Chapter 10: Mixing

This chapter is taken from
Sound for Film and Television,
2nd edition by Tomlinson
Holman, Focal Press.

Introduction

The process of mixing sound tracks together has slightly different terminology applied to it depending on the field of use. In film work, mixing is called *re-recording* or *dubbing*. In television production, the words *mixing* or *sweetening* are more commonly employed to name the process.

The term *re-recording* is perhaps the most self-defining of these terms. The term *mixing*, on the other hand, can be applied to a wide variety of tasks, from live sound mixing at concerts, to combining the tracks of a multitrack recording into fewer tracks for release of a music compact disc, to making adjustments in level in transferring one two-track master to another. *Re-recording* is a more limited term, meaning taking something already recorded and distilling it down by mixing processes to a more convenient representation from many units, elements, or tracks to pre-mixes, or from pre-mixes to final mixes. *Dubbing* is synonymous with *re-recording* when applied to the mixing process, but it is also applied to any copying process that may occur in audio or video recording.

Laydown, *layover*, and *layback* are terms that essentially mean x-copying at the various stages of a television production. Laydown means capturing from the production sound medium into an editing environment, usually a workstation. Layover is synonymous with dub (used as the second definition above), and with x-copy, but these two terms are more associated with film than video. Layback is the copy from the sound master to the composite video master, usually the last step in the process.

Re-recording is typically done on a dubbing stage by from one to three mixers working simultaneously on the program material. Arranged to be like a theater, with the mixers occupying the best audience seats, the basic idea is to combine the tracks of the relevant stage of processing, while simultaneously manipulating them for the best sound quality and the desired effect, while hearing the movie under the same conditions as an audience would in a theater.

In post production, it is necessary to have a system that can "rock and roll," that is, respond to commands and go forward or backward, at high or normal speed, while maintaining sync. In fact, one of the distinguishing factors of film dubbing compared to other methods is the ability to "play in reverse," which is revered by mixers who work on the sound processes during reverse play and so maximize their productivity.

Film has traditionally been projected for film dubbing, but its slow speed access to different parts of the reel has given way in recent years to dubbing while showing a projected video image, which may be available from a nonlinear editing system and thus able to "reset" to another time instantly. Entering the mixing room has also been the Digital Audio Workstation, usually of the same type used by the sound editors. In this way, a sound editor can be on standby to make rapid editorial changes if called for. Also, digital film dubbers offer offset capabilities, so sound can be slipped in sync relative to other sounds during dubbing without tying up editorial facilities.

Sweetening for television programs is typically done in a relatively small studio by one or two mixers working simultaneously, which is arranged to be closer to a living room than to a theater in size; however, more expensive TV dubbing is done in larger rooms.

Sound source devices used in re-recording

Traditional dubbers played magnetic film across virtually any number of machines interlocked by a signal called *biphase*. Magnetic film could carry from one to six tracks per strand of film, and interlocking multiple machines meant a very large number of audio tracks could be played simultaneously. Large studio sound departments had machine rooms with up to about 100 machines, each of which could theoretically carry up to six tracks, but this centralized room served multiple dubbing stages simultaneously, and none of them could handle 600 tracks, although 200 was not unknown. Devices were built to provide a slip sync function for one machine at a time, so sound effects could be quickly resynced if relatively simple changes were needed. Edit/change rooms were usually made available so that sound editors could quickly revise units more thoroughly as needed. Television post production has relied on 2-in. 24-track recording for many years as standard, until recently.

Modern post production has mostly done away with the magnetic film dubber in the last few years, although a few low budget films still use them, and they will be in use for a very long time to come to play back legacy recordings. The number of 2-in. machines in use has also begun declining. At first the replacement was a generation of multichannel digital audio tape recorders called MDM as a class, for *modular digital multitrack*. The colloquial name for these machines is "DA-88," named for the first machine of the class, but the tape format is more formally named DTRS. These 8-track machines cost about one-tenth the price of an analog dubber, so were quickly adopted. They could be ganged by control functions, so that a large number of tracks could be supported. The sound quality was decent, although the machines were 16-bit linear PCM, that did not have as much dynamic range as the magnetic film or tape with Dolby SR that they replaced (see Table 6.3). Combining many 16-bit tracks results in a dynamic range of less than 16 bits, so this is seen as a limitation. Also, the first machines lacked confidence head recording, although this was added to later models.

Modular digital multitracks were a way to store a lot of digital content cheaply, but access to that content was limited to the linear domain. To go from the end to the beginning of a reel, the tape had to be rewound, and then on pressing play, would have to "sync up." Although the machines were relatively cheap, faster operation could be achieved if a random-access based system were to be used. Also, MDMs could not play in reserve as dubbers could, so mixers considered them to be a step backwards.

Thus digital dubbers were developed, characterized by random access, quick lock up, play in reverse, and control integration features for dubbing, and control over record insertions called punch ins, described later in this chapter. The format of storage is hard disk or magneto-optical drives. With this development, software was made available to read some standard DAW file formats directly, so putting sound up on a digital dubber usually means simply plugging in a SCSI drive mounted in a carrier.

Digital dubbers are on the order of double or triple the cost of MDM machines, but being purpose built with the features desired for the film and video post production market, have been successful. Meanwhile, digital audio multitrack machines based on hard disk drives have come on the market too. Usually in a 24-track format on hard disks, they are used more widely than dubbers for all kinds of recording. Their large-scale of manufacture makes them cost about what an MDM machine costs, with the added benefit of random access. However, they have limitations in dubbing applications, such as no play in reverse function, etc. Their low cost makes them attractive, though, for those willing to give up on features.

DAWs are also increasingly making their way onto dubbing and mixing stages. An advantage is instantaneous ability to re-edit. This is also a disadvantage if the process is not disciplined. Using an expensive dubbing/mixing stage as an editing room is not a productive use of the producer's cash. DAWs may deliver their output over digital audio interfaces such as AES3, in which case the workstation is tied up supplying sound to the console, or their media can be demounted and placed in a digital dubber or multitrack, assuming there is file format interchangeability between the DAWs and console. Some dubbing stages have been built recognizing this trend, with booths open to the stage on one side, but surrounded on the others.

Mixing consoles

Mixing consoles used for dubbing are often large and intimidating, with hundreds to thousands of controls. Luckily, there is a great deal of duplication among the controls, so by learning just one area of a console one learns nearly all of the areas. Fundamentally, what goes on inside consoles can be broken down into two ingredients, processing and configuration. Sound processes are the devices used to affect the sound, including all the way from simple internal level controls to sophisticated outboard reverberation units. Configuration issues are about signal routing from the input to the output of the console through the various processes. Of the two, it is processes that are easier to study, because they are represented by the knobs, switches, and other controls on the console. Configuration, on the other hand, is more obscure, since modern consoles are so dense with controls there is no room to draw diagrams on them showing the electrical order of the controls. Given this outlook, we will look into processes in some depth and give the general principles of organizing them by configuration.

It is worth pointing out that each new console faced by a professional is just as much a sea of knobs to them, as it is to the pre-professional until the array can be broken down into logical units and addressed singly. While a professional will recognize a great many of the knobs for the processes they represent, everyone inevitably needs training on the configuration of the processes.

Before describing processes, one organizing feature is worth noting. In many consoles, the construction is such that a series of processes that are associated with one input are arranged vertically in a console *slice*. This means that a primary issue in configuration is accounted for by this fact: An entire column of knobs is likely to be associated with processing the signal from one source. With that in mind, various processes that may appear in an individual slice are first described, and then variations from this standard.

Obscuring the classical distinction between editing and mixing is the fact that DAWs today have many mixing features, and may even have more potential different processes available as software "plug-ins" than major consoles. Console control surfaces that operate the functions of DAWs are becoming popular. The distinction between DAWs equipped with a control surface and large consoles is usually that, if the console is digital, it will have dedicated digital signal processors for each channel, and thus may be designed not to "overload" under the burden of signal processing, and possibly crash or lose signals. DAWs are more likely to dynamically assign resources like digital audio signal processing power, so could run out if a great many signal processes were in simultaneous use. This can often be solved by plugging in more hardware to the DAW, but then its cost may approach that of a console.

Processes

Level

Setting the level of each of the elements of a mix is surely the single most important item to be done in mixing. Even the simplest of equipment has a means to adjust the relative level (also called *gain*) or volume of the individual elements by way of *faders*, also called *potentiometers* or *pots*. The reason is simple. If each recording has been made to make best use of the medium, then a Foley recording of footsteps, for example, will be recorded about as loud as a dialog recording. When re-recording, though, it is necessary to get these various elements into balance with one another, so inevitably the Foley element will be turned down relative to the dialog element in order to assume its proper relationship in the mix.

The main level control for each input is given more weight than any other console process by the placement and type of control. On re-recording consoles, the main channel fader is always the control that is largest and closest to the operator, and is usually a vertical slider type of control with markings for resetability.

Another related primary control is called a *mute*, which is simply a switch that kills the signal altogether, allowing for a speedier turn-off than turning the fader all the way down rapidly. Mutes are probably more commonly used during multitrack music recording than during film mixing because in music all tracks are on practically all of the time, whereas workstations produce silence when there is no desired signal, thus accomplishing muting right at the source. Music mixes may mute individual channels for whole sections of the mix, say, the string channels during a brass solo, to prevent audible crosstalk into the unused channels from the open, but unused, mikes. The mute function also provides a means of identification of where a sound might be. By activating the mutes in turn during a trial run, the mixer can learn where the various sounds are, in case the cue sheets are faulty or unclear.

In cases where there has been too much sound cut for a sequence, and not enough time to change it editorially, a mute function may become valuable. Here, a computer follows the action of switches throughout a reel, keeps track of what is muted and what is unmuted, and performs the mutes on subsequent passes. Like fader automation, mute automation can be built up into a complex pattern over many passes.

These two functions, level and mute, are so important that they are the first functions to be automated in more elaborate consoles. A means to mute all other channels and to monitor only what one channel is contributing to the mix is very useful. This function is called *solo*. Pressing the solo button on a channel will make the monitor mute all other channels. Pressing more than one solo button will produce a mix of only those channels.

> There are variations on the solo idea. Most solo systems offer only a "cue" or "audition" function, so the signal processing, such as the position in a stereo mix, is lost when solo is activated. Some offer "solo in place," representing the sound being soloed correct spatially. Many solo systems are "destructive," that is, the output mix of the console is affected by the solo function so it cannot be used during principal mixing, but others are "nondestructive," affecting only what is monitored, not what is

recorded. So individual consoles vary greatly in their possible solo functions.

Multiple level controls in signal path

On its way through a modern console, a single signal may well pass through a large number of level controls—individual channel fader, subgroup master fader, master fader, and monitor volume control. This multiplicity of controls, while offering high utility and flexibility, also creates problems. The problems are similar to those of recording on a medium: If a tape is under-recorded, when the level is subsequently restored by "turning it up," noise will become evident. Conversely, if over-recorded, the resulting distortion is permanent and will not be removed by "turning it down" at a later stage. While consoles generally have a wider dynamic range than recorders, hitting the dynamic range of each of the intermediate stages correctly is an important issue to avoid excessive noise or distortion.

On the best professional consoles, with their multiplicity of controls, attacking this problem of the correct setting of the variety of controls is accomplished relatively easily. The scale on them is the clue, with 0 dB the nominal setting of the controls. Many of the controls have "gain in hand," which goes above 0 dB, that is, one can "turn it up" from the nominal in order to reach for something under recorded as needed, but the nominal setting is clear. On consoles that lack this feature, it is necessary to determine which settings of all of the controls are the nominal ones. It is usually the channel fader for each slice on which most of the actual mixing is performed. The other controls, such as submasters or master level controls, are used for slight trims to the overall section-by-section balance, or for the main fade-ins and fade-outs of the overall mix. On the other hand, since the individual channel slice gain controls can be used to set the "balance" among the parts of an effect, a submaster can be used to set the overall level of an effect.

One problem with using DAWs as mixing consoles is that they don't typically provide very much "gain in hand." Their design assumes properly recorded levels coming in. Older film consoles had as much as 25 dB gain available over the nominal to accommodate under recorded material, whereas today, with wider dynamic range available on source media so that under recording can be better tolerated, some DAWs only have 6 dB. The designers probably do this to avoid clipping distortion that would otherwise be a strong potential for some material—to keep the operator out of trouble. Other digital console designs have as much as 45 dB gain in hand, so are designed to better tolerate under-recorded sources.

Dynamic range control

Compression

Each track used for re-recording has a volume range. When tracks are combined in mixing, the problem of unintentional masking of one signal by another arises. Let us say that we have a dialog recording, with a volume range, and a music recording, with its own volume range. Starting at the beginning of a show, we have music as the foreground, but it is not theme music in this case that fades out before the dialog begins, but rather source music, fading under the dialog. The problem is that while most of the time the music will lie underneath the dialog, there may be a point in time at which the peaks of the music correspond to the minimum level of the dialog and the dialog is obscured. On the other hand, there will also be times when the music is faded under so that it seems to go away altogether. The alternating presence and absence of the music is distracting.

In order to solve this problem, we could "ride the gain," turning the level of the music down during its higher level passages and up during its softer ones, to maintain a more even level behind the dialog, but this would be tedious and time consuming. The process can be automated by a device called a *compressor*, which does just what we have described. A compressor is equipped with a number of controls to vary the volume range over which the action of the compression occurs, the amount of the compression, and how fast or slow the compressor acts. Each of these devices is fairly idiosyncratic as to control functions, so the number of knobs associated with them varies.

A typical compressor may have the following control knobs:

1. A "threshold" control, below and above which the compressor exhibits a different "transfer function." Usually below the threshold, the compressor acts as a linear amplifier, such that each deciBel in yields a deciBel out, and above the threshold each deciBel in results in less than one deciBel out.

2. A compression ratio control with markings such as 2:1, 4:1, 20:1, or more. This is the ratio between input and output in deciBels above threshold. A 4:1 compression ratio would mean that a 4 dB change in the input produces a 1 dB change at the output, above threshold.

3. An output level control. Because compression often has the effect of lowering the overall level, this control is used to make up the gain and to raise the overall level after the compression process.

4. An attack time control. This control modifies how quickly the controller circuitry responds to an increased level. If it changes too quickly, short sounds that do not reach full perceived loudness may control the gain excessively and "pumping," (audible gain riding) may result. If it changes too slowly, then loud attacks may be heard followed by a level change downward, also leading to pumping. A typical starting value for this control is 80 msec.

5. A release time control. This control modifies how the controller acts when the signal decreases. If the control function is made too fast, the gain will change within one cycle of the signal, which leads to harmonic distortion. If the gain change is set too slow, then soft sounds following loud ones could be lost completely. A typical starting value is 0.5 seconds.

Of course, technically speaking it would be equally possible to compress the dialog into a narrow volume range in order to keep it above the music continuously, but even with the music in correct overall balance then, it may seem to come and go, which can draw attention to it. The principle of "least treatment" for this scene would say that the mixer should process the background sound first rather than the foreground sound, to leave the fewest artifacts present. On the other hand, I have found is useful to use small amounts of compression on dialog to make it sound more natural. The reason for this may be that we are recording dialog typically at one point in space, but we hear at two points. The spatial averaging of level that occurs by averaging two points tends to reduce the stronger level differences observed at one point. Thus slightly compressed sound may actually be a better representation of what we hear than "pure" uncompressed sound. But note that I am speaking of small amounts of compression, on the order of 6 to 8 dB of maximum gain reduction from the linear condition.

If the dialog and music tracks have already been combined, then compression is not an option. The louder track dominates the "thinking" of the compressor, moment by moment. A likely outcome of this condition is that the level of the music would audibly go up and down with the level of the dialog. Assuming the dialog is practically always above the level of the music, when the dialog is soft the compressor will "turn it up," bringing up the music as well; and when the dialog is loud, the dialog and music would be "turned down." This is an effect known as *pumping,* when the level of one element of a mix audibly affects another element in level and is generally undesirable. It is thus best to compress each individual source alone and then to combine sources, rather than to try to compress the entire program.

An aesthetic use for a compressor has been described for keeping music "under control" while faded under a dialog source. This use would be made despite any other consideration in the overall system because it is an aesthetic one, but there is another reason to use compression. Generally, post production works with media that have a wider dynamic range capability than that of the ultimate release format.

Another use for compression is to reduce the large dynamic range of a theatrical presentation to the smaller range felt necessary for home video. Such compression is not routinely used in video transfers, but may be used in special cases where the supervisor of the transfer is certain of the conditions of use. If all users are expected to listen to a program over a television set internal loudspeaker, then that may be given weight in a compression decision. For instance, children's VHS video and airline and hotel video copies are compressed compared with the theatrical version of the same movie. So, there are different uses for compression—in re-recording to control certain tracks to make them more manageable, and subsequently to "fit" the program material's volume range to the range needed by individual media or users.

Expansion

The opposite of a compressor is an expander. An expander increases the volume range of a source, and may do so across a wide dynamic range, or may be restricted to a narrower region by control functions. Restricting expansion to just low-level sounds is often used in re-recording. Called *downward expansion*, *noise gating*, or *keying*, this function turns the level down below a threshold set by a control. For example, all sound below say –40 dB could be rapidly turned down to essentially off. The advantage of this is that there is often low-level noise on each of a number of tracks, that is undesired, and that would be a problem if all of the noise sources were mixed together continuously. With use of a noise gate (which would be more properly called a program gate because it essentially "turns on" for signals above a certain level, and off for ones below that level), such additive noise can be reduced because only tracks above a certain level will "get through" the gate.

Noise gates have a number of audible problems. Let us say that we have a dialog recording with some air-conditioner noise in it. The threshold of the noise gate can be set to distinguish between the dialog and the air-conditioner noise because the air-conditioner noise is lower in level than the dialog. The problem is that we hear air-conditioner noise "behind" the dialog when speaking is going on, and we hear its "absence" in between lines. The dialog is pumping the level of the air-conditioning noise. The exaggerated difference between the noise being on and it being off may well draw more attention to it than just leaving it alone, unprocessed.

For this reason, traditional "all-or-nothing" noise gates are not used too often in critical re-recording tasks. They are used, however, in multitrack music studio work. For example, suppose we have recorded an orchestra and have placed the instrument groups on different tracks. Upon playback, we find that the string track is polluted with acoustical crosstalk from the brass section. In order to eliminate the brass from the string track during passages where the strings are not playing, we can use a noise gate, set so that when the strings are playing the gate is on, and when they are not it shuts the signal off. This operation depends on the string level being higher thatnthe crosstalk of the brass, and on setting the noise gate so that it can discriminate between them. In this application, the noise gate is used on playback, where the control function can be changed and repeated if necessary, rather than on the record side, where any mistake could not be corrected subsequently.

An advancement on the all-or-nothing approach is frequently used in dubbing. Some noise reduction devices work like gates but don't turn the signal fully off, so changes are less abrupt. A more sophisticated approach breaks the audio spectrum up into four or more frequency bands, and applies a downward expansion below a certain threshold separately in different frequency bands. Also, the expansion is not made as dramatic as turning the signal all the way off. These two strategies have the effect of producing much less objectionable side effects. Audible pumping is greatly reduced or eliminated by these approaches. The generic name for such a device is a *multi-band low-level expander*, although because that is such a mouthful, the actual units go by their trade names. These operate in both the analog and digital domains, with potentially many parallel frequency channels. All of them attempt to distinguish desired program material from background noise content, and decrease the level of the background noise without affecting the program material.

Single-ended noise reduction devices based on Dolby A or Dolby SR, known as the Cat. 43 and Cat. 430, respectively for their Dolby catalog numbers, are widely used. These devices use the multiple frequency bands of the associated type of Dolby noise reduction, but not in a manner complementary to what is usually associated with companding noise reduction. Here "single-ended" means that the signal to be processed has not been previously encoded. The Dolby units combine both level and frequency domains because they do their work in multiple frequency bands, basically turning down the level below a certain adjustable threshold. If a desired voice can be set above threshold and an undesired noise below threshold, the equipment can distinguish the two and turn the noise, but not the voice, down. Their use is usually dialog cleanup of production sound recordings.

Limiting

A variation on the idea of a compressor is a limiter. A limiter acts on signals above a certain threshold as a compressor does. Then above that threshold, the level is controlled so that for each deciBel of increase on the input of the limiter, the gain is reduced by the same amount. Thus, above the threshold, the level simply stays practically the same despite any increase in level.

Limiters are useful in production sound to "catch" occasional events that might not otherwise be controlled as to level, to bring them into a range where the recording medium can handle the signal linearly. They are routinely included in production recorders, such as the Nagra, and in camcorders for the utility they offer in keeping unexpected signals under control.

> Limiters are very useful for keeping unexpected high levels from distorting, such as an actor "going over the top" and shouting more loudly on a take than in rehearsal. With mixed program material, though, there may be a problem. Let us say that there is a gunshot in the middle of a scene. With a limiter in use, the gunshot will certainly "trigger" the limiter and the gain will be dramatically reduced. The difficulty is that it will be reduced for sounds coming after it as well, and the "duck" in level may well be noticeable for speech coming immediately after the shot. In such a case, it may be better to proceed without using a limiter, letting the tape distort on the gunshot briefly. In that way, a post production editor can "clean out" the gunshot from the production sound track by cutting it out, leaving no artifact other than a hole, and cut a clean gunshot into the sound effects track. The limiter must reduce the gain quickly and keep it that way for a while, because if it did not, the result would be distortion of low-frequency sound, so the "duck" after a very loud sound is a natural artifact of limiting.

Limiters are also useful in post production in several ways. They can:

- Put an upper limit on one track, so that it cannot rise in level so much as to interfere with another track. An example is Foley recordings of footsteps in which one or a few footsteps sound like they stick out of the mix, but lowering the overall level of the Foley track is not satisfactory, because then the overall impression is too weak. So applying a limiter set so that it catches the highest level footsteps and keeps them under control is a useful function.

- Probably the most common use of limiters is to control the highest level signals on a sound track so that they can be recorded on a particular medium without distortion. Examples include limiting for analog optical sound tracks, which have much less headroom than the magnetic tracks used as their source.

- Another use for a limiter is to raise the overall loudness without affecting the maximum level. By limiting the highest peaks of the program to a lower level, the average level can be raised without exceeding a particular maximum. With the proper sort of limiter, 5 to 6 dB of limiting can be practically inaudible, and this amount is a large advantage in fitting into the requirements of, say, an optical sound track or any other limited headroom medium.

De-essing

One particular specialized form of limiting is de-essing. This refers to the fact that many media are sensitive to reproduction of the "esses" in speech. By limiting strongly only on such high-frequency sounds in dialog tracks, the resulting sound track can be easier to record to certain media, such as optical sound tracks. Sibilance distortion is the result of imperfect waveform reproduction, when the high-frequency sounds in speech, especially "esses," are affected.

A de-esser is a limiter with one difference—the control function for limiting is made sensitive only for high-frequency sound, so most signals are unaffected. Should an "es" sound cross the threshold of limiting, then the level will be "ducked" for the duration of the "es." You might think that this would make the sound dull because high frequencies are being reduced in level, but de-essing can be surprisingly benign, having little audible effect except to reduce the sensitivity to such distortion at subsequent stages.

Conclusion

All of the items so far discussed affect the level of audio signals. By far the most commonly used of these processes is level controls, which is used even on the simplest of mixers. Muting and solo are also found on most re-recording consoles. Dynamics processing, including compression, expansion, limiting, and de-essing, are also frequently used in re-recording, but more rarely than level controls. Some re-recording consoles include all of these processes in every input path of the console, while others do not have dynamics controls in each input, but rather route signals from the inputs to separate dynamics processors, either built into the console or external to it.

Processes primarily affecting frequency response

Processes that affect principally the frequency response of the signal are probably second in importance only to level control. These processes can "clean up" the audio signal, make it more interchangeable with other signals (for instance, adjusting the timbre of production sound and ADR recordings to be more equal), adjust for the loudness effect (by adding bass for sound portrayed at lower than its original level), and generally make the sound more intelligible or pleasant, or, for effect, deliberately worse.

Most of the ear training necessary to become a re-recording mixer is involved with level and frequency response processing, with other factors such as dynamics playing a subsidiary role. The reason that this is so is that the maintenance of good continuity depends much of the time on the sound track not drawing attention to itself through unexpected changes to level or frequency response. The changes must be "smoothed out" so that the audience is not distracted by them.

The two principal frequency response determining processes are *equalization* and *filtering*. Although either of these processes may affect any frequency band from low bass through high treble, there is a fundamental difference between them. Filtering is done essentially to eliminate certain frequency ranges from the output, and thus the action of a filter changes abruptly across frequency. In fact, filters are rated for their frequency and their slope in deciBels per octave. The slope of a practical filter, how much its output changes across frequency, is generally greater than equalizers.

Equalization

Almost everyone has some experience with bass and treble controls on a stereo system. Equalization is the professional name for the process of altering the frequency response in a manner similar to what tone controls do on a stereo system. However, only the simplest professional equipment devotes only two controls to this important task. For professional *equalizers*, often found in each input channel slice in re-recording consoles, and possibly in other parts of the chain, the audio spectrum is more frequently broken up into three or four parts, which might be called low-bass, mid-bass, mid-treble, and high-frequency controls. Alternatively, the two middle controls could be labeled midrange 1 and 2.

Each of the main equalizer controls, called EQ, are rated for how much variation they produce in deciBels when set to their extremes, such as ±12 dB. These main controls may be supplemented by other continuously variable knobs or switches to provide more flexibility in use. If provided, these extra controls affect the parameters of frequency and curve shape, resulting in the name *parametric equalizer* for the device.

The first of these subsidiary controls usually available is one that changes the *frequency* range in which the EQ control is most active. Note that the frequency range of the various controls may be set to overlap, producing the possibility of up to twice as much boost or cut as for one control on analog consoles, or up to potentially four times as much on a four-band digital equalizer with a range of 20 Hz to 20 kHz in each section.

The second most likely control is one that changes the general shape of the curves produced by the equalization control. The two shapes offered are bell shaped and shelving. Bell-shaped curves are centered around a specific frequency, and at maximum boost or cut look like the outline shape of a bell, right side up or upside down respectively. Shelving curves are like conventional tone controls: Once boosted or cut, the whole frequency range from the center frequency of the control to the audio band extreme is equally affected. Thus, this type of control is usually found only on the lowest and highest frequency bands of a multiband equalizer.

The use of the two depends on why the equalization is being done. For example, a shelving characteristic is used to overcome a muffled sound associated with too much cloth over the mike, while a bell-shaped curve may be used for precise equalization of musical overtones of particular instruments. The shelving equalizer is a broad stroke, and the bell-shaped equalizer is more specific.

The third most likely control is called Q. This relates to the "sharpness" of the control function with respect to frequency. Two controls can have the same center frequency and boost, say +8 dB at 2 kHz, but they may get to that boost in a manner that is either very narrow, having little effect on neighboring frequency ranges, or wide, having effects well away from the center frequency of the equalizer. A "narrow" equalizer has high Q, and a wide one has low Q. Because low Q affects the response in more critical bands, it is generally more audible. Low Q is usually used for program equalization unless a particular problem in a narrow frequency band is the trouble, and then the high-Q version is valuable for having little effect away from its center frequency.

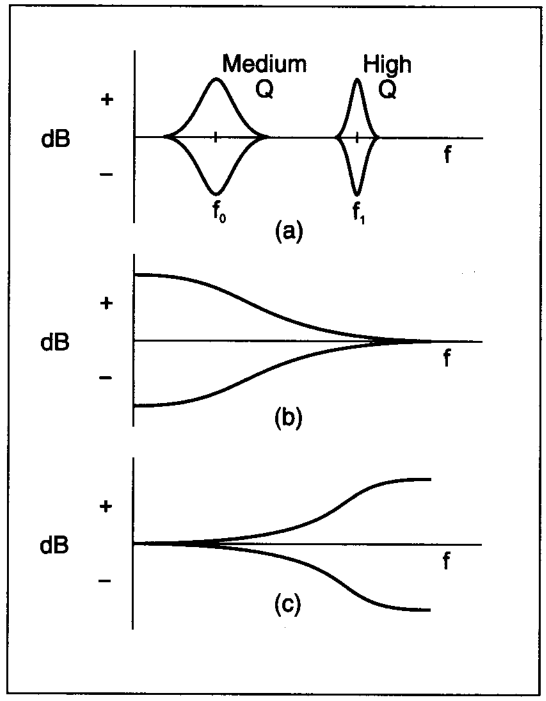

Fig. 10.1 The curves given by (a) represent commonly found "bell shaped" equalization curves, with the "Q" as well as center frequency varied between the two sections. Curves (b) and (c) represent low- and high-frequency shelving curves respectively, most commonly found as the bass and treble controls on a stereo system, but also useful in re-recording.

Professional equalizers could then have as many as 3 knobs or switches per frequency section of an equalizer, and 4 sections is not uncommon, so 12 controls affect frequency response in this scenario, and these program equalizers are usually found in every channel. On a large post production console with many inputs, controls just for equalization run into the hundreds, demonstrating the importance that equalization has for the post production process.

Another type of equalizer is also found on dubbing consoles, but usually not in each slice. *Graphic equalizers* consist of a row of multiple knobs that can be used to "draw" a desired frequency-response curve, offering a very easily grasped human interface. The number of knobs across frequency varies with the model, with six to eight being common. The curves are usually bell shaped, and most graphic equalizers offer no means to become parametric—the frequency of the controls and the curve shapes are fixed. This type is patched into the channels as needed, either into individual channels, or more likely, into groupings of channels. Due to the lack of space on the console operating surface, the use of graphic equalizers has declined in recent years.

Filtering

Filters are distinguished from equalizers by being more brute force in their action. They are intended to essentially eliminate certain frequencies from the output. The utility of filters is in corrective action generally, compensating for noise in the recording, especially low-frequency noise.

Filters may strip off any part of the audio spectrum. Probably the most commonly used filter attenuates low bass, and passes the rest of the spectrum essentially unchanged. Such a filter is called a *high-pass filter* in professional circles because it passes highs while attenuating lows. On consumer equipment, on the other hand, exactly the same device is called a *low filter* or *low-cut filter*. Equipment whose pedigree isn't certain, the category *prosumer*, uses uncertain designations, either high-pass or low-cut, to mean the same thing.

The various filter types and a typical use for them are as follows:

- *High-pass (low-cut) filter*: Used to remove excessive room noise, which is often concentrated at low frequencies.
- *Low-pass (high-cut) filter*: Used in music recording to help isolate a low-frequency instrument playing in a recording studio along with others. Isolation from crosstalk is improved by stripping off the highs from other instruments in the studio that are leaking into the open mike in front of the bass instrument.
- *Bandpass filter*: A combination of high- and low-pass filters. One use is as a "telephone filter," so-called because restricting the audible frequency range in this way sounds like one of the primary things that a telephone does to sound.
- *Notch filter*: A filter that greatly attenuates only a narrow frequency range. It is useful for removing tonal noises such as certain whines or whistles. Notch filters usually can be adjusted for center frequency and the width of the notch.
- *Hum eliminator*: A filter that has a notch at the line frequency and its harmonics, for use in reducing recorded hum.

Most of these filter types are rated by the frequency where they attenuate the signal by 3 dB, and by their slope versus frequency in deciBels per octave. Typical filter slopes are 12, 18, and, more rarely, 24 dB/octave. The notch filter is not usually rated in deciBels per octave, because its slopes are extremely steep.

Developing an ear for timbre

Perhaps the most important issue in training for mixing is developing an ear for timbre. This is quite complex on program material because it is constantly changing, and so takes a lot of accumulated impression to hear. One way to short circuit the time needed to learn to hear timbre differences is to listen to equalized pink noise, and to match the equalization with a program equalizer at hand. An unknown can be arranged by sending pink noise through a console and using one channel's equalizer to make a particular sound, and then through switching arrange for a second equalizer to be available for matching the first by ear. Of course, the first equalizer should be covered up, and shown only after a "solution" is found. Pink noise has two advantages over program material: it is constant in time, and it has all frequencies present. These two combine to simplify the experience as a starting point.

Processes primarily affecting the time domain

The former processes work practically instantaneously, in real time. Some processes, however, work deliberately to change the time characteristics of the signal, in particular, adding reverberation or deliberate echoes and echo-like effects.

Reverberators

Reverberators are very useful, like equalization, in matching the difference between production sound and ADR recordings. They are also used to "sweeten" music that may have been recorded in too "dry" a space. Another use is to help distinguish among auditory objects; thus, all sound having one reverberant character will be categorized together by hearing. This is an important feature in "layering" sound in depth from in "front" of the screen to behind it.

In the early days of filmmaking, shooting stages were made very dead acoustically, in part because theaters were reverberant, and any added reverberation in recordings detracted from speech intelligibility when heard in these theaters. In the same film, however, it was more pleasant to hear music with added reverberation, so scoring stages were built with moderate reverberation times. This dichotomy started a feature that is still present today: Dialog is generally less reverberant than the orchestral score underlying the scene, partly for speech intelligibility and partly for the aesthetics of music listening. If the music was recorded in too dead a space, then reverberation was added to the recording by playing the recording over loudspeakers located in a reverberant room, and the reverberation was picked up with a microphone, amplified, and added back to the direct recorded sound in a rerecording console. Thus, the reverb chamber became a part of film sound technique as early as the mid-1930s. While these lasted well into the 1970s, the real estate they took up became very valuable and they were not very flexible (one could change the reverberation time only by changing the absorption in the space). Starting then in the 1970s, mechanical and then fully electronic reverberators came to dominate the scene. There are many types of reverberators available today, most of which are based on digital electronics. Most reverberators are designed for music enhancement, and have "good sounding" reverberation, but these often do not have an adequate range of reverberation types for film sound because many spaces, needed to be synthesized to make the scenes portrayed, range from acoustically good to downright bad and from short to long reverberation times. For synthesizing smaller and less reverberant spaces, a *room simulator*, which is a variant on a reverberator using many of the same techniques, may be applied.

Echo effects

A digital delay line (DDL) is a device that simply delays sound by converting audio into samples and storing the samples in a digital computer memory, and then withdrawing the samples at some later time and converting them back into audio. *DDLing*, as this has come to be known, adds a variety of effects depending on the delay time and the relative strengths of the direct sound and the artificial reflection. At relatively short times, between 1 and 20 msec, strong timbral effects, like speaking into a barrel, are heard. This "thickening" of the sound is what makes the distinctively metallic voice of C-3PO in *Star Wars*, for example. Longer time delays and stronger artificial reflections sound progressively more like separate echoes.

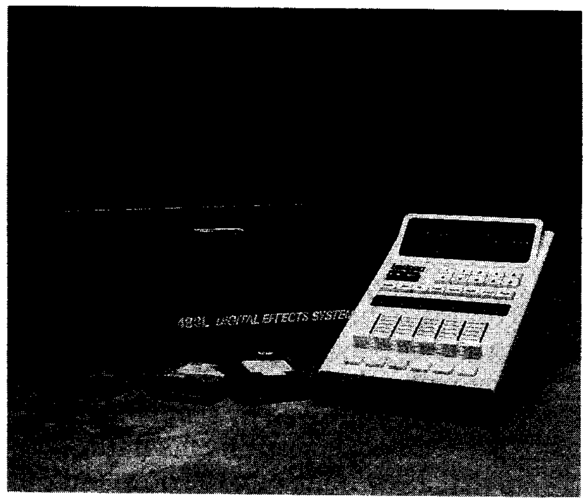

Fig. 10.2 One popular reverberation device and its control panel, which often finds its way onto the tops of rerecording console controls. This is the Lexicon 480 and its LARC controller. Photography courtesy Lexicon, Inc.

Pitch shifters and subharmonic synthesizers

Pitch shifters are digitally based units which take in sound of one frequency range, and translate it up or down to a different frequency range. Pitch shifters are sometimes known by the trade name Harmonizers. They are useful in affecting voices that are meant to be in a different range than the actor's natural voice, and to change the quality of sound effects.

Subharmonic synthesizers are devices that find the fundamental in program material, and synthesize subharmonics, as described in Chapter 1. They are useful for adding desired "weight" to effects.

Combination devices

There are devices on the market that combine two or more of the above-mentioned processes for use doing one task, such as equalizing and compressing for a vocal track to be heard in a mix with other music tracks. These are given names by their manufacturers that emphasize the purpose for which they are built, such as Vocal Stressor.

Configuration

Each of the above-mentioned processes may be needed to manipulate a channel during the mixing process, but no practical console has all of the processes present on each and every input channel. Thus, the configuration of the console becomes important in order to organize the various processes and to reduce the number of control functions to a (barely) manageable number.

Early re-recording consoles

Early film sound consoles had relatively simple signal paths. Each console input channel was equipped with a fader and patch points to insert processing equipment (which at that time included all equalizers, filters, and the like). Processing equipment was patched into input channels as needed. The input channels were summed to produce the output and were sent to a recorder. A loudspeaker monitor system was switchable between the signal sent to the recorder and the return from it, in the manner of three-head tape machine monitoring described in Chapter 6.

If reverberation was needed, the summed output of the channels was additionally sent to a reverberation chamber, and the amplified output of the chamber's microphone was summed together with the dry sound in an added summing stage before the recorder.

Adding mix in context

One difficulty with this arrangement is that it is hard to make predubs[1] having all of the necessary fades because the mixers can only hear the elements belonging to one predub at a time. Thus, if a sound-xeffects premix needs a fade under to accommodate narration, and only the effects can be heard while making the predub, it is difficult to judge the timing and the amount of the fade that is needed. This is the origin of a technique called *mix in context,* which uses two consoles in effect, although they are typically in one housing.

The actual mixing for recording is done on the first console, with its output sent to the recorder. This output is also sent to the second console, along with the other existing predubs. The second console is for monitoring purposes only, and little actual mixing is performed on it, because its output is not recorded but rather is sent to the monitor loudspeakers. The second console is usually set for "unity gain" on all of its active inputs, so the predubs are represented in 1:1 level correspondence to each other. By such a setting, the mixing that is occurring on the first console is "in context" with the rest of the existing material. If the dialog predub is played through the mix-in-context inputs, then the premix that is occurring on the first console can be done with respect to all of the dialog.

Busing

In order to account for the separation of D M & E (as they are traditionally abbreviated), the absolute minimum console must contain three signal buses, one for each part.

---

1. Synonym for premix.

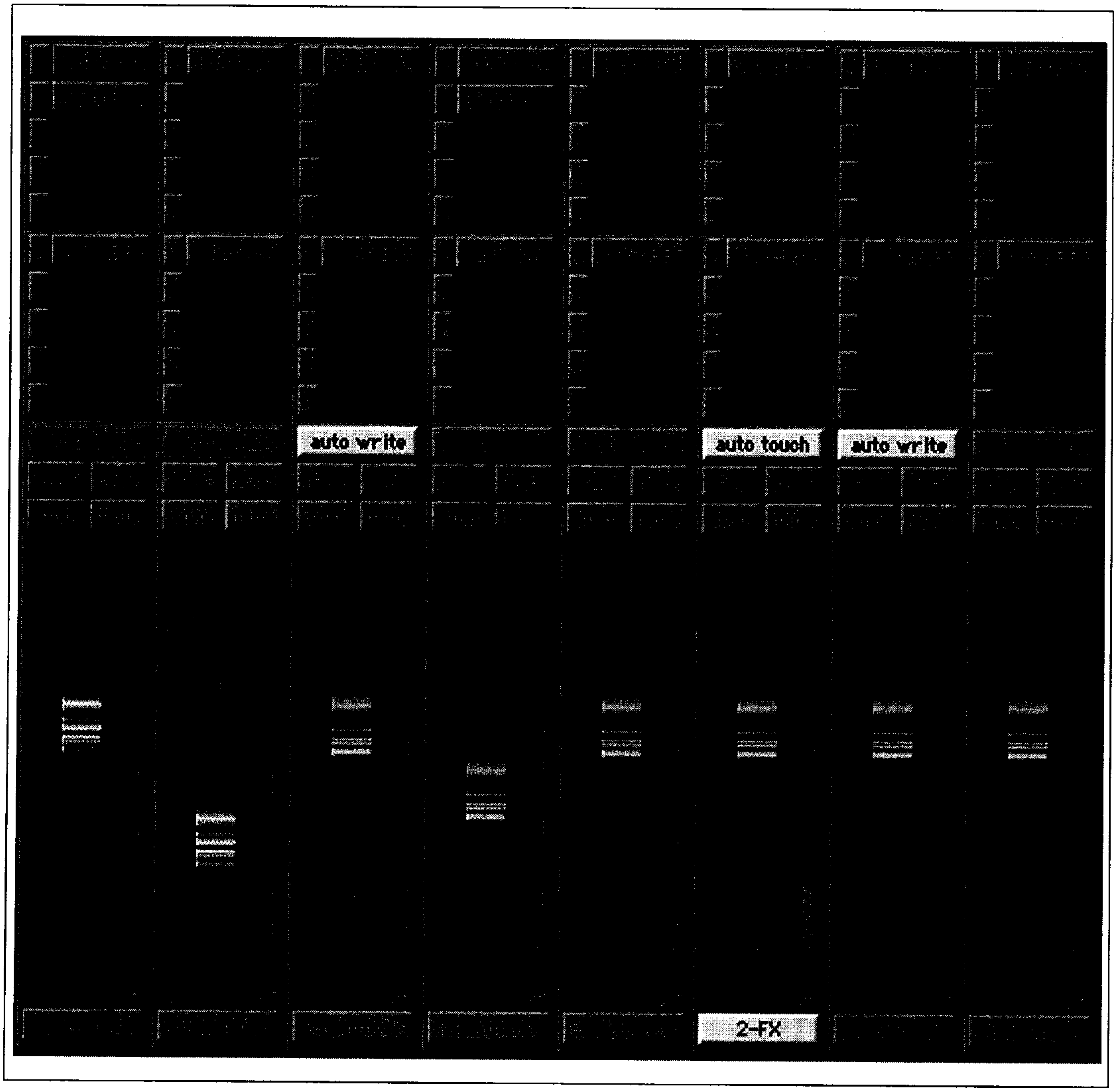

Fig. 10.3 The mixing screen of a Digital Audio Workstation.

A *bus* is an electrical connection that brings together and sums multiple inputs, much like a milk tanker driving among farms, picking up milk and delivering it to the processing plant. Note that once mixed together, we can no longer separate out the individual contributions—summed is summed, and reversing the process is practically impossible. This puts summing at the heart of the sound mixing process, because it is necessary to combine sources in order to simplify mixing, but going back on a mix is very difficult, usually requiring work to be redone.

To account for stereo production, each of the three *stems* (DM&E) needs multiple channels to represent at least left (L), center (C), right (R), and surround, with two surround channels (LS, RS) the norm today, as well as a possible extra low-frequency—only enhancement channel for effects. Each of the stems is thus recorded, either on a separate 6-track film recorder, or alternative methods provide the same functionality (e.g., 18 tracks on a 24-track machine).

Each of the three sections of major re-recording consoles usually have a minimum of six output buses in order to produce multichannel outputs capable of being panned anywhere in the sound field: left, center, and right screen channels, left and right surround channels, and the auxiliary low-frequency effects channel (that would not be panned to but rather "hard assigned").

The term *tracks* applies in several ways and may be ambiguous. Technically, a track is a space on a medium assigned to carry one signal. So a tape recorder capable of recording more than two signals[2] is called a *multitrack recorder.* The term also applies to the cut units or elements, or to the overall *sound track* of a film.

The term *channel* refers to a conceptual model of a signal path. We would think of a microphone channel as originating with the microphone and winding up on a track of a multitrack machine. An *input channel* of a console includes all the signal processes from the input to the buses, which may also be called *output channels*. The assignment of an input channel to buses is set by either hard switches or by panning among the output channels. (*Channel* is the more global term, allowing for signal processing after the summing accomplished by a bus.)

Patching

Because all facilities are not generally available to all input channels directly, there needs to be a means to *patch* specialized processes into the signal flow within a channel. All large consoles provide a way to do this, usually by way of *patch bays*, with many jacks, permitting insertion points within an input channel, or an output channel, so that processing may be applied to an individual track, or to a sum, such as to all the dialog. Alternatively, on digital consoles specialized software programs called *plug ins* may be "patched" into the signal flow. Alternate terminology is *insertion point,* into which processes may be patched.

Each physical console has its own specific rules governing patching. Some general rules that usually apply are as follows:

- Inserting a piece of equipment into a single console input channel involves patching the equipment from a source signal (called *insert send*) to the input of the process, and from the output of the process back to the signal path (called *insert return* or *receive*), usually to the point in the chain immediately after that point from which it was originally detoured.
- Patching that connects two outputs together is impermissible: They short circuit each other.
- Patching that forms loops around equipment is not permissible: The signal must always progress, not backtrack. The consequence of forming loops is the potential instability oscillation, which audibly or inaudibly (ultrasonically) is feedback, much like the acoustic feedback we hear when a public address system has too much gain and it feeds back.
- Some processes, such as older passive filters, must be patched into points with known source and load impedances so that the filter characteristics (in that case) are maintained correctly. This comes from the notion of matching versus bridging impedance discussed in Chapter 3. Although such devices are disappearing, there are still some in use, so source and load impedance conditions must be observed on those units to produce the correct results.

Panning

A fundamental part of configuration in a multichannel re-recording console is panning. *Pan pots* are devices that place sound among the channels described earlier: L, C,

2. 1-track are mono recorders; 2-track are stereo.

R, LS, and RS. There may be several knobs, such as three panning among L, C, R; front/back; and LS/RS. A joystick somewhat like that used with video games may also be used. Each input channel of the console is typically equipped with a panning section or, at the very least, with switches that hard assign an input channel to an output bus. In some cases specialized panners such as joysticks are provided on a console separate from the input channel slices and without a particular channel assignment; in these cases they are designed to be patched in as needed. In digital consoles fitted with joysticks, the stick is a controller that can be switched to take over the pan function for a slice to make the adjustment of the panning control more visceral. When the setting has been made and the moves, if any, recorded, the joystick can be switched to another channel. Some digital consoles use trackballs or graphics tablets instead of a joystick.

Many source sounds are monaural, single-channel recordings. When re-recording these sounds in a multichannel production, there are rules governing placement. These arise from aesthetic considerations based on psychoacoustics and practical considerations based on the nature of film production and viewing. One of the psychoacoustic issues was described in Chapter 2 in reference to dialog panning in the section *Speech for film and television.*

Some of these rules areas follow:

- Dialog in ongoing conversations is usually either centered or kept close to center because otherwise sound edits that match picture edits cause the sound to noticeably "jump" around the screen.
- Off-screen lines are usually panned hard left or right, as makes sense; panning them to the surrounds in the auditorium breaks the "box" of the frameline too much.
- Lines that are isolated from others in time may be panned.
- Foley is routinely recorded in mono and panned into place to match.
- Ambience is most often from original stereo recordings that are placed from two or more source channels into two or more output channels. The principal aesthetic concern of ambience panning is whether or not to include the surrounds, depending on whether or not the audience is supposed to be sharing the space portrayed on the screen.
- Cut effects could be either mono or stereo source recordings. Mono cut effects are panned into position to match their on-screen image. Stereo effects are usually two-channel recordings that are also panned in place, such as into left and center for a stereo effect located left of center. A danger arises when stereo effects are panned left and right, skipping center, discussed in Chapter 11.

Auxiliary and cue buses

So far the kind of bus discussed is called an *output bus*. Destined ultimately to reach a loudspeaker monitoring channel such as Left, such buses may be further separated by discipline, such as a left sound effects bus or channel, as we have seen. The idea of these output buses does not, however, cover all possible purposes for which we may need to combine channels. For example, we may wish to send signals from multiple input channels of the console to the same reverberator, especially for sounds that share the same space, and because reverberators are generally more expensive than other processing equipment, it is useful to share resources. For this reason, auxiliary buses have been developed. Auxiliary buses have two primary purposes, effects send and cue send. Effects send auxiliary buses are used for the purpose already described, to gather signals from multiple inputs and to deliver them to a processor. Effects return modules, similar to input modules,[3] then direct the return signal, such as reverberation, to the main buses. See Fig. 10.5.

3. But hopefully lacking, above all else, effects sends themselves, because if an effects return module could send to an effects unit, an unstable feedback loop is likely to be formed.

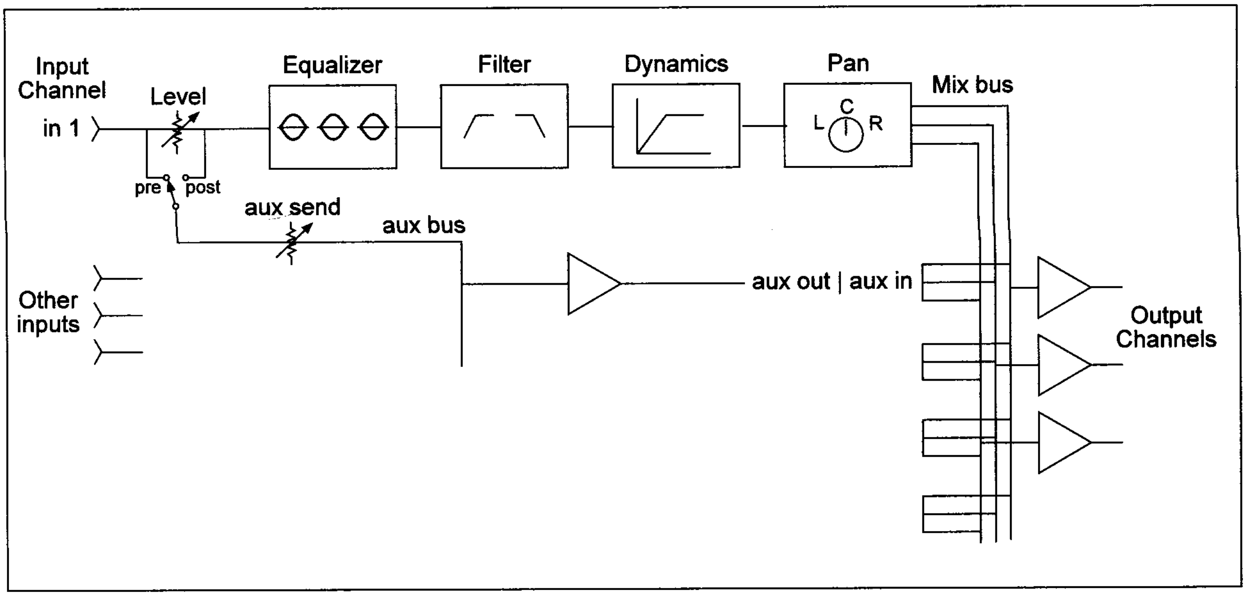

Fig. 10.4 A single-line drawing for a simple re-recording console. A main signal path routes the input channels to the output channels by way of various signal processors, while the auxiliary system provides a means to send a proportion of the signal to an output processor, such as a reverberator, and return it to the main mix bus.

The second purpose is cue send. *Cueing*, in general, in film and television production means any activity meant to alert an actor, newsperson, or musician to a timed event, so that he or she can start on time and perhaps even maintain time. In a multitrack music studio, it has a more specific meaning. There, the first track recorded is often a *click track*, the output of an electronic metronome, used for keeping musicians on time. Then, during recording the cue send feature of the console is used on the click track to send the clicks to headphones on the musicians. Because the musicians cannot play well without hearing themselves, some cue send level is added from their own instrument tracks so that they can hear themselves. Because different parts of the orchestra may wish to hear different balances among instruments (usually with themselves louder of course), there may well be multiple cue send buses so that separate mixes can be made for different groups of musicians.

Punch-in/punch-out (insert) recording

Punch-in recording is a very important concept in post production, because it allows updating mixes without remixing whole reels. Punch-in recording relies on the ability of a mixer to achieve console settings that are identical to those used during a prior mixdown. This is assured by switching back and forth between the source (the mixdown from units produced by the console) and play back off the recording. This switching is given the possible names Direct–Film, Bus–Film or PEC–Direct as described in Chapter 5.

Once throwing the direct-film switch back and forth reveals no sonic difference, the mixer can punch in, that is, activate timed erase and record circuitry so that a new recording is begun seamlessly. Then the mixer can proceed to remix a portion of a reel and, coming once again to a place where he or she can match the console settings with the original, can *punch out*, yielding a newly mixed section without abrupt changes at the transition between the new and existing mixes. This process saves a great deal of time and money in post production and relies on the use of three-head machines to know what the prior recording was, and what the new recording will be, by activating the source/film switch.

Punch-in recording is also used in music studios

> working on multitrack recorders. Let us say that all tracks are all right, except one, which contains a wrong note. What is done is to punch in and out just before and after the note, having the musician play the right note at the right time. The musician is "cued" by listening to sound from the other tracks, and he or she performs continuously (because it is hard to play just one note at the right time). The engineer punches in and out at the right moments to substitute a new performance of the one note.

One difficulty occurs in both of these cases due to use of three-head machines. With separate record and playback heads, if one listens to playback and plays in sync with the music, the musician will be in sync with the sound at the playback head, but not at the record head where sync is needed. This problem was solved early in the history of multitrack recording when guitarist Les Paul devised a method to listen to tracks *played back* from the other channels of the *record* head for cueing purposes. Called *sel-sync, sel-rep*, or by other trade names, this method of recording ensures sync is maintained despite the three-head configuration. In TV sweetening it is commonplace to use machines in this mode, so that all the signals are always coming from or going to the record head, to prevent any sync errors.

Automation

Increasingly, mixes have gotten so complex and are under such time pressure that it is impossible to "perform" the mix by operating the controls in real time, even with more than one mixer doing the job. While some complex mixing can be done during premixing because of the multistage nature of film and video sound re-recording, nevertheless it is important to have sophisticated control over all of the premixes at the final mix because this is the stage at which the producer, director, and others become most involved.

There are several levels of automation possible. The most fundamental automation is over level control, because this is by far the most active part of the console typically. Several types of fader automation are available, basically breaking down between moving-fader automation and voltage-controlled amplifier automation. In moving-fader systems, the re-recording mixer performs moves on the faders during one pass of the film, and a computer memorizes the moves and reperforms them during subsequent passes. On these later passes, more faders can be brought into play and the computer continues to move already established faders. In this way, one person can do an extremely complicated mix, by adding more and more sophistication over time. Voltage-controlled amplifier automation systems also accomplish the same thing, but without the physically moving faders and thus are somewhat harder to update, because the physical position of the faders must be matched to an old setting in order to take over the adjustment task. The position of the faders is indicated by some kind of metering function of the console.

Usually fader automation is accompanied by automation of mutes, as described previously. A few consoles exist wherein the control surface is not directly connected to the circuits they are controlling, but rather operate as a user interface, sending all control functions as messages to a rack of equipment located in a machine room. These consoles may use "rotary shaft encoders" for their control functions, which means that the setting is no longer tied to the physical setting of the controls but rather to changes in the controls. In this system it is possible to take a snapshot of all of the controls and to restore the console to a precisely known former condition as long as a snapshot was taken. Full automation is possible, with all the control functions affected.

On fully automated digital consoles, it is possible to work on a program without committing it to being recorded, because the console will continue to reproduce all the fader moves, equalization changes, et cetera, in all subsequent passes through the material. Once the program is finalized, then it can be recorded to the medium.