8 Getting to Work

8.1 Sound for All Occasions
This book is intended to be useful to sound designers and technicians in theatre, film, and music production. It is also aimed at computer scientists who would like to collaborate with such artists and practitioners or design the next generation of sound hardware and software for them.

We've provided a lot of information in previous chapters moving through the concepts, applications, and underlying science of audio processing. In the end, we realize that most of our readers want to do something creative, whether that be designing and setting up the sound backdrop for a theatre performance, creating a soundtrack for film, producing a music album, or writing programs for innovative sound processing. It's doubtful and not even necessary that you mastered all the material in the previous chapters, but it serves as a foundation and a reference for your practical and artistic work. In this chapter, we pull things together by going into more detail on the artists' and practitioners' hands-on work and offering more suggestions for creative collaborations that would engage the programming and mathematical expertise of the computer scientist.

We considered dividing this chapter into Sound for Theatre, Studio Music Production, and Sound for Film or Video, since each reader may be particularly interested in one of these three areas. The division is problematic, however, because there is a great deal of overlap among the three. Microphone choice is a first step in all three areas. Recording in a studio is a central part of studio music production, but it is also done as sound is prepared for theatre or film. Live sound mixing is extremely important in theatre performances and music concerts, but an
equivalent software-based process is part of post-production of studio-recorded sound, whether it's destined for a music album, a theatre sound effect, or a film soundtrack.

Whatever you're particular focus in sound creation and editing, your workflow can be divided into steps: **pre-production**, when you design your project and choose your basic equipment; **production**, when you record and edit the sound; and **post-production**, when you deliver the sound, either live or on a permanent storage medium. Thus, a good way to begin this chapter is to give an overview of the workflow in sound creation for theatre, music production, and film based on the three main steps. We can then fill in details on the processes that are shared by the three areas, including microphone choice, studio recording, and mixing.

Before looking at the three main steps in sound creation for music, theatre, or film, we need to cover two activities common to all three areas – capturing sound, and mixing sound from different tracks.

### 8.2 Workflow in Studio Music Production

With more information about microphones and mixers, we're now ready to look more closely at the pre-production, production, and post-production steps involved when you work in different areas of sound and music.

Here's an overview of the workflow in studio music production:

- **Pre-Production**
  - Sound design (music arrangement and production aspects)
  - Analyzing your recording needs (Choosing microphone, hardware and software, recording environment, etc.)
  - Setting up

- **Production**
  - Recording

- **Post-Production**
  - Audio processing (Applying EQ, dynamics processing, special effects, stereo separation, etc.)
  - Mixing
  - Mastering

### 8.2.1 Pre-production

In the pre-production stage you are preparing everything you’ll need in order to complete the recording. Of course the first things you should do is compose the music you plan to record, hire the musicians, rent the studio, etc. Once you have all of these things in place you can begin preparing for the recording.

An important thing to consider is whether all the music will be recorded live or created ahead of time using MIDI sequencing and software synthesizers. For example if you’re just going to be recording singers against a MIDI sequenced backing track, you’ll need to make sure you have that backing track created ahead of time. At least enough that the singers will be able to sing on pitch and in time with the music. You can clean up the music later in post-production but it’s more difficult to fix a vocal recording that is off-pitch or off-tempo.

If you’ll be recording all the music in the studio using real instruments, you’ll need to decide whether to record everything at the same time or record each instrument separately. The advantage to recording everything at the same time is that the musicians can play together and feel their way through the song. That will almost always produce a better performance and
usually the musicians will prefer this. The downside to this method is that with all the musicians in the same room together playing, you’ll have a hard time getting good isolation between each instrument in the recording. For example, you’ll be picking up the sound of the drums on the vocalist’s microphone. In post-production if you want to make the vocals louder in the mix, the drums will get louder too.

8.2.2 Production

8.2.3 Post-production

Mastering is the process of adjusting the dynamics and frequency response of the final mix, preparing it for its final output form. The storage medium could be CD, DVD, disk drive for online access, or even an analog format. The term *mastering* comes from the idea of a master copy from which all other copies are made.

When you've completed the mixing process of a recording project, the next step is mastering the mixed-down audio. Mastering is the process of adjusting the dynamics and frequency response of a mix in order to optimize it for listening in various environments and prepare it for storage on the chosen medium. In some ways you could describe the mastering process as making the mix sound louder. When mixing a multitrack recording, one thing you watch for is clipped signals. Once the mix is completed, you may have a well-balanced mix but overall the mix sounds quieter than other mixes you hear on commercial recordings. What is typically happening is that you have one instrument in your mix that is a bit more dynamic than the others, and in order to keep the mix from clipping, you have to turn everything down because of this one instrument. One step in the mastering process is to use a multi-band compressor to address this problem.

A multi-band compressor is a set of compressors that compress a limited frequency band without affecting other frequency bands. A traditional compressor will attenuate the entire mix when one frequency exceeds the threshold. A multi-band compressor will attenuate an instrument that is dynamic in one frequency band without attenuating other frequency bands. This is often much more effective than using a simple EQ because the processing is only applied when needed, whereas an EQ will boost or cut a certain range of frequencies all the time. This allows you to let the less-dynamic frequencies take a more prominent role in the mix, resulting in the entire mix sounding louder. Figure 8.21 shows an example of a multi-band compressor.

8.3 Workflow in Sound for Film or Video

Producing sound for film or video involves essentially the same steps as studio music production. The main difference lies in the design and the post-production steps because the sound has to be synchronized with speech and visual action in the film.

- **Pre-Production**
  - Sound design
  - Analyzing your recording needs
  - Setting up
- **Production**
  - Recording
  - Foley sound effects creation
- **Post-Production**
8.3.1 Pre-Production
When you begin a recording project, the first decision you need to make is how you'll acquire the sound of whatever you plan to record. The goal is to find a balance between fidelity, isolation, cost, and convenience. It's very difficult to get all four of these things in every scenario, so be prepared to make some compromises. The solution that offers the best isolation will likely not be very convenient and may be fairly expensive. For example, if you want to record thunder without picking up the sounds of rain, dog barks, cars, birds, etc., you need to find an outdoor location that is far from roads, houses, and trees but also offers shelter from the rain for your equipment. Once you find that place, you need to predict when a thunderstorm will happen and get there in time to set up your equipment before the storm begins. Since this is not a very practical plan, you may have to make some compromises and be prepared to spend a long time recording storms in order to get a few moments where you manage to get the sound of thunder when nothing else is happening. Every recording situation will be different, but if you understand the options that are available, you can be better prepared to make a good decision.

8.3.2 Production
Production audio refers to the sound captured during the production process. The production process is when you actually shoot the video or film. At that time, there may be various sounds you're trying to capture. This could be the voices of the actors, environmental sounds, sounds of props or other scenic elements. When recording in a controlled sound stage studio, you can capture the production audio reasonably well with good quality, but when recording on location you will constantly be battling background noise. The challenge in either situation is to capture the sounds you need without capturing the sounds you don't need.

All the same rules apply in this situation as in other recording situations. You need good quality microphones and you need to get them as close as possible to the thing you're trying to record. This can be challenging in a production environment where high definition cameras are close up on actors and can see a great level of detail. Typically the microphone needs to be invisible or out of the camera shot. Actors can wear wireless lavaliere microphones as long as they can be hidden under some clothing. This will affect the quality of the sound being picked up by the microphone, but in the production environment compromises are a necessity. The primary emphasis is of course on capturing the things that would be impossible or expensive to change in post-production, like performances or action sequences. For example, if you don’t get the perfect performance from the actors, or the scenery falls down, it's very difficult to fix that without completely repeating the entire production process. If you don’t catch those problems until the post-production process, rebuilding that set and bringing back the cast and crew is very expensive. On the other hand, if the production audio is not captured with a high enough quality, the actor can be brought back in alone to re-record the audio without having to re-shoot the video. So in the production environment, the picture and the performance are the most important things. Capturing the production audio is ultimately of less-importance because it's easier to fix in post-production.

With this in mind, most production audio engineers are willing to make some compromises by putting microphones under the clothing. The other option is to mount a small
directional microphone to a long pole called a boom pole. Someone outside the camera shot holds the pole and he or she can get the microphone fairly close to the person or thing they're trying to capture without getting the microphone in the shot. Because re-shooting is so expensive, the most important job of the boom pole operator is to keep the microphone out of the shot. Picking up usable production audio is secondary.

The other challenge with production audio is synchronization. The audio is typically captured on a completely different recording media than the video. They are put back together in the post-production process. In the analog domain this can be a very tricky process. The early attempt at facilitating this synchronization was the clapboard slate. The slate has an area where you can write the information needed to identify this particular recording such as the name of the show, the scene number, and the take number. There is also a block of wood connected to the slate with a hinge. This block of wood can be raised and then lowered quickly onto the slate to make a loud clap sound. The person holding the slate would read out loud the information written on the slate while holding the slate in front of the camera and then drop the clapper. In post-production the slate info can be seen on the film and heard on the audio recording. This way you know that you have the correct audio recording with the correct video. The clap sound can be easily heard on the audio recording, and on the video you can easily see the moment that the clapper closes. The editor can line up the sound of the clap with the image of the clapper closing, and then everything after that is in sync. This simple and low-tech solution has proven to be quite effective and is still used in modern filmmaking along with other improvements in synchronization technology. In the analog domain, one synchronization problem was called drift. Because the analog equipment does not perform exactly the same every time, even though you line up the clap, slight variations in speed of the tape player or the film camera could cause the audio to slowly drift out of sync with the video. Digital audio synchronization doesn't have that problem.

A time code format called SMPTE has been developed to address the issue of synchronization. The format of SMPTE time code is described in Chapter 6. The idea behind time code synchronization is that the film has a built-in measuring system. There are 24 frames or still pictures every second in traditional motion picture film with each frame being easily identified. The problem is that on an audiotape there is no inherent way to know which part of audio goes with which frame of video. Part of the SMPTE time code specification included a method of encoding this time code into an audio signal that could be recorded on a separate track of audio on the tape recorder. This way, the entire audio recording is linked to each frame of the video. In the digital domain, this time code can be encoded into the video signal as well as the audio signal and the computer can keep everything in sync. The slate clapper has even been updated to display the current time code value to facilitate this synchronization in post-production.

8.3.3 Post-Production

After the production shooting is over, the post-production process begins. This is where the video is edited, special effects are added, and the soundtrack is completed. The sound track includes the production audio tracks along with any overdubs that are necessary to fix problems in the production audio, sound effects, and music. The work on music and sound effects is somewhat limited until the editing process is completed. The process of overdubbing the production audio is often called ADR. Depending on who you ask, this stands for Automated Dialog Replacement or Additional Dialogue Recording.
During this process, an actor is brought in to a recording studio, looks at the scene that was filmed during the production process, and listens to the performance they gave. The actor then attempts to recreate that performance vocally. This is typically done in small chunks in a loop so the actor has multiple attempts to get it right. He's trying to recreate not only the sound but also the speed of the original performance so that the new recording is synchronized with the movement of the lips on the screen. System clicks and streamers can be used to help the actor. Clicks (sometimes called beeps) are a rhythmic sound, like a metronome, that count down to a certain point when the actor needs to start or hit a particular word. Streamers are a visual reference that follows the same speed of the clicks. The streamer is a solid line across the screen that moves in time with the clicks so you can see when the important synchronization events occur. Clicks and streamers are also used in other post-production audio tasks for synchronizing sound effects and music during recording sessions.

For sound effects there are two basic strategies. With digital editing systems, one is simply to create and recording the various sound effects and then place them in the correct location digitally to ensure proper synchronization. Traditionally the process of adding sound effects involves a process called Foley sound, named after Jack Foley, who did the original work on this technique in the early days of silent films. Foley artists are a special breed of filmmaker who create all the sound effects for a film manually in a recording session. Foley stages are recording studios with all kinds of toys, floor surfaces, and other gadgets that make various sounds. The Foley artists go into the stage and watch the film while performing all the sounds required, the sounds ranging from footsteps and turning doorknobs to guns, rain, and other environmental sounds. The process is a lot of fun and some people build entire careers as Foley artists.

Composers for film will typically work first on developing a few musical themes that will be used throughout the film while the editing process is still happening. Once a full edit is completed the composer will begin to take the various themes that have been composed and create versions of various lengths to fit with the timing of the edited scenes. Sometimes this is done entirely with electronic instruments directly in the DAW with the video imported to the project file for a visual reference. Other times, a recording session is conducted where an orchestra is brought into a special recording studio called a scoring stage. The film is projected in the studio and the composer along with a conductor will perform the various musical passages for the film using clicks and streamers to help them synchronize the important moments in the music and the video.

With all the various audio elements recorded, and the video edited, the next step is mixing. With all these different sounds happening at once, it can be quite challenging to mix them together in a way that doesn’t result in complete chaos. One way of taming the mix is to use surround sound. Mixing the various elements to different loudspeakers can help each sound to be heard in the mix. For example, the voices are typically mixed to the center channel and music and sound effects are mixed to the four different surround channels. Loudness and dynamics are also an issue that gets close attention in the mixing process. In some cases you may need to meet a specific average loudness level over the course of the entire video. The mix engineer will typically create stems (similar to busses or groups) to help with the process. For example, there might be a vocal stem, a music stem, and a sound effects stem. These stems can then be manipulated for various delivery mediums. For example, the audio needs of the television broadcaster are going to be different from the needs of the person authoring the DVD for home distribution. The mix may need to be adjusted for each of these delivery media, and it
is much easier to do this using the stems than to return to the original multitrack source that may involve several hundred tracks.

8.4 Workflow in Sound for Theatre or Live Music Performance

Working with sound for theatre or music performance is different from the other two areas because of its "live" aspect. Even so, there is overlap in that some elements of the sound may be pre-recorded. The basic steps are these:

- **Pre-production**
  - Sound design
  - Evaluating the performance environment
  - Designing the sound delivery system
  - Setting up
- **Production**
  - Recording
- **Post-production/delivery**
  - Sound processing of recorded sound
  - Putting sound into an automatic sound playback system
  - Live sound processing and mixing

8.4.1 Pre-Production

8.4.1.1 Designing the Sound

When working in theatre sound, one role of the sound designer is to design all the sound content for the production. This can include music and sound effects for a play. For musical theatre, the music is typically composed and performed live by others but the sound designer is still responsible for ensuring that the music is delivered to the audience clearly and audibly. The process for designing the sound effects and music includes a creative process that is beyond the scope of this text, but we can describe here some of the audio processes for producing the content once you know what is needed for the show.

When designing sound effects, the first step is to acquire some source material to work with. Commercial sound effect libraries are available for purchase and there are some online sources for free sound effects, but the free sources are often of inconsistent quality. Sometimes you may need to go out and record your own source material. The goal here is not necessarily to find the exact sound you are looking for. Instead, you are trying to find source material that has some of the characteristics of the sound you are looking for. Then you can edit, mix and process the source material to achieve the exact sound you need.

There are a few common tools used to transform your source material into the exact sound you are looking for. One of the most useful is pitch shifting. Spaghetti noodles breaking can sound like a tree falling when pitched down an octave or two. When using pitch shift, you will get the most dramatic transformative results when using a pitch shifter that does not attempt to maintain the original timing. In other words, when a sound is pitched up, it should also speed up.

Video Tutorial: Creating A Sound Effect
Another strategy is to mix several sounds together in a way that creates an entirely new sound. If you can break down the sound you are looking for into descriptive layers, you can then find source material for each of those layers and mix them all together. For example, you would never be able to record the roar of a Tyrannosaurus Rex, but if you can describe the sound you’re looking for, perhaps as something between an elephant trumpeting and a lion roaring, you’re well on your way to finding that source material and creating that new, hybrid sound.

Sometimes reverb or EQ can help achieve the sound you are looking for. If you have a vocal recording that you want to sound like it is coming from an old answering machine, using an EQ to filter out the low and high frequencies but enhance the mid frequencies can mimic the sound of the small loudspeakers in those machines. Making something sound farther away can be accomplished by reducing the high frequencies with an EQ to mimic the effect of the high frequency loss over distance, and some reverb can mimic the effect of the sound reflecting from several surfaces during its long trip.

8.4.1.2 Designing the Sound Delivery System

In a live performance it is quite possible that when the performers on the stage create their sound, that sound does not arrive at the audience in a way that is loud enough or clear enough to understand. In most cases, a sound designer or sound engineer is hired to design a sound reinforcement system to address this problem. The basic process is to use microphones near the performers to pick up whatever sound they are making and then play that sound out of loudspeakers that are pointed to the audience that can distribute that sound louder and clearer than the natural sound from the performer.

There are several things to consider when designing and operating a sound reinforcement system:

- You need loudspeakers that can faithfully generate a loud enough sound.
- You need microphones that can pick up the source sound as faithfully as possible without getting in the way.
- The loudspeakers need to be positioned in a way that will direct the sound to the listeners without sending too much sound to the walls or back to the microphones. This can impact intelligibility and gain.
- Ideally the sound system will deliver a similar listening experience to all the listeners regardless of where they sit.

Many of these considerations can be predicted prior to purchasing and installing the sound equipment. Using mathematical predictions, you can alter the plan before you have to spend any money on equipment. Once the equipment is installed, the system can be analyzed for these considerations and adjusted in order to improve the performance of the system. These adjustments can include repositioning microphones and loudspeakers to improve gain and frequency response, replacing equipment with something else that will perform better, and adjusting the settings on processing equipment such as equalizers, compressors, crossovers, and power amplifiers.

Most loudspeakers have a certain amount of directivity. Loudspeaker directivity is typically described in terms of the 6 dB down point – a horizontal and vertical angle off-axis corresponding to the location where the sound is reduced by 6 dB. The 6 dB down point is significant because, as a general rule of thumb, you want all of the audience to hear the sound within at most a 6 dB range. In other words, the seat on the end of the aisle shouldn’t sound more than 6 dB quieter than the seat in the middle of the row, or anywhere else in the audience.
The issue of loudspeaker directivity is complicated by the fact that loudspeakers naturally have a different directivity for each frequency. A single circular loudspeaker driver is more directional as the frequency increases because the loudspeaker diameter gets larger relative to the wavelength of the frequency. This high-frequency directivity effect is illustrated in Figure 8.1. Each of the six plots in the figure represents a different frequency produced by the same circular loudspeaker driver. Frequencies having a wavelength that is longer than the diameter of the loudspeaker are dispersed very widely. Once the frequency has a wavelength that is equal to the diameter of the loudspeaker, the loudspeaker begins to exercise some directional control over the sound. This directivity gets narrower as the frequency increases and the wavelength decreases.

![Figure 8.1 Directivity of circular radiators, diagrams created from actual measured sound](image)

This varying directivity per frequency for a single loudspeaker driver partially explains why most full-range loudspeakers have multiple drivers. The problem is not that a single loudspeaker can’t produce the entire audible spectrum. Any set of headphones uses a single driver for the entire spectrum. The problem with using one loudspeaker driver for the entire spectrum is that you can’t distribute all the frequencies uniformly across the listening area. The listeners sitting right in front of the loudspeaker will hear everything fine, but the listeners sitting to the side of the loudspeaker will hear the high frequencies much quieter than the low frequencies. If you add a second loudspeaker driver, considerably smaller than the first, and use a crossover to direct the high frequencies to the small driver and the low frequencies to the large driver, you can achieve a much more uniform directional dispersion as shown in Figure 8.2. In this case, the larger of the two drivers is 5” in diameter and the smaller driver is 1” in diameter.
Wavelengths corresponding to frequencies of 500 Hz and 1000 Hz have larger wavelengths than 5", so they are fairly omnidirectional. The reason that frequencies of 2000 Hz and above have consistent directivity is that the frequencies are distributed to the two loudspeaker drivers in a way that keeps the relationship consistent between the wavelength and the diameter of the driver.

There are many other strategies used by loudspeaker designers to get consistent pattern control, but none are able to avoid this simple concept that the size of the loudspeaker has a direct correlation to the range of frequencies over which it has directional control. You can simply look at any loudspeaker and easily determine the lowest possible directional frequency based on the size of the loudspeaker.

Understanding how a loudspeaker exercises directional control over the sound it radiates can also help you decide where to install and aim a loudspeaker to provide consistent sound levels across the area of your audience. Using the inverse square law in conjunction with the loudspeaker directivity information, you can find a solution that provides even sound coverage over a large audience area using only a single loudspeaker. Consider the example 1000 Hz vertical polar plot for a loudspeaker shown in Figure 8.3. If you were going to use that loudspeaker in the theatre shown in Figure 8.4, where would you aim the loudspeaker?
Most beginning sound system designers will choose to aim the loudspeaker at seat B thinking that it will keep the entire audience as close as possible to the on-axis point of the loudspeaker. To test the idea, we need to calculate the dB loss over distance using the inverse square law for each seat and then subtract any additional dB loss incurred by going off-axis from the loudspeaker. Seat B is directly on axis with the loudspeaker, and according to the polar plot there is approximately a loss of $-2$ dB at 0 degrees. Seat A is 33 degrees down from the on-axis point of the loudspeaker, corresponding to 327 degrees on the polar plot, which shows an approximate loss of $-3$ dB. Seat C is 14 degrees off axis from the loudspeaker, resulting in a loss
of –6 dB according to the polar plot. Assuming that the loudspeaker is outputting 100 dB SPL at 1 meter (3.28 feet), we can calculate the dBSPL level for each seat as shown in Equation 8.1.

<table>
<thead>
<tr>
<th></th>
<th>Seat A dB SPL</th>
<th>Seat B dB SPL</th>
<th>Seat C dB SPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100 dB + (20 ( \log_{10} \frac{3.28'}{33.17'} )) – 3 dB</td>
<td>100 dB + (20 ( \log_{10} 0.1 )) – 3 dB</td>
<td>100 dB + (20 ( \log_{10} 0.04 )) – 6 dB</td>
</tr>
<tr>
<td></td>
<td>77 dB SPL</td>
<td>74.25 dB SPL</td>
<td>66.55 dB SPL</td>
</tr>
</tbody>
</table>

Equation 8.1 Calculating dBSPL of a given loudspeaker aimed on-axis with seat B

In this case the loudest seat is seat A at 77 dBSPL, and seat C is the quietest at 66.55 dBSPL, with a 10.45 dB difference. As discussed, we want all the audience locations to be within a 6 dB range. But before we throw this loudspeaker away and try to find one that works better, let’s take a moment to examine the reasons why we have such a poor result. The reason seat C is so much quieter than the other seats is that it is the farthest away from the loudspeaker and is receiving the largest reduction due to directivity. By comparison, A is the closest to the loudspeaker, resulting in the lowest loss over distance and only a 3 dB reduction due to directivity. To even this out let’s try having the farthest seat away be the seat with the least directivity loss, and the closest seat to the loudspeaker have the most directivity loss.

The angle with the least directivity loss is around 350 degrees, so if we aim the loudspeaker so that seat C lines up with that 350 degree point, that seat will have no directivity loss. With that aim point, Seat B will then have a directivity loss of –3 dB, and Seat A will have a directivity loss of –10 dB. Now we can re-calculate the dBSPL for each seat as shown in Equation 8.2.

<table>
<thead>
<tr>
<th></th>
<th>Seat A dB SPL</th>
<th>Seat B dB SPL</th>
<th>Seat C dB SPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100 dB + (20 ( \log_{10} \frac{3.28'}{33.17'} )) – 10 dB</td>
<td>100 dB + (20 ( \log_{10} 0.1 )) – 10 dB</td>
<td>100 dB + (20 ( \log_{10} 0.04 )) – 6 dB</td>
</tr>
<tr>
<td></td>
<td>70 dB SPL</td>
<td>70 dB SPL</td>
<td>66.55 dB SPL</td>
</tr>
</tbody>
</table>
In this case our loudest seat is seat B at 73.25 dB SPL and our quietest seat is seat A at 70 dB SPL, for a difference of 3.25 dB. Compared with the previous difference of 10.55 dB, we now have a much more even distribution of sound to the point where most listeners will hardly notice the difference. Before we fully commit to this plan, we have to test these angles at several different frequencies, but this example serves to illustrate an important rule of thumb when aiming loudspeakers. In most cases, the best course of action is to aim the loudspeaker at the farthest seat, and have the closest seat be the farthest off-axis to the loudspeaker. This way, as you move from the closest seat to the farthest seat, while you're losing dB over the extra distance you're also gaining dB by moving more directly on-axis with the loudspeaker.

Fortunately there are software tools that can help you determine the best loudspeakers to use and the best way to deploy them in your space. These tools range in price from free solutions such as MAPP Online Pro from Meyer Sound shown in Figure 8.5 to relatively expensive commercial products like EASE from the Ahnert Feistel Media Group, shown in Figure 8.6. These programs allow you to create a 2D or 3D drawing of the room and place virtual loudspeakers in the drawing to see how they disperse the sound in the room. The virtual loudspeaker files come in several formats. The most commonly found format is the EASE format. Fortunately, while EASE is the most expensive and comprehensive solution out there, most other programs have the ability to import EASE loudspeaker files. Another format is the Common Loudspeaker Format (CLF). CLF files use an open format, and many manufacturers are starting to publish their loudspeaker data in this format. Information on loudspeaker modeling software that uses the CLF format can be found at the website for the Common Loudspeaker Format Group [http://www.clfgroup.org](http://www.clfgroup.org).

---

**Aside:** EASE was developed by German engineers ADA (Acoustic Design Ahnert) in 1990 and introduced at the 88th AES Convention. That’s also the same year that Microsoft announced Windows 3.0.
Figure 8.5 MAPP Online Pro software from Meyer Sound
8.4.1.3 Checking Frequencies

With practice you can hear and identify frequencies, but sometimes being able to see the frequencies can help you to diagnose and solve problems. This is especially true when you’re setting up the sound system for a live performance in a theatre.

In previous chapters, we showed examples of frequency response graphs. You might wonder why the word “response” is used for the graph. The implication is that frequencies are responding to something. In practice, the primary concern with sound frequencies is how they respond to the environment they’re in, or as we’ll see later, how they respond to a filter applied to them. Suppose you’re generating certain frequencies in a sound you’re creating, say for a theater performance in an auditorium. How are these frequencies heard by the audience once they pass through loudspeakers and travel through space encountering obstructions, varying air temperatures, and so forth? Is each frequency arriving at the audience’s ears at the desired amplitude? Are certain frequencies too loud or too quiet? If the high frequencies are too quiet, you could sacrifice the brightness or clarity in the sound. Low frequencies that are too quiet could result in muffled voices. There are no clear guidelines on what the “right” frequency response is because it usually boils down to personal preference, artistic considerations,

\[\text{Aside: Acoustic systems} \text{ are systems in which the sounds produced depend on the shape and material of the sound-producing instruments. Electroacoustic systems produce sound through electronic technology such as amplifiers and loudspeakers.}\]
performance styles, and so forth. In any case, before you can decide if you have a problem, the first step is to analyze the frequency response in your environment.

A sound analysis system is one of the fundamental tools for ensuring that frequencies are being received at proper levels. The system consists of a computer running the analysis software, an audio interface with inputs and outputs, and a special analysis microphone. An analysis microphone is different from a traditional recording microphone. Most recording microphones have a varying response or sensitivity at different frequencies across the spectrum. This is often a desired result of their manufacturing and design, and part of what gives each microphone its unique sound. For analyzing acoustic or electroacoustic systems, you need a microphone that measures all frequencies equally. This is often referred to as having a flat response. In addition, most microphones are directional. They pick up sound better in the front than in the back. A good analysis microphone should be omni-directional so it can pick up the sound coming at it from all directions. Figure 8.7 shows a popular analysis microphone from Earthworks.

![Figure 8.7 Earthworks M30 analysis microphone](image)

One application of a sound analysis system is to measure the frequency response of a loudspeaker when heard from a listening location. A common way to do this is to use an analysis system to generate a sine wave sweep through a range of frequencies that you define. A sine wave sweep is a sound that begins at a low frequency sine wave and smoothly moves up in frequency to some given high frequency limit. The sweep lasts a few seconds or less. The audio output of your analysis system that generates the sine sweep is sent by a direct cable connection to the loudspeaker. You then place your analysis microphone at the listening location you want to analyze. The microphone listens for each frequency as it is radiated by the loudspeaker and stores the information in a file called an impulse response. The impulse response is a graph of the sound wave with time on the x-axis and the amplitude of the sound wave on the y-axis. This same information can be displayed in a frequency response graph, which has frequencies on the x-axis and the amplitude of each frequency on the y-axis. (In Chapter 7, we’ll explain the mathematics that transforms the impulse response graph to the frequency response graph, and vice versa.) Figure 8.8 shows an example frequency response graph created by the procedure just described.
Figure 8.8 Frequency response graph showing a low frequency boost

The sine wave sweep generates all frequencies equally, so the frequency response graph, in the ideal, should be flat. But notice in the graph that the frequencies between 30 Hz and 500 Hz are 6 to 10 dB louder than the rest. It’s up to you to decide if this is a problem you want to solve. Keep in mind that the goal isn’t necessarily to make the frequency response graph be a straight line, indicating all frequencies are of equal amplitude. The goal is to make the right kind of sound. Before you can decide what to do, you need to determine why the frequency response sounds like this. There are many possible reasons. It could be that you’re too far off-axis from the loudspeaker generating the sound. That’s not a problem you can really solve when you’re analyzing a listening space for a large audience, since not everyone can sit in the prime location. You could move the analysis microphone so that you’re on-axis with the loudspeaker, but you can’t fix the off-axis frequency response for the loudspeaker itself. In the example shown in Figure 8.8, the loudspeaker system that is generating the sound uses two sets of sound radiators. One set of loudspeakers generates the frequencies above 500 Hz. The other set generates the frequencies below 500 Hz. Given that information, you could conclude that the low-frequency loudspeakers are simply louder than the high frequency ones. If this is causing a sound that you don’t want, you could fix it by reducing the level of the low-frequency loudspeakers.

Figure 8.9 shows the result of this correction. The grey line shows the original frequency response and the black line shows the frequency response after reducing the amplitude of the low-frequency loudspeakers by 6 dB.
The previous example gives you a sketch of how a sound analysis system might be used. You place yourself in a chosen position in a room where sound is to be performed or played, generate sound that is played through loudspeakers, and then measure the sound as it is received at your chosen position. The frequencies that are actually detected may not be precisely the frequency components of the original sound that was generated or played. By looking at the difference between what you played and what you are able to measure, you can analyze the frequency response of your loudspeakers, the acoustics of your room, or a combination of the two. The frequencies that are measured by the sound analysis system are dependent not only on the sound originally produced, but also on the loudspeakers’ types and positions, the location of the listener in the room, and the acoustics of the room. Thus, in addition to measuring the frequency response of your loudspeakers, the sound analysis system can help you to determine if different locations in the room vary significantly in their frequency response, leaving it to you to decide if this is a problem and what factor might be the source.

There are many choices for analysis software, but they all fall into two main categories: signal dependent and signal independent. Signal dependent sound analysis systems rely on a known stimulus signal that the software generates, as was illustrated in the previous example, where the signal is a sweep through all frequencies in a given range. The software sends the signal into the room and then compares what the analysis microphone picks up against what the analysis system knows it sent out. The advantage to a signal dependent system is that it’s easy to use, and with it you can get a good general picture of how frequencies will sound in a given acoustic space with certain loudspeakers. You also can save the frequency response graphs to refer to and analyze later. The disadvantage to a signal dependent analysis system is that it uses only artificially-generated signals like sine sweeps, not real music or performances.

Figure 8.10 shows a screenshot from FuzzMeasure Pro. FuzzMeasure Pro is a signal dependent analysis program that runs on the Mac operating system. The frequency response is on the top and the impulse response is at the bottom. These graphs show that what is heard differs from the sound that was originally generated, where all frequencies were played at the same amplitude.
Signal independent analysis systems use a stimulus sound signal that you provide. This means that you can use recorded music, voice, sound effects, or even live performances in the acoustic space you want to analyze. In contrast to systems like FuzzMeasure, which know the precise sweep of frequencies they’re generating, signal independent systems must be given a direct copy of the sound being played so that the original sound can be compared with the sound that passes through the air and is received by the analysis microphone. This is accomplished by taking the original sound and sending one copy of it to the loudspeakers while a second copy is sent directly, via cable, to the sound analysis software. The software presumably is running on a computer that has a sound card attached with two sound inputs. One of the inputs is the analysis microphone and one is a direct feed from the sound source. The software compares the two signals in real time – as the music or sound is played – and tells you what is different about them. The advantage of the signal independent system is that it can analyze “real” sound as it is being played or performed. However, real sound has frequency components that constantly change, as we can tell from the constantly changing pitches that we hear. Thus, there isn’t one fixed frequency response graph that gives you a picture of how your loudspeakers and room are dealing with the frequencies of the sound. The graph changes dynamically over the entire time that the sound is played. For this reason, you can’t simply save one graph and carry it off with you for analysis. Instead, your analysis consists of observing the constantly-changing frequency response graph in real time, as the sound is played. If you wanted to save a single frequency response graph, you’d have to do what we did to generate Figure 8.11 – that is, get a “screen capture” of the frequency response graph at a specific moment in time – and the information you
Figure 8.11 was produced from a popular signal independent analysis program called Smaart Live, which runs on Windows and Mac operating systems. The graph shows the difference, in decibels, between the amplitudes of the frequencies played vs. those received by the analysis microphone. Because this is only a snapshot in time, coupled with the fact that noise is measured as well, it isn’t very informative to look at just one graph like this. Being able to glean useful information from a signal independent sound analysis system comes from experience in working with real sound – learning how to compare what you want, what you see, what you understand is going on mathematically, and – most importantly – what you hear.

8.4.1.4 System Documentation

Once you've decided on a loudspeaker system that distributes the sound the way you want, you need to begin the process of designing the systems that capture the sound of the performance and feed it into the loudspeaker system. Typically this involves creating a set of drawings that give you the opportunity to think through the entire sound system and explain to others – installers, contractors, or operators, for example – how the system will function.
The first diagram you're likely create is the System Diagram. This is similar in function to an electrical circuit diagram, showing you which parts are used and how they're all wired up. The sound system diagram shows how all the components of a sound system connect together in the audio signal chain, starting from the microphones and other input devices all the way through to the loudspeakers that reproduce that sound. These diagrams can be created digitally with vector drawing programs such as AutoCAD and VectorWorks or diagramming programs such as Visio and OmniGraffle.

The United States Institute for Theatre Technology has published some guidelines for creating system diagrams. The most common symbol or block used in system diagrams is the generic device block shown in Figure 8.12. The EQUIPMENT TYPE label should be replaced with a descriptive term such a CD PLAYER or MIXING CONSOLE. You can also specify the exact make and model of the equipment in the label above the block.

![Figure 8.12 A generic device block for system diagrams](image)

There are also symbols to represent microphones, power amplifiers, and loudspeakers. You can connect all the various symbols to represent an entire sound system. Figure 8.13 shows a very small sound system, and Figure 8.14 shows a full system diagram for a small musical theatre production.

![Figure 8.13 A small system diagram](image)
While the system diagram shows the basic signal flow for the entire sound system, there is a lot of detail missing about the specific interconnections between devices in the system. This is where a patch plot can be helpful. A patch plot is essentially a spreadsheet that shows every connection point in the sound system. You should be able to use the patch plot to determine which and how many cables you’ll need for the sound system. It can also be a useful tool in troubleshooting a sound system that isn’t behaving properly. The majority of the time when things go wrong with your sound system or something isn’t working, it's because it isn’t connected properly or one of the cables has been damaged. A good patch plot can help you find the problem by showing you where all the connections are located in the signal path. There is no industry standard for creating a patch plot, but the rule of thumb is to err on the side of too much information. You want every possible detail about every audio connection made in the sound system. Sometimes color coding can help make the patch plot easier to understand. Figure 8.15 shows an example patch plot for the sound system in Figure 8.13.
8.4.1.5 System Optimization

Once you have the sound system installed and everything is functioning, the system will need to be optimized. System optimization is a process of tuning and adjusting the various components of the sound system so that

- they're operating at the proper volume levels,
- the frequency response of the sound system is consistent and desirable,
- destructive interactions between system components and the acoustical environment have been minimized, and
- the timing of the various system components has been adjusted so the audience hears the sounds at the right time.

The first optimization you should perform is optimizing the gain structure of the sound system. When working with sound systems in either a live performance or recording situation, gain structure is a big concern. In a live performance situation, the goal is to amplify sound. In order to achieve the highest potential for loudness, you need to get each device in your system operating at the highest level possible so you don’t lose any volume while traveling through the system. In a recording situation, you're primarily concerned with signal-to-noise ratio. In both of these cases, good gain structure is the solution.

In order to understand gain structure, you first need to understand that all sound equipment makes noise. All sound devices also contain amplifiers. What you want to do is amplify the sound without amplifying the noise. In a sound system with good gain structure, every device is receiving and sending sound at the highest level possible without clipping. This solves the noise problem because the level of the sound is significantly higher than the level of the noise. Each piece of equipment in your sound system has a maximum input level and a maximum output level as specified by the manufacturer. In a system with good gain structure, these input and output levels are lined up in such a way that the entire system is operating at a unity gain. In other words, −10 dB on the input meter of your mixing console is also −10 dB on the input meter of your compressor, power amplifier, etc.

Lining up the gain for each device involves lining up the clip points. You can do this by starting with the first device in your signal chain – typically a microphone or some sort of playback device. It is easier to set up gain structure using a playback source because you can control the output volume. Start by playing something on the CD, synthesizer, computer, iPod or whatever your playback device is in a way that outputs the highest volume possible. This is usually done with either normalized pink noise or a normalized sine wave. Turn up the gain preamplifier on the mixing console or sound card input so that the level coming from the playback source clips the input. Then back off the gain until that sound is just below clipping. If you're recording this sound, your gain structure is now complete. Just repeat this process for each input. If it's a live performer on a microphone, ask them to perform at the highest volume they expect to generate and adjust the input gain accordingly.
If you're in a live situation, the mixing console will likely feed its sound into another device such as a processor or power amplifier. With the normalized audio from your playback source still running, adjust the output level of the mixing console so it's also just below clipping. Then adjust the input level of the next device in the signal chain so that it's receiving this signal at just below its clipping point. Repeat this process until you've adjusted every input and output in your sound system. At this point, everything should clip at the same time. If you increase the level of the playback source or input preamplifier on the mixing console, you should see every meter in your system register a clipped signal. If you've done this correctly, you should now have plenty of sound coming from your sound system without any hiss or other noise. If the sound system is too loud, simply turn down the last device in the signal chain. Usually this is the power amplifier.

Setting up proper gain structure in a sound system is fairly simple once you're familiar with the process. The Max demo on gain structure associated with this section gives you an opportunity to practice the technique. Then you should be ready to line up the gain for your own systems.

Once you have the gain structure optimized, the next thing you need to do is try to minimize destructive interactions between loudspeakers. One reason that loudspeaker directivity is important is due to the potential for multiple loudspeakers to interact destructively if their coverage overlaps in physical space. Most loudspeakers can exercise some directional control over frequencies higher than 1 kHz, but frequencies lower than 1 kHz tend to be fairly omnidirectional, which means they will more easily run into each other in the air. The basic strategy to avoid destructive interactions is to adjust the angle between two loudspeakers so their coverage zone intersects at the same dBSPL, and at the point in the coverage pattern where they are 6 dB quieter than the on-axis level, as shown in Figure 8.16. This overlap point is the only place where the two loudspeakers combine at the same level. If you can pull that off, you can then adjust the timing of the loudspeakers so they’re perfectly in phase at that overlap point. Destructive interaction is eliminated because the waves reinforce each other, creating a 6 dB boost that eliminates the dip in sound level at high frequencies. The result is that there is even sound across the covered area. The small number of listeners who happen to be sitting in an area of overlap between two loudspeakers will effectively be covered by a virtual coherent loudspeaker.

When you move away from that perfect overlap point, one loudspeaker gets louder as you move closer to it, while the other gets quieter as you move farther away. This is handy for two reasons. First, the overall combined level should remain pretty consistent at any angle as you move through the perfect overlap point. Second, for any angle outside of that perfect overlap point, while the timing relationship between the two loudspeaker arrivals begins to differ, the loudspeakers also differ more and more in level. As pure comb filtering requires both of the interacting signals to be at the same amplitude, the level difference greatly reduces the effect of the comb filtering introduced by the shift in timing. The place where the sound from the two loudspeakers arrives at the same amplitude and comb filters the most is at center of the overlap, but this is the place where we aligned the timing perfectly to prevent comb filtering in the first place. With this technique, not only do you get the wider coverage that comes with multiple loudspeakers, but you also get to avoid the comb filtering!
Single loudspeaker with 90° horizontal coverage between -6dB downpoints

Two loudspeakers with 90° horizontal coverage between -6dB downpoints that are rotated so the -6dB downpoints intersect creating a combined coherent coverage of 180°

Figure 8.16 Minimizing comb filtering between two loudspeakers
What about the low frequencies in this example? Well, they’re going to run into each other at similar amplitudes all around the room because they’re more omnidirectional than the high frequencies. However, they also have longer wavelengths, which means they require much larger offsets in time to cause destructive interaction. Consequently, they largely reinforce each other, giving an overall low frequency boost. Sometimes this free bass boost sounds good. If not, you can easily fix it with a system EQ adjustment by adding a low shelf filter that reduces the low frequencies by a certain amount to flatten out the frequency response of the system. This process is demonstrated in our video on loudspeaker interaction.

You should work with your loudspeakers in smaller groups, sometimes called systems. A center cluster of loudspeakers being used to cover the entire listening area from a single point source would be considered a system. You need to work with all the loudspeakers in that cluster to ensure they are working well together. A row of front fill loudspeakers at the edge of the stage being used to cover the front few rows will also need to be optimized as an individual system.

Once you have each loudspeaker system optimized, you need to work with all the systems together to ensure they don’t destructively interact with each other. This typically involves manipulating the timing of each system. There are two main strategies for time aligning loudspeaker systems. You can line the system up for coherence, or you can line the system up for precedence imaging. The coherence strategy involves working with each loudspeaker system to ensure that their coverage areas are as isolated as possible. This process is very similar to the process we described above for aligning the splay angles of two loudspeakers. In this case, you're doing the same thing for two loudspeaker systems. If you can line up two different systems so that the 6 dB down point of each system lands in the same point in space, you can then apply delay to the system arriving first so that both systems arrive at the same time, causing a perfect reinforcement. If you can pull this off for the entire sound system and the entire listening area, the listeners will effectively be listening to a single, giant loudspeaker with optimal coherence.

The natural propagation of sound in an acoustic space is inherently not very coherent due to the reflection and absorption of sound, resulting in destructive and constructive interactions that vary across the listening area. This lack of natural coherence is often the reason that a sound reinforcement system is installed in the first place. A sound system that has been optimized for coherence has the characteristic of sounding very clear and very consistent across the listening area. These can be very desirable qualities in a sound system where clarity and intelligibility are important. The downside to this optimization strategy is that it sometimes does not sound very natural. This is because with coherence optimized sound systems, the direct sound from the original source (i.e. a singer/performer on stage) has typically little to no impact on the audience, and so the audience perceives the sound as coming directly from the loudspeakers. If you’re close enough to the stage and the singer, and the loudspeakers are way off to the side or far overhead, it can be strange to see the actual source yet hear the sound come from somewhere else. In an arena or stadium setting, or at a rock concert where you likely wouldn’t hear much direct sound in the first place, this isn’t as big a problem. Sound designers are sometimes willing to accept a slightly unnatural sound if it means that they can solve the clarity and intelligibility problems that occur in the acoustic space.
Optimizing the sound system for precedence imaging is completely opposite to the coherence strategy. In this case, the goal is to increase the clarity and loudness of the sound system while maintaining a natural sound as much as possible. In other words, you want the audience to be able to hear and understand everything in the performance but you want them to think that what they are hearing is coming naturally from the performer instead of coming from loudspeakers in a sound system. In a precedence imaging sound system, each loudspeaker system behaves like an early reflection in an acoustic space. For this strategy to work, you want to maximize the overlap between the various loudspeaker systems. Each listener should be able to hear two or three loudspeaker systems from a single seat. The danger here is that these overlapping loudspeaker systems can easily comb filter in a way that will make the sound unpleasant or completely unintelligible. Using the precedence effect described in Chapter 4, you can manipulate the delay of each loudspeaker system so they arrive at the listener at least five milliseconds apart but no more than 30 milliseconds apart. The signals still comb filter, but in a way that our hearing system naturally compensates for. Once all of the loudspeakers are lined up, you’ll also want to delay the entire sound system back to the performer position on stage. As long as the natural sound from the performer arrives first, followed by a succession of similar sounds from the various loudspeaker systems each within this precedence timing window, you can get an increased volume and clarity as perceived by the listener while still maintaining the effect of a natural acoustic sound. If that natural sound is a priority, you can achieve acceptable results with this method, but you will sacrifice some of the additional clarity and intelligibility that comes with a coherent sound system.

Both of these optimization strategies are valid, and you'll need to evaluate your situation in each case to decide which kind of optimized system best addresses the priorities of your situation. In either case, you need some sort of system processor to perform the EQ and delay functions for the loudspeaker systems. These processors usually take the form of a dedicated digital signal-processing unit with multiple audio inputs and outputs. These system processors typically require a separate computer for programming, but once the system has been programmed, the units perform quite reliably without any external control. Figure 8.17 shows an example of a programming interface for a system processor.

Aside: While your loudspeakers might sit still for the whole show, the performers usually don't. Out Board's TiMax tracker and soundhub delay matrix system use radar technology to track actors and performers around a stage in three dimensions, automating and adjusting the delay times to maintain precedence and deliver natural, realistic sound throughout the performance.
8.4.2 Production

Whether an accomplished performer, a classically trained musician, a budding new composer, or music enthusiast who wants to bring their work to a professional level, there is no doubt that digital audio and computer tools have and will continue to change the way music is created. As music tools have grown in quality, flexibility, and accessibility, it seems as almost anyone can begin composing and producing music should they desire. Certainly, the more musical knowledge and ability you have the better. Yet even those with the right skillsets but perhaps without the available resources, or knowledge of those resources, now have more tools on hand then ever before to create high quality compositions.

Traditionally, composing is a process of conceiving of melodies and harmonies and putting them down on paper, called a **musical score**. This way of writing music is certainly still widely used, and many new digital tools have been created to help streamline and improve the scoring process. We talk a little more about these tools, such as Finale, Sibelius, or the open source MuseScore, back in Chapter XXX. These written musical scores are then often played by live musicians either as part of a performance, or during a recording session. Nowadays, these scoring software will also let you play back the score and even generate digital audio files using software instruments, either with general MIDI sounds or via more powerful and realistic sound samples. Most software also lets you export your musical score as individual MIDI tracks, which can be then imported, arranged, and mixed in your DAW of choice. The musical quality of these auto generated audio files is largely dependent on the quality of the triggered audio samples themselves, as well as the quality and detail of the scoring performance data. While the amount of musical control and intuitiveness within scoring software continues to improve, you can’t really expect the software to interpret “allegro con moto,” or even a crescendo or fermata the way an actual musician would. While the overall result may vary in quality, the musicality generally associated with scoring generated music is very “canned”, and mechanical sounding. Still, these audio files are often useful as scratch or temp music in a production or film, and can give a decent sense of the composition to collaborators, musicians, and audiences until the final score is ready to be produced. There are also a number of ways to tweak and improve this computer generated music, which if done well can achieve a surprisingly realistic result.
Before we get into these digital composition techniques, let’s make it clear that this knowledge is not intended to bear any opinion on the ethical considerations for employing professional live musicians, or devalue the work of classically trained musicians. There are a number of reasons using computer generated music may be best suited for a project, be it time, budget, resources, or style. The intention of this material is purely to show the functionality of these digital music tools and how to get the most out of them.

As mentioned, the quality of the sound samples plays directly into the quality of the computer generated music. What’s just as important to achieving the most realistic result, is the power and flexibility of the sample playback system. This could be the firmware on a hardware device, a standalone software application, or a software plug-in that runs within your DAW of choice. A basic sampler plays back a specific sound according to a received trigger such as pressing a key on your MIDI keyboard, or receiving note data from a music scoring application. These samplers may not even play a unique sound for each different key, often using manipulating one sample file to produce multiple notes. More information on working with samplers can be found in Chapter 6. More powerful samplers may allow multisampling, where multiple variations of a sample or note will be played back, for instance, depending on how hard/loud a note is played (note velocity). These samplers may also utilize other performance parameters and data to manipulate the sound for a more dynamic, realistic feel. For instance, the mod wheel could be linked to an LFO that imparts a controllable vibrato characteristic to the sound sample. While these sampler features greatly improve the realism of sample playback, the most powerful samplers go far beyond those.

Another potent feature provided by some samplers is a round robin feature. Even when playing two of the same notes one after another on an instrument, although the instrument is the same, the pitch is the same, and the force the note was struck or plucked is the same if not very close, the note will never sound exactly the same way twice. With round robin, when a note is triggered, the sampler will automatically cycle playback from among a set of similar samples, simulating that effect. In order to have round robin capability, you’ll of course need to have multiple takes of the audio samples for every note, and for multi-sampled sounds that means multiple takes for every velocity layer sample as well. From a memory and file size standpoint, this feature directly multiplies the size and system use of your sampler patch. The number of round robin sound samples included will vary, though you’ll need at least two, and you probably won’t find more than 4 or 5 at most. Some sampler systems instead use subtle processing to vary the way a sample sounds each time it is played back, simulating the round robin effect without the need for additional samples, though this may not quite achieve the same realism.

Another feature of high-end sampler instruments and sampler systems is multiple articulations. For example guitar has a different timbre depending if it’s played with a pick, the finger, if it’s palm muted, hammered on, etc. Classical stringed instruments have even more articulations than do guitars. Rather than having a separate sampler instrument for each articulation, all of these articulations are layered into one instrument, yet maintain individual control. Typically the sampler will have a number of keyswitches that will switch between the articulations. These keyswitches are often keys on the keyboard that are perhaps above or below the instrument’s musical range, so rather than play back a note, when pressed they provide an easy way to select the instrument articulation while playing. An example of keyswitches on a sampler can be seen in Figure 8.18.
Some sampler systems even have intelligent playback that will switch articulations automatically depending on how notes overlap or are played in relation to each other. In order to get the most realistic sound and range from an instrument, the more articulations the better. However, knowing how and when to employ those articulations is often limited by your familiarity with the particular instrument, so musical experience still plays an important role here.

![Sampler instrument with Keyswitches shown in blue](image)

**Figure 8.18 Sampler instrument with Keyswitches shown in blue**

As you can see, a lot of what results in more powerful, realistic computer generated music is not only having quality, but also quantity and diversity of content to play back. The amount of samples these virtual instruments require demands powerful and intelligent sampler playback systems, not to mention the computer hardware specs to support them. As an example of how extensive a well-crafted virtual instrument can be, one company’s upright bass instrument alone contains over 21,000 samples. For optimum performance, some of these larger sampler systems even allow networked playback, out-sourcing the CPU and memory load of samples to other dedicated computers.

With the power and scope of the virtual instruments emerging today, it is possible to produce full orchestral scores from your computer DAW. As nice as these virtual instruments are, they will only get you so far without a good understanding of sampling and sequencing along with solid digital music production skills. Even with the standard sample set that may come with your DAW or sampler software, there are some important production tips and tricks that can help you get the most out of your digital compositions.

Aside: In addition to the round-robin and multiple articulation samples, some instruments also include release samples, such as the subtle sound of a string being released or muted, which are played back when a note is released to help give it a natural and realistic ending. Even these release samples could have round-robin variations of their own!

With the complexity of and diversity of the instruments that you’ll attempt to perform on a simple MIDI keyboard, it may take several passes at individual parts and sections to capture a performance you’re happy with. With MIDI of course, merging performances and editing out
mistakes is a simple task, and doesn’t require dozens of layers of audio files, messy crossfades, and the inherent issues of recorded audio. So use as many takes as you need, and break down the parts as much as you feel helps to improve your performance. If your timing is inconsistent between performances, you can always quantize the notes. Unlike quantization in digital audio, which refers to forcing the audio sample values to a fixed amplitude level according to the bit depth (more on digital audio quantization can be found in Chapter 5), quantization in MIDI refers to adjusting the performance data to the nearest selectable time value, be it whole note, half note, quarter note, etc. While quantizing can help tighten up your performance, it is also a main contributor to your composition sounding mechanical and unnatural. While you don’t want to be off by a whole beat, tiny imperfections in timing is natural with human performance and can help your music sound more convincing. Some DAW sequencers in addition to a unit of timing will let you choose a degree of quantization, in other words how forceful do you want to be when pushing your note data toward that fixed value. This will let you maintain some of the feel of your actual performance. Most sequencers also have a collection of other MIDI processing functions, such as randomization. You can select a group of notes that you’ve quantized, and tell it to randomize the starting position and duration of the note (as well as other parameters such as note velocity) by some small amount, introducing some of these timing imperfections back into the performance. These processes are ironically sometimes known as humanization functions, which nevertheless can come in handy when polishing up a digital music composition.

When actual musicians play their instruments, particularly longer sustained notes, they typically don’t play them at a constant level and timbre. Part of that could be an actual dynamic response to the music or score, such as a crescendo, or a particular articulation/style of playing. But it’s also a simple human factor – sometimes a trumpet player pushes the air a little harder to get the note started or at the end of a long note when they’re running out of breath, and the bowing motion of a stringed instrument varies slightly as the bow moves back and forth. Only a machine could perfectly bow a note exactly the same every time. These nuances aren’t captured in timing or note velocity information, but often these samplers will respond to MIDI control data in order to achieve these more dynamic performances. As the name might imply, an important one of these is MIDI controller 11, expression. Manipulating the expression value is a great way to add a natural, continuous arc to longer sustained notes, and is a great way to simulate dynamics swells as well as the variation in string bowing. In many cases it simply affects the volume of the playback, although some samplers are programmed to respond to it in a more dynamic way. Many MIDI keyboards have an expression pedal input, which will allow you to control the expression with a variable foot pedal. If your keyboard has knobs or faders you could also set those to write expression data, or you can always draw it in to your sequencer software by hand with a mouse or track pad. An example of expression data captured into a MIDI sequencer region can be seen in Figure 8.19. MIDI controller 1, the modulation wheel, is also often linked to a variable element of the sampler instrument. In many cases, it will default as a vibrato control for the sound, either by applying a simple LFO to the pitch, or even dynamically crossfading it with a recorded vibrato sample. When long sustained notes are played or sung live, often a note pitch will start out constant and as the note goes on, an increasing amount of vibrato will be applied, giving it a nice fullness. Taking another pass over your music performance with the modulation wheel in hand, you can bring in varying levels of vibrato at appropriate times, enhancing the character and natural feel of your composition.
As you can see, getting a great piece of music out of your digital composition isn’t always as simple as playing a few notes on your MIDI keyboard. With all of the nuanced control available in today’s sampler systems, crafting even a single instrument part can require careful attention to detail and is often an extreme exercise in patience. Perhaps hiring a live musician is sounding much better at the moment? Certainly, when you have access to the real deal, it is often a simpler and more effective way to create a piece of music. Yet for those who enjoy total control over their production, and are deeply familiar and experienced with their digital music production toolset, virtual instruments can be a fast and powerful means of bringing one’s music to life.

8.4.3 Post-Production

8.4.3.1 Multi-Channel Playback

Mid-Side can also be effective as a playback technique for delivering stereo sound to a large listening area. One of the limitations to stereo sound is that the effect relies on having the listener perfectly centered between the two loudspeakers. This is usually not a problem for a single person listening in a small living room. If you have more than one listener, such as in a public performance space, it can be difficult if not impossible to get all the listeners perfectly centered between the two loudspeakers. The listeners who are positioned to the left or right of the center line will not hear a stereo effect. Instead they will perceive most of the sound to be coming from whichever loudspeaker they are closest to. A more effective strategy would be to set up three loudspeakers. One would be your Mid loudspeaker and would be positioned in front of the listeners. The other two loudspeakers would be positioned directly on either side of the listeners as shown in Figure 8.20.
If you have an existing audio track that has been mixed in stereo, you can create a reverse Mid-Side matrix to convert the stereo information to a Mid-Side format. The Mid loudspeaker gets a L+R audio signal equivalent to summing the two stereo tracks to a single mono signal. The Side+ loudspeaker gets a L-R audio signal, equivalent to inverting the right channel polarity and summing the two channels to a mono signal. This will cancel out anything that is equal in the two channels essentially, removing all the Mid information. The Side- loudspeaker gets a R-L audio signal. Inverting the left channel polarity and summing to mono or simply inverting the Side+ signal can achieve this effect. The listeners in this scenario will all hear something similar to a stereo effect. The right channel stereo audio will cancel out in the air between the Mid and Side+ loudspeakers and the left channel stereo audio will cancel out in the air between the Mid and Side- loudspeakers. Because the Side+/− loudspeakers are directly to the side of the listeners, they will all hear this stereo effect regardless of whether they are directly in front of the MID loudspeaker. Just like Mid Side recording, the stereo image can be widened or narrowed as the balance between the Mid loudspeaker and Side loudspeakers is adjusted.

You don’t need to stop at just three loudspeakers. As long as you have more outputs on your playback system you can continue to add loudspeakers to your system to help you create more interesting soundscapes. The concept of Mid-Side playback illustrates an important concept. Having multiple loudspeakers doesn’t mean you have surround sound. If you play the same sound out of each loudspeaker, the precedence effect takes over and each listener will source the sound to the closest loudspeaker. To create surround sound effects, you need to have different sounds in each loudspeaker. The concept of Mid-Side playback demonstrates how you can modify a single sound to have different properties in three loudspeakers, but you could also have completely different sounds playing from each loudspeaker. For example, instead of having a single track of raindrops playing out of ten loudspeakers, you could have ten different
recordings of water dripping onto various surfaces. This will create a much more realistic and immersive rain effect. You can also mimic acoustic effects using multiple loudspeakers. You could have the dry sound of a recorded musical instrument playing out of the loudspeakers closest to the stage and then play various reverberant or wet versions of the recording out of the loudspeakers near the walls. With multiple playback channels and multiple loudspeakers you can also create the effect of a sound moving around the room by automating volume changes over time.

8.4.3.2 Playback and Control

Sound playback has evolved greatly in the past decades, and it’s safe to say tape decks with multiple operators and reel changes are a thing of history. While some small productions may still use CD players, MiniDiscs, or even MP3 players to playback their sound, it’s also safe to say that computer-based playback is the system of choice, especially in any professional production. Already an integral part of the digital audio workflow, computers offer flexibility, scalability, predictability, and unprecedented control over audio playback. Being able to consistently run a performance and reduce operator error is a huge advantage that computer playback provides. Yet as simple as it may be to operate on the surface, the potential complexity behind a single click of a button can be enormous.

Popular computer sound playback software systems include SFX by Stage Research for Windows operating systems, and QLab by Figure 56 on a Mac. These playback tools allow for many methods of control and automation, including sending and receiving MIDI commands, scripting, telnet, and more, allowing them to communicate with almost any other application or device. These playback systems also allow you to use multiple audio outputs, sending sound out anywhere you want, be it a few specific locations, or the entire sound system. This is essential for creating immersive and dynamic surround effects. You’ll need a separate physical output channel from your computer audio interface for each loudspeaker location (or group of loudspeakers, depending on your routing) in your system that you want to control individually.

Controlling these systems can be as simple as using the mouse pointer on your computer to click a GO button. Yet that single click could trigger layers and layers of sound and control cues, with specifically timed sequences that execute an entire automated scene change or special effect. Theme parks use these kind of playback systems to automatically control an entire show or environment, including sound playback, lighting effects, mechanical automation, and any other special effects. In these cases, sometimes the simple GO isn’t even triggered by a human operator, but by a timed script, making the entire playback and control a consistent and self-reliant process. Using MIDI or Open Sound Control you can get into very complex control systems. Other possible examples include using sensors built into scenery or costumes for actor control, as well as synchronizing sound, lighting, and projection systems to keep precisely timed sequences operating together and exactly on cue, such as a simulated lighting strike. Outside of an actual performance, these control systems can benefit you as a designer by providing a means of wireless remote control from a laptop or tablet, allowing you to make changes to cues while listening from various locations in the theatre.

Using tools such as Max or PD, you can capture input from all kinds of sources such as cameras, mobile devices, or even video game controllers, and use that control data to generate MIDI commands to control sound playback. You’ll always learn more actually doing it than simply reading about it, so included in this section are several exercises to get you going making your own custom control and sound playback systems.
You may also choose to apply some overall EQ in the mastering process to suit your taste. In some cases you may also manipulate the stereo image a bit to widen or narrow the overall stereo effect. You may also want to add a multi-band limiter at the end of the processing chain to catch any stray clipped signals that may have resulted from your other processes. If you're converting to a lower bit depth, you should also apply a dither process to the mix to account for any quantization errors. For example, CDs require 16-bit samples, but most recording systems use 24 bits. Even if you are not converting bit depth you may still want to use dither since most DAW programs process the audio internally at 32 or 64 bits before returning back the original 24 bits. A 24-bit dither could help you avoid any quantization errors that would occur in that process. Figure 8.22 shows an example of a multi-band limiter that includes a dither processor.
8.5 References
Programming Exercise:
Title of Exercise