is to put a more-or-less continuous signal through the gate, such as heavily distorted chords from an electric guitar, and then use the external key to chop it up into a rhythm. Like this:

- Connect the guitar, through a distortion unit, to the gate in the normal way.
- Connect a drum machine, or other rhythmic source synchronized to the track, to the external key input.
- Switch the gate to external key.
- Set the threshold so that the gate triggers on a signal from the external key.
- Adjust the attack, hold and release controls to achieve the desired envelope.

Powerful though MIDI sequencing may be, you can't get the same sound in any other way. This is well worth trying.

Another useful gate effect is to compress the sound of an individual drum, then gate it. This works particularly well on drum samples which have a little bit of reverb on them. The compressor can shape the envelope of the sound by emphasizing the attack (by setting a slow attack time on the compressor, allowing the initial transient to get through unaltered), or by allowing the reverb to increase in level as the drum dies away. The noise gate can then further process the envelope using the attack, hold and release controls.

Gated Reverb

First popularized in the 1980s, gated reverb has become something of a cliche. But as a technique, it is still well worth knowing about. It goes like this:

- Connect the snare drum mic (say) to the mixing console in the normal way.

- Through an auxiliary send, send some of the signal to a reverb unit.

- Bring the output of the reverb back to a channel with an insert point. (If your console's auxiliary returns have insert points, then they will work fine).

- Connect the noise gate to the insert point send and return of the reverb channel.

- Connect the insert send of the snare channel to the gate's external key input. (You could alternatively derive this signal from another auxiliary send). Set the gate to external key (EXT).

- Set the hold and release controls so that the reverb extends beyond the end of the dry snare drum sound, but then dies away suddenly (long hold/short release).

You now have Phil Collins-style gated reverb! You could use a distant mic as the reverb source, as an alternative to the reverb unit.

Check Questions

- Is a noise gate usually effective on a mixed stereo signal?
- Is a noise gate effective on a signal where the instrument plays all the time?
- Is a noise gate effective on a single noisy signal where there gaps in the playing?
- Why is it often considered beneficial to gate some or all of the mics on a drum kit (except the overheads)?
- How many gates are often used in live sound: none, a few or many?
- How is the noise gate connected to the mixing console?
- What is the function of the Threshold control?
- What is the function of the Range control?
- Comment on the Attack, Hold and Release (Decay) controls.
- What would happen if a stereo signal was gated, but the stereo link function was not selected?
- Why are side-chain filters beneficial?
- Comment on the use of an external key to improve the reliability of gating.
- What is an expander?
- Why are compressors and gates often used together?
- Comment on envelope shaping using a noise gate.
- Comment on gated reverb.

Chapter 6: Delay and Reverb

Any enclosed space will cause natural reverberation, where sound bounces off surfaces cause a million reflections, and an infinite number of reflections of reflections. In the open air, a hard flat surface, such as a wall, the side of a building or a cliff-face, will cause a single echo. These phenomena are well-known and understood by both our conscious and unconscious mind.

Since natural echo and reverberation are a common feature of our aural experience, it is inevitable that they will be incorporated into our recording techniques. For some reason, these days we call 'echo' delay. 'Echo' is occasionally, and confusingly, used to mean reverb.

The simplest device commonly used in the past to generate delay, is the trusty Revox B77 (and formerly the A77) analog tape recorder. Because this is a three head machine, there is a gap - and therefore a delay - between the record and the playback heads. The set up is thus:

- Send signal from the channel to which you want to apply delay via a post-fade auxiliary send to the Revox.

- Switch the Revox to monitor off-tape.

- Bring back the output of the Revox to another channel.

Like magic, you now have a single echo on the signal which you can vary in level using either the auxiliary send or the Revox's channel's fader. The latter would be preferable since the noise level could be optimized.

The problem with this is that there isn't a great deal of control over the timing of the delay. At a faster tape speed, the delay time will be shorter. The varispeed control will also allow a small variation in timing. However, before the days of digital delay, people would go to considerable lengths, and it was actually possible to construct an external varispeed control that would allow almost any delay time that could be musically desirable. At longer delay times, the tape speed would be slower and therefore the quality of the delay signal would be poor, but that was considered to be part of the charm of the technique. Even now, analog tape delay sounds subjective better than digital delay.

A variation of this technique was sometimes known as spin echo, or repeat echo. With the setup as above, if you send some signal via the auxiliary send of the Revox channel back to the tape recorder, it will feed back on itself in a loop, causing repeated echoes rather than a single delay. If too much signal is sent back, the level of the echoes will increase until the signal goes into severe distortion. This was actually a much-used technique of the Jamaican dub producers of the 1960s and 1970s.

When digital delay is used to simulate analog tape delay, there is often a control to reduce high frequency content of the delays. This isn't simply a low pass filter on the output, it is in the loop and therefore each repeat has less HF energy. It still doesn't sound the same though.

Particularly in the 1960s, there were a number of tape echo machines available, although these were primarily intended for performance rather than studio use.

Binson Echorec

Calculation of Delay Times

Using an analog tape recorder with varispeed to produce delay, it is impossible to calculate the delay time, so people just set whatever timing they thought was artistically appropriate. Digital delay units however have a numeric readout, and they are also less intuitive to set, therefore calculation of delay time is appropriate. It works like this:

Start with the number 60,000 and divide it by the tempo of the music in beats per minute (BPM).

Divide this number by 2, 3, 4 or 5 - any of these could provide a musically useful delay. Just choose which sounds best.

Delay in PA

Delay is a musically useful technique, but it is also used in public address to counter a real problem. Imagine an outdoor rock concert with an audience of tens of thousands. Obviously, people further back in the audience won't be able to hear a loud enough sound from the speaker stacks on stage (unless they were deafening for the people close to the stage), so additional speakers are commonly mounted on towers to serve the rear section of the audience. These are often called 'delay towers'.

The problem now is that people at the back of the audience hear the sound from the towers, but they can also hear the sound from the stage after a delay of several tens of milliseconds. This is unpleasant, therefore the feed to the amplifiers for the towers is delayed to allow sound from the stage time to catch up.

The delay time can be calculated from measurement of the distance between the stage and the delay towers. Since sound travels at approximately 340 m/s, the delay required equals distance divided by 340 x 1000, giving an answer in milliseconds. An alternative 'rule of thumb' is that sound travels approximately one foot in one millisecond. This is useful if you have been brought up with imperial measurements.

Calculation of the delay time gives and approximate solution, but to be completely accurate the delay has to be fine tuned by ear. One way is to use a clicker device, as sometimes used by dog trainers, to generate an impulse on stage. It is easy to hear when the correct delay time has been set, as the clicks from the stage and towers will merge into one.

The History of Artificial Reverberation

The generation of believable artificial reverberation was seen as an important goal in sound engineering, which some would say hasn't really been achieved properly even now. Analog tape-based units never could approach anything that sounded like reverberation - they just created a sequence of delays.

Natural Echo Chamber

The natural echo chamber is a room with hard irregular surfaces, in which are installed a loudspeaker and two microphones (for stereo). The microphones should not point at the loudspeaker. Signal is sent via an auxiliary send and power amplifier to the loudspeaker, and returned from the microphones to the console where it can be mixed in with the dry signal. The natural echo chamber never sounds realistic, firstly because loudspeakers never sound realistic anyway, and it is generally too small to produce really good-sounding reverberation. Even so, the result is quite unlike any simulation of reverb as it is a much denser sound, and it worth experimenting with even today. It has also been known for classical music recordings to be 'improved' by using a large auditorium as a natural echo chamber. If electrostatic loudspeakers are used then the result can be of excellent quality.

Assisted Resonance

Assisted resonance is a technique used for improving the acoustic characteristics of concert halls. It was first used for the Royal Festival Hall in the UK. This hall, opened in 1951 was found on completion to have too dry an acoustic. Several alternative methods of increasing the reverberation time were considered but found to be impractical or too expensive. The chosen solution of assisted resonance incorporates 220 microphones in resonating cylinders of various sizes placed strategically in the auditorium Each microphone drive its own channel of amplification and loudspeaker. The gain of each channel is set to allow the system to 'ring', but without approaching the point of howlround. Although the technology of the time did not allow such high quality microphones and loudspeakers as we enjoy today, this doesn't matter as each channel operates over only a narrow band of frequencies. This system actually works very well and was employed at the Royal Festival Hall from 1968 to 1998 with near-total success.

Plate Reverb

EMT plate reverb unit

The plate reverb consists of a thin metal plate suspended in a sound proofed enclosure. A transducer similar to the motor of a moving coil loudspeaker drive unit is mounted on the plate to cause it to vibrate. Multiple reflections from the edges of the plates are detected by two (for stereo) microphone-like transducers. Reverb time is varied by a damping pad which can be pressed against the plate thus absorbing its energy more quickly. The plate would typically be mounted outside of the control room, with a remote control for reverberation time.

The sound quality of a good plate is actually very smooth and dense, and is still much in demand for vocals. A top recording studio would still keep an EMT plate in addition to digital reverb units.

Spring Reverb

The spring reverb is an electromechanical device very much like the plate. Instead of a metal plate however it uses long springs in which the signal is allowed to vibrate and reflect. A good spring reverb can give a very good sound on strings, but is hopeless on anything percussive as the springs tend to twang. Spring reverbs are now hardly ever used.

Digital Reverberation

Obviously now digital reverberation is in almost universal use, and it can produce an almost totally believably natural sound, or it can be used to go beyond naturalness and creatively enhance a sound. The main parameters of digital reverb are these:

Reverberation time. As in acoustics, this is defined as the time it takes for the reverberation to decay by 60 dB.

Early reflections. As in acoustics again, the earliest reflections are perceived separately. The character of these early reflections tell our ears a lot about the space we are in, and digital reverberation units will various combinations to simulate different acoustic environments.

Pre-delay. It is often useful to separate reverberation in time from the original signal. Particularly with vocals, this allows a higher level of reverb while maintaining intelligibility.

High frequency damping. It is common for real rooms with soft surfaces coverings to have a shorter reverberation time at HF than at LF. Digital reverb units simulate this.

Although any good digital reverberation unit will offer these parameters and many more, it is the overall character of the unit that is the most significant factor. The most well-respected manufacturer of digital reverb units is Lexicon, whose 480L and 224XL are 'must have' equipment for any top recording studio.

Check Questions

- Describe the signal flow in a three-head analog tape recorder used to produce a single delay.
- In analog tape delay, using a tape recorder, how is the delay time controlled?
- What is 'spin echo'?
- How was the technique of spin echo used by Jamaican dub producers of the 1960s and 1970s?
- At a tempo of 120 BPM, what delay times are most likely to produce musically useful results? Give four answers.
- What are 'delay towers'?
- If you use imperial measurements, what is the rule of thumb for calculating delay time for delay towers? (If you are not familiar with imperial measurements, you need not answer this question).
- How can the delay time for delay towers be fine tuned?
- What is a natural echo chamber?
- Briefly, what is assisted resonance?
- Briefly, what is a plate reverb?
- List four important parameters that may be controlled in a digital reverberation unit.

Chapter 7: Disk Recording

Tape recording had a long history of success from the 1940s to the 1990s. Even though tape still survives - analog tape as a medium with its own sonic character, DTRS and DASH machines as bulk storage devices - it is clear that in almost every way imaginable, disk recording is superior and will eventually become universal. (It's worth pointing out that digital video tape will be around very much longer simply because of the requirement to store such a great bulk of data).

Disk recorders come in two types, one is as a complete system, the other is as a software and peripherals to operate in conjunction with a standard personal computer. Complete systems have certain advantages:

- The entire system is a known quantity to the manufacturer, who can guarantee performance

- If there is a problem, you just go back to the manufacturer for the solution.

PC systems have the advantage that they can be much cheaper because they use standard computer industry hardware and operating systems, but they have the distinct disadvantage that if you come across a problem, who do you turn to for help?

- The audio software designer?

- The audio hardware manufacturer?

- The computer manufacturer?

- The disk drive manufacturer?

- The operating system designer?

- The hard disk driver software designer?

As likely as not, they will all blame each other. Professional users like PC-based systems for their cheapness, but they will protect themselves by installing and configuring a system with the assistance of a dealer who has a good knowledge of all the components used. Once the system is working and any problems ironed out, they will just leave it as it is and resist any temptation to upgrade, for example, to a newer operating system. This way, they have a system that can work day in day out until it finally breaks down, or is superseded by a more efficient system.

Home recordists have a tendency to bring problems upon themselves in a variety of ways:

- Installing and using software other than music and audio software

- Regularly upgrading the operating system to a new set of bugs

- Installing games and 'utilities' from magazine cover CDs

- Tinkering with the system settings.

Any of the above actions could get you fired from a commercial studio.

It is worth, at this point, considering some of the available options.

- AMS-Neve Audiofile
- DAR SoundStation
- Euphonix R-1
- Fairlight Prodigy
- IZ RADAR
- Digidesign Pro Tools

All of the above would be considered professional systems, and of course there are others. It's worth pointing out that systems such as Logic Audio, Digital Performer and Cubase VST are considered by audio professionals to be musicians' playthings. Perhaps that is just an opinion, but it is worth considering the motto, "Think Professional. Act Professional. Be Professional". All of the above manufacturers serve the pro audio market, while the audio sequencer manufacturers sell mainly into the bedroom studio market.

Of the above, some - like the AMS-Neve Audiofile - are considered to be recording and editing systems and their main use would be in audio post production. The Euphonix R-1 however is marketed as a replacement for the traditional multitrack recorder. Of course it has editing facilities, but recording is its main function. The IZ RADAR, and other similar machines, are 'drop in' replacements for the old 24-track recorders. Digidesign Pro Tools is far and away the most common professional system and, in the music recording studio, is mainly seen as an editing accessory rather than the primary recorder, although this may change. Pro Tools (since Digidesign is owned by Avid) is also widely used in audio post production for film and video.

Inputs, Tracks, Outputs

The way a hard disk recorder interfaces with the rest of the studio is different to a tape recorder, analog or digital. All tape recorders have the same number of inputs, tracks and outputs. So for example, a 24-track recorder always has 24 inputs, 24 tracks and 24 outputs. There wouldn't be any point in any other arrangement.

The nature of tape demands that each track be given its own area across the width of the tape. So a tape recorder always has a fixed number of tracks. There are no such constraints in disk recording. Data is recorded on the disk in three dimensions, and not necessarily contiguously or in sequence. The number of tracks of which a system may be capable depends on how fast the disk can access data, and how fast the rest of the system can process it. The speed of the disk is the main limitation, and when the data is fragmented (more on this later), the access times slow. The result is that the number of tracks a hard disk system might be capable of is not a constant. Some manufacturers have made it appear so - so that the IZ RADAR for example is presented as a 24-track recorder - but this is probably just to help ease the transition from tape for Luddite engineers.

Once the number of tracks ceases to be a constant, the question of how many inputs and how many outputs is thrown wide open. The fewest could be two inputs and two outputs. If the disk is fast enough, this could be a 24-track system, or even more, but you wouldn't be able to record more than two tracks at once, nor take any outputs other than the main stereo output. In this case, using an external reverb unit via an aux send and return would not be possible.

Having eight inputs and eight outputs is a good compromise and allows several tracks to be recorded simultaneously, and also the use of outboard effects. You couldn't however use such a system as a replacement for a traditional 24-track recorder with a normal mixing console. For that you need a full complement of inputs and outputs.

Having fewer inputs and outputs than tracks means that an input has to be assigned to the track you want to record on - a procedure that isn't necessary with a tape recorder. Also, tracks and auxiliaries have to be assigned to outputs.

Session Files

With a conventional tape recorder, you would start a session by taking out a new reel of tape and loading it on the machine. There has to be an equivalent to this in disk recording.

With removable hard disk cartridges available, the process could be very similar. Take out a new hard disk and connect it to the system. But within this there is also the process of creating a new 'session file'. Manufacturers tend to have their own terminology, but it's reasonable to use Digidesign's terms since Pro Tools is so common. The session file contains a lot of information about the project:

- Project title
- Audio files used
- Edits
- Track assignments
- Mix and plug-in data

Typically each new song would be given its own session file. One difference between the session file and a reel of tape is that the session is not so easily transportable. For instance, a session might be recorded and edited on a Pro Tools system with 16 inputs and outputs and mixed within Pro Tools using a certain combination of plug-in effects. If you try and play that back on another Pro Tools system with only 8 outputs, and with a different combination of plug-ins, then it will take some effort to get the session sounding similar to the way it did originally. With tape, you know that you can record on one machine and play back on another without any problems, assuming the machines are properly aligned, and you don't expect the tape to contain mix data, so it is a completely natural process simply to reconstruct a rough mix.

Also, session files are highly incompatible between manufacturers. There are moves towards compatibility, or at least common transfer processes, but you can't just take a disk recorded on one system and expect to get any sense out of it at all on another.

Disk Media

The choice of disk media isn't as tricky as it once was. There used to be the problem that hard disks need to recalibrate themselves every so often to account for the heating and expansion of the disk platters but the manufacturers have now mostly found workarounds to this. It is still necessary for a disk to achieve a certain level of performance, and the rotational speed of the platters should ideally be at least 7400 revolutions per minute, preferable 10,000 rpm. Most modern disks will be capable of simultaneous record and playback of at least 24 tracks. Note that the very small disks used in laptop computers may have a rotational speed of only 4800 rpm so either the track count will be low, or the recording software may recognize that the disk is not qualified and refuse to use it.

The type of interface used is still significant. The most common disk interface is ATA, formerly known as IDE or EIDE. ATA disks are cheap, but they are not the fastest. The ATA interface does not intrinsically lend itself to having disks that are external to the computer so expansion of storage capacity is limited. Some systems allow the ATA disk to be installed in a removable cartridge, which does come close to the convenience of tape.

SCSI is commonly regarded as an old interface on the verge of extinction. This is not so. SCSI does allow disks to be external to the computer (the host), and disks can be daisy-chained on a bus so that several can be connected at the same time. SCSI also has the advantage that, being a parallel interface, data travels along several wires rather than the single data conductor of Firewire or USB and can therefore be faster than either. It is likely that SCSI will continue to be developed for the foreseeable future.

Firewire, or IEEE 1394 (or iLink as Sony call it) is a fast serial interface, meaning that data travels along only one conductor in the Firewire cable. Firewire is fast, if not as fast as modern versions of SCSI, and has a number of features that SCSI lacks:

- Self-configuring - no ID numbers to be set by the user.
- Hot swap - devices can be connected and disconnected without powering down the system.
- Bus power - the cable can transfer power from the host to the peripheral device, although this might not be sufficient for all devices.
- Thin, flexible cable and robust connectors.
- Target mode - where the internal disk of a computer can be connected to another computer as though it were just an external disk.

There is in fact no such thing as a Firewire disk. Firewire is simply a means of data transfer whereas ATA and SCSI intelligently determine the manner in which the data is allocated to the disk's platters. An external Firewire disk is therefore simply an ATA disk in an enclosure together with a 'Firewire bridge'.

USB is currently not of relevance as a disk interface for audio recording as it is too slow, and faster versions must compete with the longer established Firewire.

It is better to dedicate a separate disk to audio rather than allow audio and software to share disk space. The advantage is that the disk can periodically be erased completely so that excessive fragmentation does not occur. Fragmentation is where the disk is unable to allocate contiguous disk space and occurs when the disk is nearly full or has been used over a long period of time without being completely erased. Discontiguous recording forces the disk mechanism to work harder and the system will be capable of fewer simultaneous tracks under this condition.

Storage and Archiving

Until recently, disk storage was too expensive to be able to store or archive a project on the original media. Disks are now cheaper than 2-inch analog tape so this is no longer a restriction. However, the archival life of a recording on disk is currently not adequately proven. Recordings can be archived to tape media, typically Exabyte, or to CD-ROM. Optical storage is felt to provide the best assurance of archival stability. The only problem with CD-ROM is the small amount of data (<700 Megabytes) that each disk can hold. If an archiving utility such as Retrospect is used, then successful restoring of the archive depends on the availability of that software, and compatibility with current systems. It would be doubtful whether this would be so in thirty years time, and the relevance and marketability of a recording could easily last a much longer time than that.

Often when recording to a disk system, there ends up being a lot more audio on the disk than is actually used. For instance, inadequate takes that would have been erased on tape are automatically kept by the disk recording system. Systems often allow unused takes to be selected, then deleted. However, there will also be unused audio remaining that consists of the trimmings from regions of audio that have been 'topped and tailed'. This can amount to a considerable quantity of data, so the system may offer a 'compact' function where the superfluous material is deleted. If so, there should be the provision to retain 'handles'. Handles are sections of audio before the start and after the end of each region that are not actually displayed on the screen and do not sound. These can be useful if it proves necessary to adjust an edit, or to create crossfades. A handle duration of 2 seconds is reasonable for most purposes.

Operation

The process of making a hard disk recording of a conventional rock band on a system such as Pro Tools would go something like this:

- Create a new session file.
- Create the required number of tracks for the basic tracks of the song.
- Assign inputs to tracks.
- Assign tracks to outputs.
- Either assign each track to its own output, if enough output channels are available, and operate as a conventional multitrack recorder. Or...
- Assign all tracks to the main stereo outputs and mix internally.
- Record the basic tracks, doing several takes until the band has produced its best performance.
- Edit the best sections of the takes together, cutting across all tracks.
- Fix small problems - for example a wrong note in the bass guitar - by editing from the appropriate track on an alternative take.
- Perform overdubs.
- Typically, up to six takes of the lead vocal would be recorded on separate tracks. They can then be compiled, or 'comped' together to produce one near-perfect version.
- The singer can punch-in any remaining problem lines. The hard disk recording system will make it easy to set punch-in and punch-out points, do as many takes as necessary, then select the best one.
- Mixing can be via a conventional mixing console, if the hard disk system has sufficient outputs, or mixed internally within the system.

The above isn't that different from conventional multitrack tape recording. The advantages are that there is no rewind time, and editing is much more flexible. If mixing is done internally, then the entire project can be stored in the session file and recalled at any later date. It is a very time consuming procedure, with a conventional mixing console, to store settings and reset the console at a later date.

Hard disk recording systems can seem complex because they offer so many options. But it's worth remembering that everything should be in the service of music, or whatever audio application. There are systems that are so complex that they obscure what should be a clear path towards a successful recording. Here are a few selected features of hard disk operations:

Reliability of Recording

There is often a greater question mark over the progress of a disk recording than a tape recording. An engineer recording a live event on tape can be almost guaranteed that the recording will be made successfully. There is always a doubt with a disk recording system that the recording will suddenly cease, accompanied by an error message, or that when the recording is stopped it will be found that there is no audio on the disk. This will change in time.

Virtual Tracks

All tape recorders, as said earlier, have the same number of inputs, tracks and outputs. Disk recorders have no such fixed relationship. Further than that, beyond the limit of the number of tracks that are replayable simultaneously, many disk recorders can store additional so-called 'virtual' tracks that are in synchronization with the regular tracks, but cannot be allocated to an output. Real tracks and virtual tracks can easily be exchanged. As systems become capable of replaying more and more tracks simultaneously, the relevance of virtual tracks will decrease.

Editing

Whereas multitrack tape recorders only deal with continuous tracks, a track recorded on a hard disk system can be split into a number of regions. Typically, this would be to separate out useful audio, and junk audio that should not be heard. When the useful regions are separated out, the junk audio remains on the disk, but is silenced. This can be extended to audio that might be useful; you can silence a segment of audio that you don't think you will need, and then retrieve it if you realize that you do. Trimming the start and end points of regions (topping and tailing) is a common activity.

When trimming regions, it would be possible to cut at such a point in the waveform that a click would be produced on entry to or exit from the region. Often a short fade will be generated automatically to eliminate the click. In some systems however this needs to be selected as an option.

When an audio file is recorded, subject to the whims of the SCSI or ATA controller, it is recorded in long contiguous sections on the disk. The disk does not have to work hard to retrieve the data therefore many tracks can be replayed simultaneously. When material is edited, then the regions that are to be replayed are probably not contiguous on the disk, so the heads have to move more often to retrieve the data. Ultimately, there can come a point where this affects the number of simultaneous tracks. The term 'edit density' is used to mean the number of regions, their length and proximity to each other (shorter, more densely packed, regions make the disk work harder) and how far they are physically separated on the disk. The greater the edit density, the fewer tracks will play. If high edit density is causing a problem, two or more disks should be used so that tracks can be shared among them.

Slipping and Spotting

The following represents Pro Tools terminology. Other systems often have equivalent functions:

Slip: Since there is no fixed time relationship between tracks, any track or region can be 'slipped' with respect to the others. If it is known that a region (or whole track if no separate regions have been defined) is in the right place, usually it can be locked so that it can't be slipped accidentally.

Spot: When a region is recorded, on some systems, it is given a time stamp. Even if it is subsequently moved, it can always be 'spotted' back to that position, or to another position defined by a numerical time reference.

Grid: It is usually possible to define a grid, so that regions will always 'snap' to grid positions when moved.

Shuffle: It is often desirable to be able to move a region so that its start point exactly coincides with the end point of the previous region, or to be able to insert a region quickly between two others and move them apart by the precise duration of the inserted region. This can be done in Slip mode but it takes time. Shuffle mode makes it almost instantaneous.

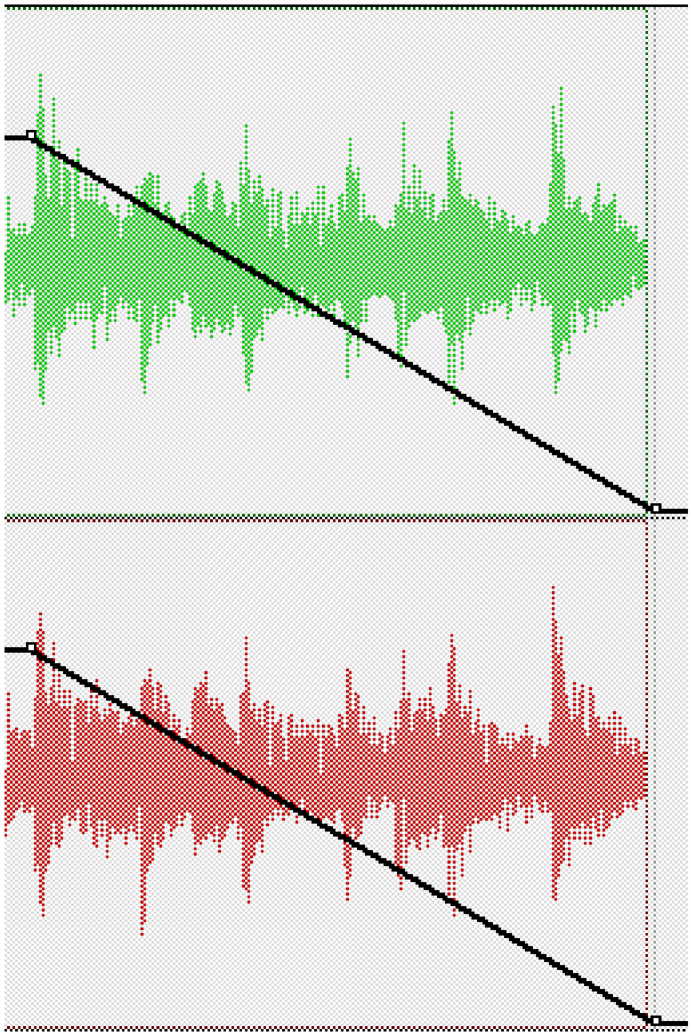
Fades and Crossfades

A region may have a fade-in and a fade-out of any duration. Normally it is beneficial to have a short fade-in and fade-out of around 5-10 milliseconds to ensure that there are no clicks. Some systems do this automatically, others provide automatic short fade-in and fade-out as an option.

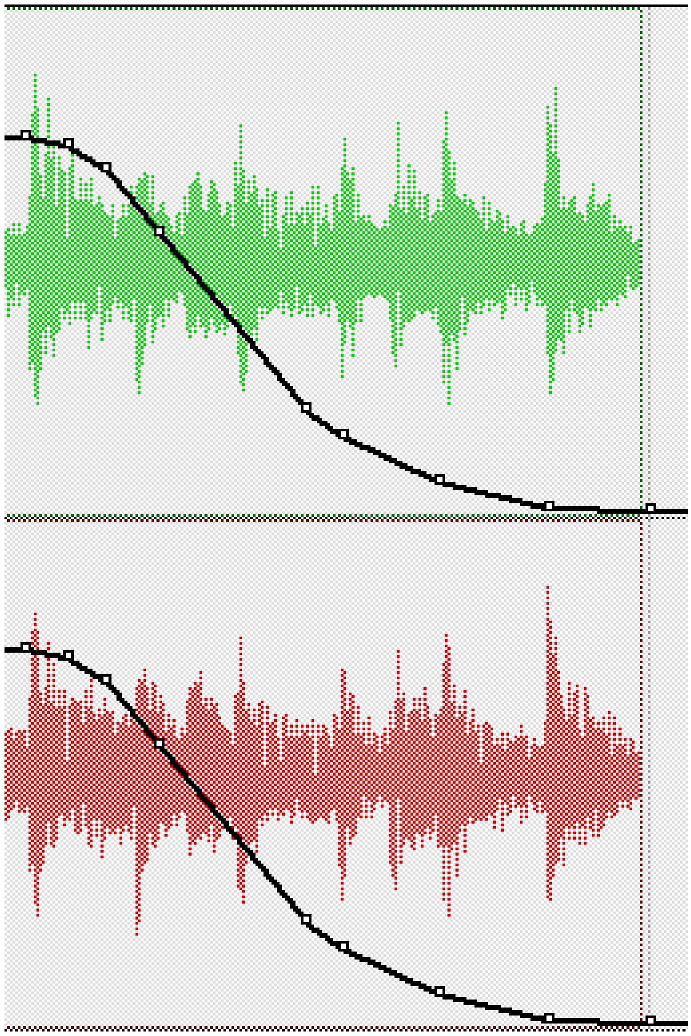
A fade can be given a profile: it can fade out as a straight line, reducing by a consistent number of decibels throughout the duration of the fade. Or it can start to fade slowly and end up fading quickly, or start to fade quickly then end up fading slowly. An 's' shaped profile often works well, where the fade starts slowly but then gets quicker, slowing down again before fading out completely.

When two regions are butted together, a crossfade may be created. It is important to remember here that there has to be audio on the disk that extends beyond the region boundaries. It is this material that is used to create the crossfade. If there is no material beyond the region boundaries, or insufficient for the desired duration of the crossfade, then the crossfade cannot be created.

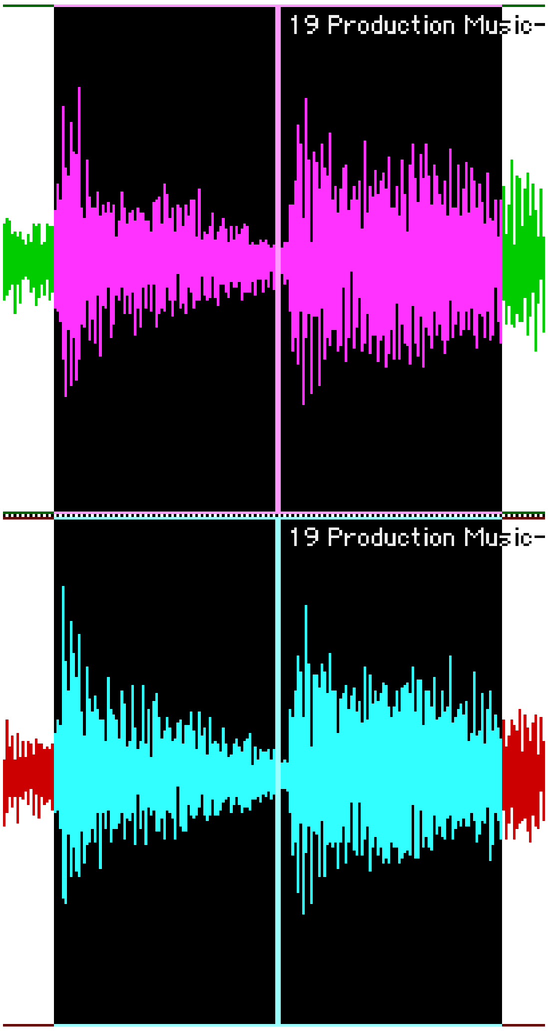
Hard disk systems perform crossfades in one of two ways - either by mixing the two sections of audio in real time, or by calculating a separate fade file. The advantage of calculating in real time is that creation of the crossfade is instant. Calculating a separate file takes a little time, but as processor speeds increase this becomes less relevant. The disadvantage of calculating in real time is that you might set a crossfade over twenty-four tracks simultaneously. The system would then have to play back forty-eight tracks simultaneously for the duration of the fade. This might not be possible.



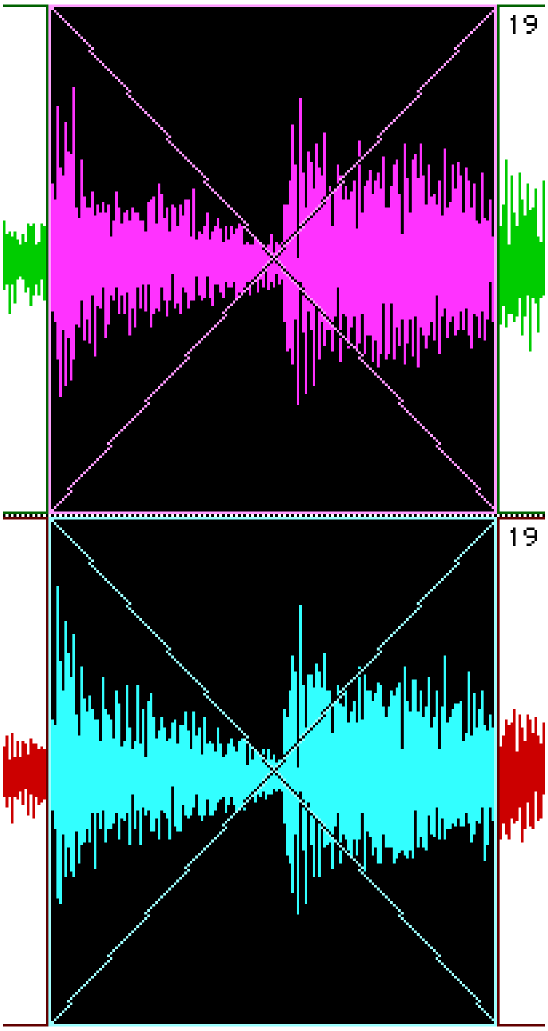
Linear fade



S-curve fade profile



Selection for crossfade - audio from the outgoing region must extend beyond the end of the region. Audio from the incoming region must extend before the start of the region, otherwise there is no material with which to create the crossfade.



Crossfade

Bouncing

The term 'bouncing' in multitrack recording normally means mixing together similar tracks, such as backing vocals, onto fewer tracks to free up the original tracks for further recording. However, in a disk system, bouncing can mean conventional bouncing, or it can mean mixing. There no need to record a mix to a separate medium such as DAT, it can simply be bounced to disk as a stereo file. In some systems, bouncing can take place faster than real time, although arguably real time bouncing is better for the proper functioning of mix automation and plug-in effects, and real time bouncing offers the chance of listening to the mix, just to make sure that it is correct.

If it is a conventional bounce that is being performed rather than a mix, it is often convenient to import the bounce back into the session. This is often available as an automatic feature.

Library

Disk recording systems have a store of audio material categorized in one of two ways:

File: The complete audio file on disk, either mono, interleaved stereo, or two separate mono files sharing the same name with suffix (typically) .L and .R

Region: A file, as above, may have a number of regions already marked, which may or may not appear in a playlist.

Playlist

Sometimes this refers to the entire editing screen, other times to a single track to which a playlist can be allocated. In the latter case the playlist is a list of regions including the timings at which they should play.

Plug-Ins and Latency

Plug-ins are additional pieces of software, sometimes from third-party manufacturers, for functions such as EQ, dynamics and effects. Every function that a disk recording system performs requires processing power - the CPU (central processing unit, or simply 'processor') of the computer has to work quickly to achieve what is required in real time. Processors are fast enough to replay many tracks and mix them with ease. EQ is also a relatively simple task, as is dynamic processing such as compression and gating. Other functions such as reverb take many calculations, and there is a limit to how much the processor can do in real time. 'Host based processing' is where the computer's own processor performs these calculations. In use, you would find that a number of plug-ins can be inserted, but there comes a point where the system can't perform, and (hopefully) it informs you that you have asked too much and a plug-in needs to be removed. There is only one way to increase the performance of a host-based system, other than replacing the computer with a faster one. That is to set up a larger 'buffer'. You can think of the buffer as a short time period during which the processor can perform its calculations

Since the demands on the processor will vary from moment to moment, it gives the processor the opportunity to 'catch up' if it had been struggling for a moment. On record or playback, the size of the buffer is of no consequence - you just have to wait a little longer between pressing play and hearing audio. When you are overdubbing however, the size of the buffer is critical. Anything over 10 milliseconds is clearly audible, and distracting for the performer. Manufacturers often call this 'latency'. Latency occurs in the sound card, in the software itself, and in the buffer, and without doubt it is a 'bad thing'.

The alternative to host-based processing is DSP, or digital signal processing. Audio signals are different in structure from standard computer data and benefit from having DSP chips specially designed for the purpose. Digidesign's Pro Tools, for example, uses DSP cards to perform mixing and plug-in effects. The latency of a DSP-based system is very low, of the order of 3 to 4 milliseconds, which for most purposes is unnoticeable. The cost of a DSP-based system may be considerable, but the benefits of eliminating problems due to latency are great.

One point to consider when using plug-ins is that each plug-in delays the signal slightly. Suppose one had recorded an acoustic guitar in stereo, and decided that one of the mics needed EQ, which might easily happen. If an EQ plug-in was inserted into only one channel, there would be a noticeable delay between the channels. The solution would be to insert EQ into both channels, even if only one was to be used. This would be of particular significance with acoustic drum tracks. It is common to EQ, compress and gate drums. It is necessary to insert the same plug-ins into every drum channel, even if they are not all going to be used. Some systems offer specific time-adjusting plug-ins which consume less DSP or processing power.

Check Questions

- What is the advantage of hard disk recording systems that do not use a standard personal computer over those that do?
- What is the advantage of PC hard disk recording systems over those that do not use a standard PC?
- How do home computer users often take actions that are to the detriment of reliable disk recording?
- Comment on the difference between recording/editing disk systems and multitrack tape replacement systems.
- What is the minimum configuration of inputs and outputs for hard disk recording?
- What would be a reasonable configuration of inputs and outputs for a hard disk system, where mixing will be internal but outboard effects units are used?
- What configuration of inputs and outputs would be necessary for a system that was intended to replace a multitrack tape recorder in a conventional studio?
- What information is stored in a session file?
- Does the session file contain audio data?
- Are session files commonly transportable between different systems?
- Comment briefly on ATA disks.
- Comment briefly on SCSI disks.
- Comment briefly on Firewire disks.
- What is the problem in archiving recordings made on disk systems?

- Regarding the section titled 'Operation', what differences in procedure are there in disk recording compared to multitrack tape recording?

- What is a 'virtual track'?

- Describe the two methods a disk recording system may use to create a crossfade.

- What are 'handles'?

- What is 'latency'?

- Comment on the delay produced by inserting a plug-in into the signal path.

Chapter 8: CD and DVD-Video

CD

Since 1983, CD has been the medium of choice for home audio entertainment, and is also suitable for in-car and personal use. On it's introduction, it was by no means certain that the market would take to it. Both discs and players were expensive, and vinyl had been entrenched since the mid-1950s. However, particularly for classical music, its advantages soon came to the public's attention:

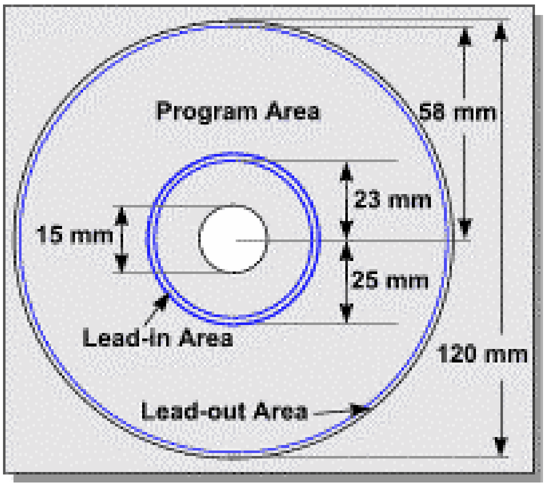
- No clicks due to dust and scratches
- No wear
- Long playing time (no turn over necessary)
- Compact size.

The initial marketing phrase 'Pure, perfect sound forever' may not have been the entire truth, but in comparison to vinyl it was close enough. CD does have some problems:

- Muting and sometimes glitching due to fingerprints (although the disc can be cleaned)
- Skipping in portable players
- Severe abuse can damage the disc - it isn't indestructible
- The small size limits the impact of the artwork.

Even so, CD is a very solid and reliable medium that does credit to its co-inventors, Philips and Sony.

Overview



Playing time

The maximum playing time of a CD was, apparently, decided after consultation with conductor Herbert von Karajan, who though that a single disc should be able to contain all of Beethoven's Ninth Symphony (evidently at Karajan's tempo, not Otto Klemperer's). Hence the standard maximum playing time was set at 74 minutes, 33 seconds. A longer duration can be achieved, up to around 80 minutes, but this is at the expense of increased risk of rejects in the factory and faulty playback in some players. A factory would be entitled to specify that production would be at the customer's risk if the duration were more than the standard maximum.

Data storage

Data is stored as a sequence of pits impressed into a plastic (normally polycarbonate) substrate, which is coated with reflective aluminum to be read by a laser. In a typical disc, there are around two billion pits, each of the order of half a micron in size, set in a spiral track that reads from the center outwards. The width of a human hair would cover about thirty turns of the spiral and its total length is almost five kilometers.

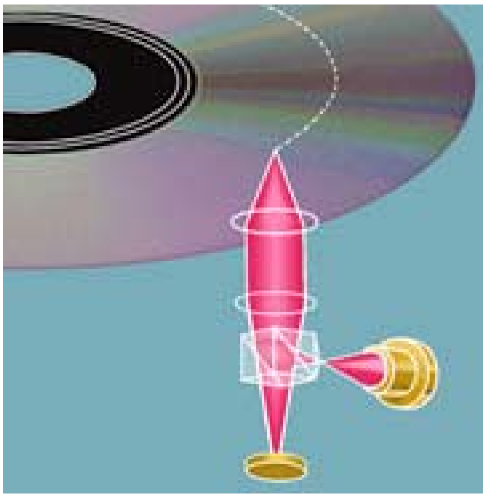
Although the playing surface of the disc is covered by tough polycarbonate, the label side is coated only in a thin lacquer. A scratch on the playing surface could be polished out, but a scratch on the label side could easily damage the data layer.

Binary data is encoded into the transitions from 'land' to pit. Each transition represents a binary '1'. If there is no transition, where there could have been one, a binary '0' is encoded.

Rotation

The disc rotates at a constant linear velocity, not a constant angular velocity as does vinyl. Hence the disc rotates quickly (about 8 revolutions per second) at the start of the audio that is near the center, reducing to about 3.5 revolutions per second when the laser is tracking close to the disc's edge. The linear speed is between 1.2 to 1.4 m/s depending on the duration of the program.

Pickup



The disc is read by a laser beam, which reflects from the spiral of pits. Since there is no physical contact, there is no wear. The beam is focussed on the pits, which are below the actual surface of the disc. Any dust or scratches on the surface are therefore out of focus and tend not to affect the beam. Additionally, the beam is split into three by a diffraction grating so that additional signals are used to facilitate tracking, the player being able to tell if the playback beam is wandering to either side of the desired path. Some players use three separate beams. There is also an auto-focus mechanism to compensate for the inevitable lack of perfect flatness of the disc.

Coding

Data is coded from pure binary into a form that is suitable for the compact disc medium. Coding ensures that the is always a sufficient quantity of pits to ensure reliable tracking (otherwise a data stream consisting of many consecutive zeroes would lead to a complete absence of pits), and to allow for error detection, correction and concealment. The coding method used also facilitates recovery of a clock signal from the data, which is used to control the speed of the motor.

Subcode

In addition to the audio data, there is room in the bitstream for eight channels of subcode, known as P, Q, R, S, T, U, V and W. Only the P and Q subcodes have achieved widespread usage.

Data is grouped in frames, each containing 192 bits of audio, plus other data such as sync bits and error correction data. In each frame there is one subcode bit for each channel. The P channel designates the presence or absence of audio, as follows:

- A sequence of 0s designates the presence of audio
- 1 designates the start of a track, and also pauses between tracks
- A sequence of 0s before the first track designates the lead-in area
- Alternating 1s and 0s after the last track designate the lead-out area.

The P channel was designed for budget players to be able to find the tracks easily. Many players now do not use the P channel.

The Q subcode channel provides information as follows:

- Whether the disc is 2-channel or 4-channel (4-channel has hardly ever been used. There is no 1-channel mode, precluding the production of double duration mono discs of archive material)
- Pre-emphasis on or off. (Pre-emphasis is the technique of boosting high frequencies on record, and cutting them back on replay, thus reducing high frequency noise. It is now seldom used on CD,

although it remains in use in analog tape recording, FM radio and television sound)

- Copy prohibit on or off
- Disc contents information such as number of tracks and their start times. In the lead-in area, this becomes a specific table of contents (TOC)
- Catalog data such as Universal Product Code (UPC)
- International Standard Recording Code (ISRC).

Audio specifications

Sampling rate: 44.1 kHz

Resolution: 16-bit

Frequency response: typically 5 Hz to 20 kHz ±0.2 dB

Signal to noise ratio: typically better than 92 dB

Distortion: typically better than 0.01%

[CD graphics courtesy Disctronics]

DVD-Video

[Adapted from an article by David Mellor that first appeared in Audio Media]

DVD-Video arrived first in Japan in November 1996, but some time before that the Motion Picture Studio Advisory Committee had been searching for a disc-based video carrier to replace VHS (which they realized, after some hesitation, was earning significant amounts of money from rental and sell-through). Of course video discs were already available in the form of Laserdisc but, for reasons that are not easy to understand, Laserdisc never captured mass market attention and languished as a specialized hobby interest only. Still, Laserdisc provided a benchmark, as did the Video CD format that was already available (and is still surprisingly popular in Eastern Asia). The committee decided that the new format would have to fulfil the following requirements:

- 135 minute duration on a single side
- Better video quality than Laserdisc
- Surround sound, subjectively as good as CD
- Multiple languages, up to five
- Multiple subtitles, up to six
- Pan & scan, letterbox and widescreen formats
- Parental lockout option
- Copy protection
- Ability to play CDs
- Chapter division and access
- Similar manufacturing costs to current CDs

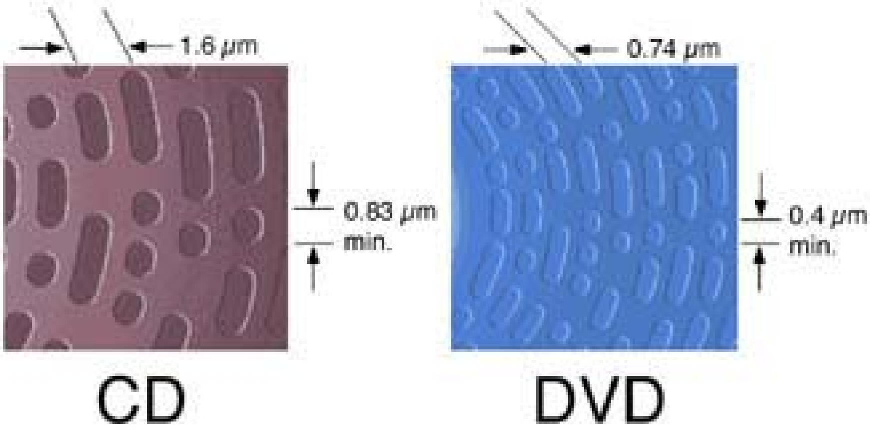
This, as it stands, seems surprisingly far-sighted. An enhancement to Video CD was considered but in reality it never would have provided sufficient data capacity to fulfil even the first two criteria, regardless of the rest. Hence a new disc format was required which came – after some tussle between rival camps of manufacturers - to be established and known as Digital Video Disc, then Digital Versatile Disc and now simply DVD. In fact DVD-Video, as we call the video application of the physical DVD format, provides more than the criteria set out by the committee:

- Up to eight languages
- Up to thirty-two sets of subtitles

- Interactive menus and program chains
- Multiple camera angles, up to nine
- Copy protection in both digital and analogue domains

Table 1	Capacity (Gigabytes)	Layers	Sides
DVD-5	4.7	1	1
DVD-9	8.54	2	1
DVD-10	9.4	1	2
DVD-18	17.08	2	2

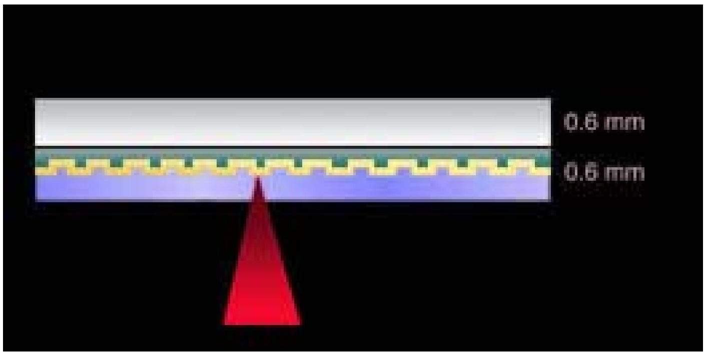
DVD-Video is based on the DVD-ROM specification with the UDF (Universal Disc Format) file system. Physically the disc is available in a number of configurations from DVD-5 up to DVD-18, as in Table 1. The smallest capacity of 4.7 Gbytes is around seven times the capacity of a CD and is sufficient to store a full 133 minutes of video, plus all the other data as mentioned above. There are certain obvious differences between a CD and a DVD. Firstly, although similar technology is used, the data is packed at around four times the density. This is made possible mainly thanks to improvements in laser technology. Oddly enough, when you consider that Laserdisc is an analogue format, the tolerances to which the laser must operate are arguably tighter than DVD. We certainly haven't reached the limit of data density yet, and I am sure we will find uses for any future increases in capacity.



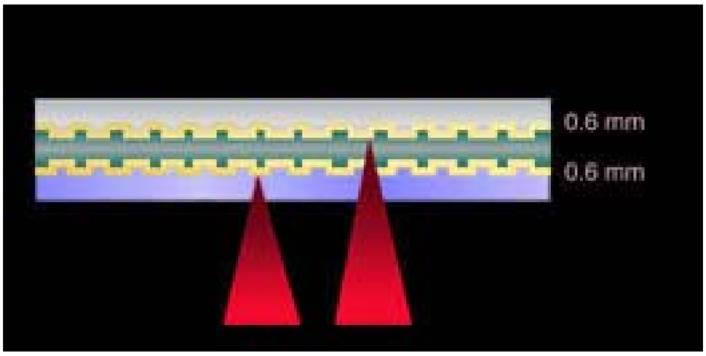
Another difference between DVD-5 and CD is that a CD is made on a single polycarbonate substrate. From the data side the disc is extremely tough, and even if it is scratched, scratches can often be polished out successfully. However the label side is protected only by a thin lacquer coat. Any damage to this side is quite likely to affect the aluminium data layer irretrievably. All DVD discs are made from two substrates (0.6mm each, compared to CD's 1.2mm, maintaining the same overall thickness). Therefore the data is locked very securely inside the disc and the actual data layer is less likely to be damaged.

DVD-9 advances technology by bonding one substrate with a reflective layer to another with a semi-transparent layer. The laser can be focussed on either layer so that nearly twice the playing time can be achieved. Why would this be necessary when one disc can already hold a complete movie? The answer is, as we shall see in more detail later, that the data rate can be varied to give varying degrees of subjective picture quality. If more capacity is available, then the data rate can be higher. Also of course, the extra capacity can be used to add value to the disc. DVD-Video producers are aware that ‘The Making of…’ style documentaries and interviews with key members of the creative team can add useful perceived value, in the mind of the potential purchaser, at very little cost compared to the movie's overall budget. And then there are Easter eggs too, but if told you about those, I would be spoiling the fun!

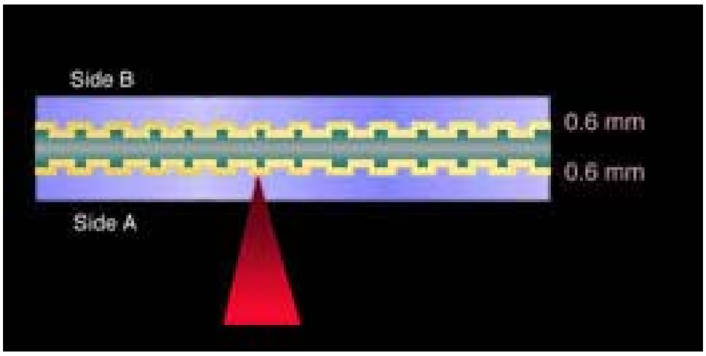
DVD-10 precisely doubles the capacity of DVD-5 by putting data on two substrates, but making it available from both sides creating a so-called ‘flipper’ disc. The value of a flipper disc can be questioned as it is just as easy for the user to take the trip to the DVD player and insert a second disc as it is to flip just one. Adding a second disc into the pack once again increases perceived value. Also the label space on a flipper disc is necessarily quite tiny.



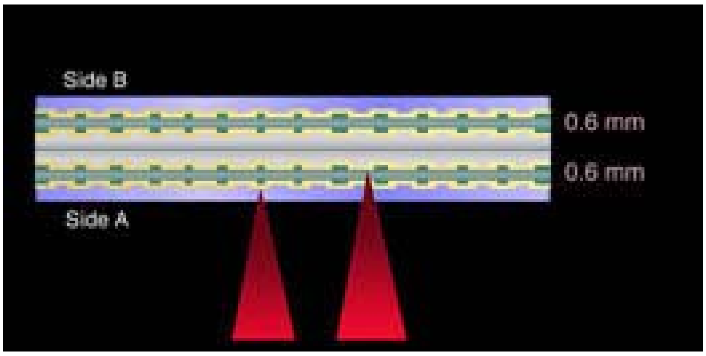
DVD-5 single-side, single-layer



DVD-9 single-side, dual-layer



DVD-10 dual-side, single-layer



DVD-18 dual-side, dual-layer

DVD-18 still hasn't quite made it in the way that was hoped. With an enormous capacity of over 17 Gigabytes, a DVD-18 disc could hold an entire movie at DVD-Video's maximum data rate and still have spare capacity for extras. The problem is that they are very difficult to manufacture successfully and yields (the proportion of discs that come through manufacture without defects) are low, making them expensive. Since they don't look any different to the purchaser there is little perceived extra value. And of course a DVD-18 disc is a flipper! DVD-18 discs will remain a comparative rarity until yields improve and perhaps even beyond.

Video Encoding

It is very well known that DVD-Video uses MPEG-2 as its video data encoding system. It isn’t quite so well known that it can use MPEG-1 as well, although you would have to speculate why anyone would want to, apart from achieving very long playing times. There are in fact a number of options, as detailed in Table 2. Of course it is the higher resolution images that are of main importance. Presumably the lower resolutions were made available to cater for any possible need for compatibility. MPEG-2 is of course capable of video good quality results, but the degree of quality depends very much on the bit rate. The maximum data rate of which a DVD-Video disc is capable is 10.08 Mbit/s to include all data streams. The video stream itself could have a maximum data rate of 9.8 Mbit/s. This doesn't seem to leave much room for audio, but if you do the math you will find that the data rate is comparable to other data-reduced audio that many people listen to quite happily. Audio, by the way, is generally Dolby Digital. Uncompressed PCM is an option, but of course eats up the bandwidth. MPEG audio is also a possibility on PAL discs. DTS is an occasional option for home cinema enthusiasts. It is of course possible to take away some video bandwidth and give it to the audio side of things, which would be nice in any case and of course is to be preferred for 5.1 channel audio output. The maximum overall bit rate of 10.08 Mbit/s isn’t achieved in practice except on material of short duration. If indeed the material is of short duration then constant bit rate (CBR) coding can be used where whatever bit rate is chosen stays fixed from beginning to end. On longer material though, choices have to be made in the form of variable bit rate (VBR) coding. Of course, capable software can decide for itself how to vary the bit rate to optimize the image at all times and will indeed examine the material twice or three times to work out the best compromises and decide which sections need to have maximum image quality and which can get by with a sprinkling of MPEG-2 low bit rate artefacts. One paradox is that images that need maximum image quality, such as close ups of the human face may, if the actor is young and smooth-skinned, encode well at a lower bit rate. On the other hand, scenes such as moving foliage (caused by the wind, or movement of the camera) are not that important to be seen in such fine detail, but the demands made on the MPEG-2 encoder mean that any artefacts that are generated will readily call themselves to the viewer's attention. To give some typical figures, a DVD-5 disc of 133 minute duration with a single stereo sound track would allocate a bit rate of 128 kbit/s to the audio and an average of just over 4 Mbit/s to the video. This would fill the disc to capacity. A DVD-9 disc encoded at the same rate would be able to have a duration up to 241 minutes.

Assets

The process of DVD-Video authoring starts with the assembly of the various assets that will be used. The video should be recorded on, or transferred to, a high quality format such as D5 (which with 10-bit uncompressed coding and component format is probably optimum), Digital Betacam or D1. Any noise present in the image will make the encoder's work harder and the result may not be satisfactory. The MPEG-2 encoder regards noise as detail, and detail obviously sends the bit budget soaring. Since none of the video formats mentioned are capable of recording 5.1 channels of audio, if surround sound is required, then audio must be supplied on a separate medium timecoded to the video. The finished DVD-Video may need still images, which should be supplied at the resolution required in the end product. The final component will be text files for captions, each of which will include the text itself and the timecode values between which the caption will be viewed.

Parallel to the collection of assets is the process of storyboarding that links the assets to menus and disc navigation. Like the storyboard of a film drama, the DVD storyboard will show the creative team where they are headed with the project, and also make sure that all the assets slot together as they should. Indeed, are all the assets available and have they been put in place? The storyboard also helps in the allocation of resources and keeps the potential complexity of the production under control. The storyboard will also lead to the production of the bit budget.

Assembly of assets according to the storyboard is the next stage of the process where video files (perhaps more than one for multiple angles), audio files, stills, subpictures and navigation are grouped together into Video Objects (VOB). The video objects form part of a hierarchy, at the top of which is the Video Title Set (VTS), each disc requiring at least one, and the Video Manager (VM) file. The video manager contains information such as region coding and the initial menu shown when a disc is inserted. Within the VOB, as well as the video and audio data, is a Program Chain (PGC) that provides interactivity via a programming language specific to DVD-Video.

Copy Protection and Regional Coding

One of Hollywood's great concerns is that of copying of DVD-Video discs. This happens at two levels: mass copying by pirates who you can take for granted will have the technology to work around any protection system (if they had to, they would point a video camera at a TV screen!). The second level is that of home copying where someone might buy a disc and copy it for a friend. There are of course issues surrounding the nature of ‘legitimate use’ – this time ‘making a copy for the car’ isn’t one of them – but let's concentrate on the technical side of things.

The DVD world is divided up into a number of regions:

- North America and US territories
- Europe, Middle East, South Africa and Japan
- Southeast Asia and East Asia
- Australia, New Zealand, Pacific Islands, Central America, Mexico, South America and the Caribbean.
- Eastern Europe, Indian subcontinent, Africa, North Korea and Mongolia
- China

This seems to cover most of the world but there are a few more regions to mention. Region 7 hasn't yet been allocated. Region 8 is for international venues such as aircraft and cruise ships. The final ‘region’ is Region 0. This unofficial region refers either to a disc that is not regionally coded, or a player that is capable of playing discs from a variety of regions. Region 0 players are in theory not allowed under the terms of the DVD-Video specification, although word has it that NASA has taken a couple to the International Space Station so that should lend it a certain degree of authority.

Regional coding was instituted so that Hollywood productions could be drip-fed into the various markets on a timescale that would suite the studios, rather than making them available worldwide simultaneously. It is interesting to consider that there is a film industry even bigger than Hollywood over in India – Bollywood. Presumably they have the same concerns. Regional coding is not an encryption system of any kind. It is just a piece of information on the disc that tells the player whether it is allowed to play it or not. The player has to be compliant for the system to work. However there are a great many players on the market that are so-called Region 0, either modified or designed as such, that could until recently play any disc. Hollywood has however hit back with Region Code Enhancement (RCE).

Region Code Enhancement first appeared on The Patriot and Charlie's Angels. Such discs will not work in players that are set to play discs of all regions. However they will work in players that are switchable between regions. A disc with RCE appears to the player to have all of its region flags set, so the player doesn't know which region it is. The disc then sends a query to the player asking what region it is, and aborts playback if it gets the wrong answer, putting up a message on-screen saying that the player may have been altered and the disc is not compatible.

It is worth saying that even despite regional coding, you can't necessarily play any disc, from anywhere in the world, in any player. The problem is that North American discs are in 525 line 30 frames per second format; European PAL discs are in 625 line 25 fps format. NTSC players normally do not play PAL discs, so it is fortunate that there aren't as many people who want to do this as want to do the reverse! PAL players will normally play NTSC discs, but the problem is in the output, which is still 525/30. Many PAL TV sets will in fact play 525/30 quite happily as long as the color carrier is modulated at the PAL standard of 4.43 MHz, which many players will obligingly do. Otherwise a true multistandard TV set will be necessary.

Copy protection is another matter. As mentioned last month, the encryption system used in DVD-Video's video object files – Content Scrambling System (CSS) – has been cracked (deCSS). Ostensibly this was so that Linux devotees could play DVDs on their computers, which wasn’t officially allowed because of the open source nature of the Linux operating system. It is a fact however that deCSS software can be obtained for the Windows operating system for which people could just buy their discs in the normal way. Still, I would imagine that people who copy discs for viewing pleasure, rather than for the pleasure of beating the system, are very few in number.

The above refers to digital copying. There is of course the possibility of copying in the analogue domain, although why anyone should want to copy a sparkling DVD-Video disc onto fuzzy VHS is difficult to understand. This is prevented by the Macrovision copy protection system built into most players. Macrovision is activated by a flag on the disc and adds spurious pulses to the signal that cause most video recorders to show colored bands, distortion or a picture that alternates between very dark and very bright. Modern TV sets (although curiously not all video projectors) should be unaffected. Once again there are means to circumvent Macrovision but it should certainly be enough to deter the casual would-be copier.

The Future of DVD-Video

Looking into the far future we can imagine, around 2006 or so, a high definition version of DVD-Video capable of data rates up to 20 Mbit/s that are necessary for HDTV. On the other hand, US experience casts significant doubt on the majority of viewers – or even a significant minority – having HDTV-capable receivers by then, so this might not be as plausible a scenario as it might appear. In the immediate future of DVD-Video there are two beacons (or warming lights, depending on your interpretation) shining. One is the ability to produce DVD-Video discs on a standard desktop computer rather than a specialized system with very high cost software. Within a couple of years this market will extend to the camcorder enthusiast outputting home movies to DVD-Video. As always, the dissemination of technology in this way initially provides opportunities, and then becomes a threat as early high-cost installations suddenly have to compete with people providing a similar service, at least technically, from a back room. Remember that the early CD writers cost several thousands of pounds/dollars and now everyone has one. The second is the fulfilment of end-to-end copy protection among all devices that are used for video including camcorders, VCRs, DVD-Video players (and recorders), computers and DVD drives. This could be in the form of Digital Transmission Content Protocol (DTCP) proposed by Intel, Sony, Hitachi, Matsushita and Toshiba. Under this regime, connected devices would exchange digital keys and authentication certificates and transfer data, if allowed to do so, in strongly encrypted form. Security can be renewed as necessary, or revoked. There are other systems with similar intent. One presumes that whichever system is adopted will be well enough thought out not to inhibit the activities of the overwhelming majority of genuine users.

In conclusion, it still bears repeating that DVD-Video has been a fantastic success and the world is prepared to adopt new formats if they are done right and presented well. Without doubt DVD-Video will continue to invade our domestic and professional lives (and maybe one day we really will be able to throw out that old VHS!).

Check Questions

- How might the playing quality of a compact disc be negatively affected, in a way that can be corrected?
- How might the playing quality of a compact disc be permanently spoilt?
- What is the maximum standard duration of a CD?
- Approximately what is the maximum duration of a CD, beyond the normal standard?
- What does 'constant linear velocity' mean, in the context of CD?
- Describe briefly the need for coding.
- Briefly, what is the function of the P subcode channel?
- List the functions of the Q subcode channels.
- Describe the difference in structure between CD and DVD (that makes DVD tougher).
- What is the duration of a DVD-5 disc?
- Describe dual-side discs.
- Describe dual-layer discs.
- What is the maximum data rate of DVD-Video?
- What is the typical data rate of a DVD-5 disc?
- What are 'assets'?
- Briefly describe regional coding.

Chapter 9: Perceptual Coding

Perceptual coding is a technique used to reduce the file size of a standard pulse code modulated (PCM) audio file. Similar techniques are also applied in video. Perceptual coding is also often known as compression (which is misleading as the data is not compressed to make the file size smaller) or data reduction. The most commonly known form of perceptual coding is MP3, which is the abbreviation for MPEG1 layer 3 audio. MPEG stands for the Motion Picture Experts Group who supervised the development of the MP3 standard. The MPEG2 data reduction system, also uses the MP3 standard for the audio. There are also layers 1 and 2, which represent earlier phases of development. Layer 2 is useful where mild data reduction is appropriate, and most MP3 decoders should be able to handle it.

Other perceptual coding techniques include Dolby AC3, Sony's ATRAC as used in Minidisc, Windows Media, and the RealAudio codec. ('Codec' is a contraction of 'coder - decoder'. A coder may also be called an encoder).

The requirement for perceptual coding comes from the high bit rate of audio compared to the bandwidth and storage capacity of distribution media. There are five significant uses:

- Internet audio: The bitrate of standard CD-quality stereo audio is approximately 1.4 Mbit/s. This is very much higher than a standard 56 kbit/s modem or typical 512 kbit/s broadband Internet connection can accommodate. Perceptual coding can reduce the file size so that an audio file may be downloaded in a reasonable time via a modem, or even 'streamed' - played live - via a broadband connection.

- Film sound: There isn't enough room on a film print to store all the data necessary for 5.1 channels of raw PCM. (The '0.1' channel is the low frequency effects channel which requires only a small bandwidth). The Dolby AC3 codec is used to reduce the bitrate so that 5.1 channels can be combined into a single 320 kbit/s bitstream and the data stored optically in the small area between the sprocket holes. Sony's SDDS system is similar. DTS writes timecode onto the film print and stores audio, also perceptually coded, on CD-ROM.

- DVD Video: The capacity of a DVD Video disk ranges from 4.7 Mbytes to 18 Mbytes according to the construction of the disk. Most of this is used for the video content. For a disk manufactured for NTSC territories, the audio format can be stereo PCM (if there is sufficient space) or Dolby AC3 (5.1 channels). It is mandatory to have one or the other of these. Additional perceptual coding formats are allowable as alternatives. Disks manufactured for PAL territories (and territories where SECAM is used for analog television) must have either stereo PCM, AC3 (1 to 5.1 channels), MPEG1 layer 2 (stereo) or MPEG2 layer 2 (1 to 5.1 channels).

- Digital television: The US ATSC standard specifies Dolby AC3. The DVB standard commonly used elsewhere in the world can allow AC3 or MPEG.

- Personal stereo: Modern personal stereos use either a solid-state memory, or a physically small hard disk. In both cases the requirement is for a great bulk of storage. MP3 coding can allow ten to twelve times as many songs to be stored in a given memory capacity than PCM, at a level of quality that is acceptable to most users.

Before perceptual coding techniques arrived in common use, there were other methods to reduce the quantity of data in an audio file:

- Combine stereo channels to mono.

- Reduce the sampling rate. Audio sampled at 44.1 kHz could be resampled to 22.05 kHz, 11.025 kHz, or other sampling rate. Naturally, this proportionately reduces the frequency range of the audio.

- Reduce the resolution (bit depth). A 16-bit file would typically be reduced to 8-bit. This increases noise and distortion significantly.

- Use non-linear coding so that high levels are quantized more accurately than low levels. This results in modulation noise that changes level according to the level of the signal.

- Use a scaling factor so that a 14-bit signal could, for example, be encoded into 10 bits with a scale factor to indicate whether those

10 bits should represent the highest signal levels, intermediate signal levels, or low signal levels. This is the technique, at least part of it, used in NICAM digital audio for television in the UK and elsewhere.

From the point of view of subjective acceptability none of these is ideal, and the trade-off between bandwidth and audio quality is evident.

Masking

The reasoning behind perceptual coding is that if it were possible to analyze a signal in such a way that only the features that were most important to the ear were coded, it shouldn't make too much difference simply to discard the rest. This is the basis of all perceptual coding techniques, and in fact up to 90% of the data in an audio signal is completely irrelevant to most listeners. The trick is to sort out the most meaningful data so that as little as possible that would have been of interest to the ear is discarded. Going beyond the 90% threshold, perceptual coding can also be used to produce an even smaller file size where the audio quality is obviously degraded. This is of use for communication, such as telephony.

The key to perceptual coding is the phenomenon known as 'masking'. To give this a rough and ready definition, masking occurs when a signal that is interesting to the ear renders another signal occurring simultaneously, or near-simultaneously, inaudible. A good example is the noise produced by a standard cassette recorder. While a song is playing, the noise is inaudible because it is masked by the signal. Between songs however, the noise is very clearly audible because it is no longer masked.

More precisely, masking occurs in both the frequency domain and the time domain:

- A low level signal that is close in frequency to a high level signal will be masked.

- A low level signal that occurs shortly before or shortly after a high level signal will be masked.

Masked signals are not removed in perceptual coding, but allocated fewer bits. Therefore the signals that the ear hears clearly may be fully described, while signals of lesser importance are 'roughly sketched'.

Encoding

Taking MP3 as an example, the original signal is split up into 32 frequency bands for analysis. Each band is analyzed separately to identify masked signals that can be described in less detail with little subjective loss. Subsequently in the time domain, the signal is split up into frames (1152 samples per frame, approximately 38 frames per second, regardless of bitrate), each of which is analyzed with a view to identifying temporally masked signals, which once again can be described in less detail. Since there is actual information loss (note that data does not necessarily equal information), this type of coding is known as 'lossy'.

From this point, there is also some further lossless coding that can be done, known as Huffman coding. Put simply, Huffman coding codes frequently appearing data sets with few bits. Data sets that occur infrequently are allowed more bits. It is rather like the old Morse code where the most commonly used letters 'e' and 't' are encoded as a single dot or dash, while less frequently used letters such as 'j' and 'q' are allocated longer combinations of dots and dashes.

Bitrate

Audio can be perceptually encoded using a variety of bitrates. The common bitrate for MP3 audio on the Internet is 128 kbit/s for a stereo file. For Dolby AC3 as used in film sound, the bitrate is 320 kbit/s for 5.1 channels. Fairly obviously the higher the bitrate, the fewer approximations the coder has to make when coding masked signals. The lower the bitrate, the more lossy the process is.

MP3 allows for both constant bitrate encoding (CBR), where the bitrate is specified before coding, and variable bitrate encoding (VBR). With variable bit rate encoding, it is necessary to set a 'quality level'. Different encoding softwares will specify this in different ways, but a scale of 1 to 10 is reasonable. The encoder will then allocate more bits to complex signals, fewer bits to simple signals, aiming to keep a consistent quality level.

Joint Stereo

An MP3 file can be mono, stereo or 'joint stereo'. In normal stereo mode, the encoder may allocate more bits to one or other channel according to which is the more complex from moment to moment. Thus a 128 kbit/s stereo file can sound better than a 64 kbit/s mono file. With joint stereo, the coder looks for similarities between the channels that can allow coding in fewer bits, at the expense of a slight degradation of the stereo image.

Metadata

Modern audio formats allow for additional data known as metadata. Metadata is information about the program material or 'data about the data'. In MP3 for example, the metadata can include the song title, composer, label etc. in AC3, the metadata can include information on how the dynamic range of the original audio is to be manipulated to suit various playback settings - full blown home cinema, late night listening, simple mono receiver etc.

Check Questions

- Why is perceptual coding necessary?
- Describe briefly the use of perceptual coding in the following:
- Internet audio
- Film sound
- DVD-Video
- Digital television
- Personal stereo
- What can be done, other than perceptual coding, to reduce bitrate?
- What is masking?
- How do perceptual coding systems handle signals that are probably going to be masked by other audio?
- What is Huffman coding?
- What is the typical bitrate for an MP3 file intended for Internet distribution?
- What is the bitrate of Dolby AC3 as used in film sound?
- What is metadata?