11

Reverberation

Reverberation (*reverb*) is one of the most often used effects in audio production. In this chapter, we will look at the causes, the main characteristics, and the measures of reverb. We then focus on how to simulate reverb. First we describe algorithmic approaches to generating reverb, focusing on two classic designs. Then we look at a popular technique for generating a room impulse response, the image source method. Next, we describe convolutional reverb, which adds reverb to a signal by convolving that signal with a room impulse response, either recorded or simulated (such as from using the image source method). The chapter concludes by looking at implementation details and applications.

Theory

In a room, or any acoustic environment, there is a direct path from any sound source to a listener, but sound waves also take longer paths by reflecting off the walls, ceiling, or objects, before they arrive at the listener, as shown in Figure 11.1. These reflected sound waves travel a longer distance than the direct sound and are partly absorbed by the surfaces, so they take longer to arrive and are weaker than the direct sound. These sound waves can also reflect off of multiple surfaces before they arrive at the listener. These delayed and attenuated copies of the original sound are what we call reverberation, and it is essential to the perception of spaciousness in the sounds.

Reverberation is more than just a series of echoes. An echo is the result of a distinct, delayed version of a sound, as could be heard with a delay of at least 40 ms. With reverberation from a typical room, there are many, many reflections, and the early reflections arrive on a much shorter time scale. So these reflections are not perceived as distinct from the sound source. Instead, we perceive the effect of the combination of all the reflections.

Reverberation is also more than a simple delay device with feedback. With reverb the rate at which the reflections arrive will change over time, as opposed to just simulating reflections that have a fixed time interval between them. In reverberation, there is a set of reflections that occur shortly after the direct sound. These *early reflections* are related to the position of the source and listener in the room, as well as the room's shape, size, and material

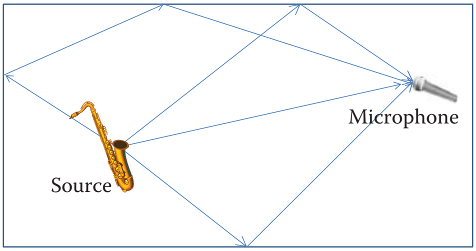

FIGURE 11.1
Reverb is the result of sound waves traveling many different paths from a source to a listener.

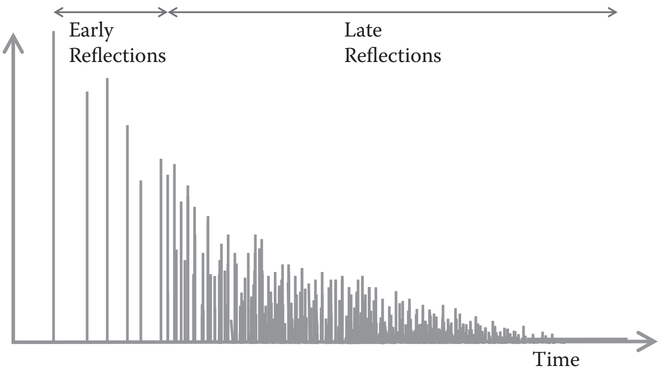

FIGURE 11.2
Impulse response of a room.

composition. The later reflections arrive much more frequently, appear more randomly, usually decay exponentially, and are difficult to directly relate to the physical characteristics of the room. These *late reflections* give rise to *diffuse reverberation*. An example impulse response for a room is depicted in Figure 11.2. Each vertical line marks when the original sound is repeated, and the height of each of these lines is the amplitude of the sound at that time.

A measure that is often used to characterize the reverb in an acoustic space is the *reverberation time*, often denoted RT_{60}. This is the time that it takes for sound pressure level or intensity to decay by 60 dB, i.e., 1/1,000,000th of its original intensity, or 1/1000th of its original amplitude (see Chapter 1). A long reverberation time implies that the reflections remain strong for a long time before their energy is absorbed. The reverberation time is associated with room size. Small rooms tend to have reverb times on the order of hundreds of milliseconds, though this can vary greatly depending on the acoustic treatment and other factors. Larger rooms will usually have longer reverberation times since, on average, the sound waves will travel a larger distance between reflections. Concert halls typically have reverberation times around

1.5 to 2 s. Cathedrals and other highly reverberant environments may have reverb times of more than 3 s.

It is possible to construct a large room with short reverberation time, and vice versa. The reverberation time is dictated primarily by the size of the room and the surfaces in the room. The surfaces determine how much energy is lost or absorbed each time the sound is reflected. Highly reflective materials, such as a concrete or tile floor, will result in a long reverb time. Absorptive materials, such as curtains, cushions, or heavy carpet, will reduce the reverberation time. People and their clothing absorb a lot of sound. This explains why a room may sound "bigger" during a sound check prior to a performance, but smaller and more intimate once the audience has arrived. The absorptivity of most materials usually varies with frequency, which is one reason the reverb time is dependent on the spectral content of the source signal (formal measurement of reverb time is performed using an impulse or turning off a noise generator). The air in the room will also attenuate the sound waves, reducing the reverberation time. This attenuation is a function of temperature and humidity and is most significant for high frequencies. Because of this, many audio effect implementations of reverb will include some form of low-pass filtering, as will be discussed later in this chapter.

Another important measure of reverberation is the *echo density*, defined as the frequency of peaks (number of echoes per second) at a certain time t during the impulse response. The more tightly packed together the reflections are, the higher the echo density. If the echo density is larger than 20–30 echoes per second, the ear no longer hears the echoes as separate events, but fuses them into a sensation of continuous decay. In other words, the early reflections become a late reverberation.

Sabine and Norris–Eyring Equations

The *mean free path* of a room gives the average distance a sound wave will travel before it hits a surface. Assuming a rectangular box room, this is given by $4V/S$, where V is the volume and S is the surface area of the room. Dividing this by the speed of sound c, the mean time τ until a sound source hits a wall is $\tau = 4V/(cS)$. So the mean number of reflections over a time t is

$$n(t) = t/\tau = t\frac{cS}{4V} \tag{11.1}$$

At each reflection off a surface S_i, α_i is the proportion of the energy absorbed. So $1 - \alpha_i$ is the proportion of the energy that is reflected back into the room. Now assume all surfaces have the same absorption coefficient, α. So after n reflections, the sound has been reduced by a factor $(1 - \alpha)^n$. Thus, after a time t, the sound energy has been reduced to

$$E(t) = (1-\alpha)^{tcA/(4V)}E(0) \tag{11.2}$$

Taking the log base (1 – α) of both sides, we have

$$\log_{(1-\alpha)}\left(E(t)/E(0)\right) = tcS/(4V) \tag{11.3}$$

Now recall that $\log_A(x) = \ln(x)/\ln(A)$. So, putting the time t on one side,

$$t = \frac{4V\log_{(1-\alpha)}\left(E(t)/E(0)\right)}{cS} = \frac{4V\ln\left(E(t)/E(0)\right)}{cS\ln(1-\alpha)} \tag{11.4}$$

Recall that the reverberation time, RT_{60}, is the time for the sound energy to decrease by 60 dB. Thus,

$$RT_{60} = \frac{4V\ln(10^{-6})}{cS\ln(1-\alpha)} \approx \frac{-0.161V}{S\ln(1-\alpha)} \tag{11.5}$$

This is known as the Norris–Eyring formula. However, it has several approximations. The most important is that it assumes a single absorption coefficient, α. A more accurate form of Equation (11.5) is

$$RT_{60} = \frac{-0.161V}{S\ln\left(1-\frac{1}{S}\sum_i S_i\alpha_i\right)} \tag{11.6}$$

Alternatively, it can be simplified even further. For small α, ln(1 – α) ~ –α. So Equation (11.5) becomes

$$RT_{60} \approx \frac{0.161V}{S\alpha} \tag{11.7}$$

This is the Sabine equation, which was first found empirically by Wallace Clement Sabine in the late 1890s.

The absorption coefficient of a material ranges from 0 to 1 and indicates the proportion of sound that is not reflected by (so either absorbed by or transmitted through) the surface. A thick, smooth painted concrete ceiling would have an absorption coefficient very close to 0. Analogous to a mirror reflecting light, almost all sound would be reflected with very little attenuation. Conversely, a large, fully open window would have an absorption coefficient of 1, since any sound reaching it would pass straight through and not be reflected, in which case Sabine's formula would be a very poor approximation, since it could still generate significant reverberation time.

As mentioned, absorption, and hence the reverberation time, is a function of frequency. Usually, less sound energy is absorbed in the lower-frequency ranges, resulting in longer reverb times at lower frequencies. Nor do the equations above take into account room shape or losses from the sound traveling through the air, which is important in larger spaces. More detailed discussion of the reverbation time formulae (with differing interpretations) is available in [85, 86].

Direct and Reverberant Sound Fields

The reverberation due to sound reflection off surfaces is extremely important. Reverberation keeps sound energy contained within a room, raising the sound pressure level and distributing the sound throughout. Outside in the open, there are far less reflective surfaces, and hence much of the sound energy is lost.

For music, reverberation helps ensure that one hears all the instruments, even though they may be at different distances from the listeners. Also, many acoustic instruments will not radiate all frequencies equally in all directions. For example, without reverberation the sound of a violin may change considerably as the listener moves with respect to the violin. The reverberation in the room helps to diffuse the energy a sound wave makes so that it can appear more uniform when it reaches the listener. But when the reverberation time becomes very large, it can affect speech intelligibility and make it difficult to follow intricate music.

A distinction is often made between the direct and reverberant sound fields in a room. The sound heard by a listener will be a combination of the direct sound and the early and late reflections due to reverberation. When the sound pressure due to the direct sound is greater than that due to the reflections, the listener is in the *direct field*. Otherwise, the listener is in the *reverberant field*.

The *critical distance* is defined as the distance away from a source at which the sound pressure levels of the direct and reverberant sound fields are equal. This distance depends on the shape, size, and absorption of the space, as well as the characteristics of the sound source. A highly reverberant room generates a short critical distance, and a nonreverberant or anechoic room generates a longer critical distance.

For an omnidirectional source, the critical distance may be approximately given by the following [87]:

$$d_c \sim \sqrt{\frac{V}{100\pi RT_{60}}} \sim 0.141\sqrt{S\alpha} \tag{11.8}$$

where critical distance d_c is measured in meters, volume V is measured in m^3, and reverberation time RT_{60} is measured in seconds. More accurate approximations are also available [88].

THE AVANT-GARDE ANECHOIC CHAMBER

An acoustic anechoic chamber is a room designed to be free of reverberation (hence non-echoing or echo-free). The walls, ceiling, and floor are usually lined with a sound absorbent material to minimize reflections and insulate the room from exterior sources of noise. All sound energy will travel away from the source with almost none reflected back. Thus, a listener within an anechoic chamber will hear only the direct sound, with no reverberation.

Anechoic chambers effectively simulate quiet open-spaces of infinite dimension. Thus, they are used to conduct acoustics experiments in "free field" conditions. They are often used to measure the radiation pattern of a microphone or of a noise source, or the transfer function of a loudspeaker.

An anechoic chamber is very quiet, with noise levels typically close to the threshold of hearing in the 10–20 dBA range (the quietest anechoic chamber has a decibel level of –9.4 dBA, well below hearing). Without the usual sound cues, people find the experience of being in an anechoic chamber very disorienting and often lose their balance. They also sometimes detect sounds they would not normally perceive, such as the beating of their own hearts.

One of the earliest anechoic chambers was designed and built by Leo Beranek and Harvey Sleeper in 1943. Their design is the one upon which most modern anechoic chambers are based. In a lecture titled "Indeterminacy," the avant-garde composer John Cage described his experience when he visited Beranek's chamber: "In that silent room, I heard two sounds, one high and one low. Afterward I asked the engineer in charge why, if the room was so silent, I had heard two sounds … . He said, 'The high one was your nervous system in operation. The low one was your blood in circulation'" [89].

After that visit, he composed his famous work entitled 4′33″, consisting solely of silence and intended to encourage the audience to focus on the ambient sounds in the listening environment.

In his 1961 book *Silence*, Cage expanded on the implications of his experience in the anechoic chamber. "Try as we might to make silence, we cannot … . Until I die there will be sounds. And they will continue following my death. One need not fear about the future of music" [90].

Implementation

Algorithmic Reverb

Early digital reverberators that tried to emulate a room's reverberation primarily consisted of two types of infinite impulse response (IIR) filters, allpass and comb filters, to produce a gradually decaying series of reflections.

Schroeder's Reverberator

Perhaps the first important artificial reverberation was devised by Manfred Schroeder of Bell Telephone Laboratories in 1961. Early Schroeder reverberators [91, 92] consisted of three main components: comb filters, allpass filters, and a mixing matrix. The first two are still used in many algorithmic reverbs of today, but the mixing matrix, designed for multichannel listening, is either not used or is replaced with more sophisticated spatialization methods.

Figure 11.3 is a block diagram of the Schroeder reverberator. This design does not create the increasing arrival rate of reflections, and modern algorithmic approaches are more realistic. Nevertheless, it provides an important basic framework for more advanced approaches.

There is a parallel bank of four feedback comb filters. The comb filter shown in Equation (11.9) is a special case of an IIR digital filter, because there is feedback from the delayed output to the input. The comb filter effectively simulates a single-room mode. It represents sound reflecting between two parallel walls and produces a series of echoes. The echoes are exponentially decaying and uniformly spaced in time. The advantage is that the decay time can be used to define the gain of the feedback loop.

The comb filters are connected in parallel. The comb filters have an irregular magnitude response and can be considered a simulation of four specific echo sequences. The delay lengths in these comb filters may be used to adjust the illusion of room size, although if they are shortened, there should be more of them in parallel according to Schroeder. They also serve to reduce the spectral anomalies. Each comb filter has different parameters in order to attenuate frequencies that pass through the other comb filters. By controlling

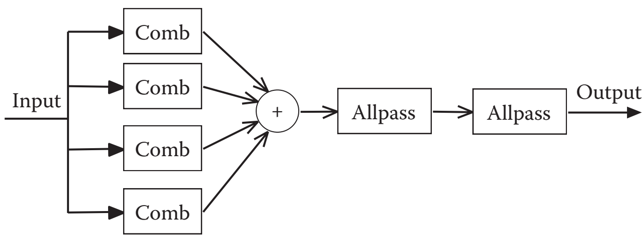

FIGURE 11.3
Schroeder's reverberator.

the delay time, which is also called loop time, of each filter, the sound waves will become reverberated.

The time domain and z domain form of Schroeder's comb filter is given by

$$
\begin{aligned}
y[n] &= x[n-d] + gy[n-d] \\
H_{Comb}(z) &= z^{-d} \big/ \left(1 - gz^{-d}\right)
\end{aligned}
\tag{11.9}
$$

where d is a delay in samples and g is the filter's feedback coefficient. The block diagram is shown in Figure 11.4a. The parallel comb-filter bank is intended to give an appropriate fluctuation in the reverberator frequency response. A feedback comb filter can simulate a pair of parallel walls, so one could choose the delay line length in each comb filter to be the number of samples it takes for a wave to propagate from one wall to the opposite wall and back. Schroeder based his approach on trying to get the same number of impulse response peaks per second as would be found in a typical room.

The comb-filter delay line lengths can be more or less arbitrary, as long as enough of them are used in parallel (with mutually prime delay line lengths) to achieve a perceptually adequate fluctuation density in the frequency response magnitude. In Schroeder's paper, four such delays are chosen between 30 and 45 ms, and the corresponding feedback coefficients g_i are set to give the desired overall decay time.

One problem with a single comb filter is that the distances between adjacent echoes are decided by a single delay parameter. Thus, the output of a comb filter has very distinct periodicity, producing a metallic sound, sometimes known as flutter echoes. This rapid echo is often found in the acoustics of long narrow spaces with parallel walls. To counteract this, Schroeder's main design connects two allpass filters in series.

The allpass filter is given below in (11.10), and a block diagram is depicted in Figure 11.4b.

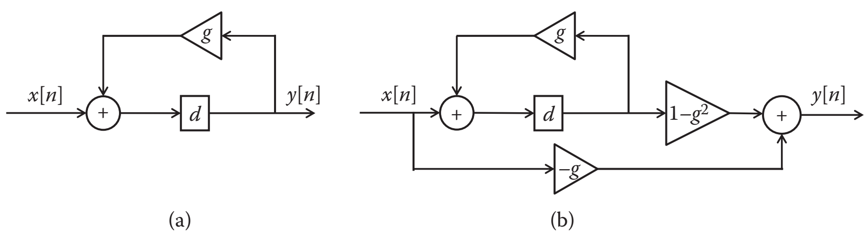

FIGURE 11.4
(a) Schroeder's comb filter. (b) A Schroeder allpass section. A typical value for g is 0.7.

$$y[n] = -gx[n] + x[n-d] + gy[n-d]$$
$$H_{AP}(z) = \frac{z^{-d} - g}{1 - gz^{-d}} \qquad (11.10)$$

The allpass filters provide "colorless" high-density echoes in the late impulse response of the reverberator. Essentially, these filters transform each input sample from the previous stage into an entire infinite impulse response, which results in a higher echo density. For this reason, Schroeder allpass sections are sometimes called *impulse diffusers*. Unlike the comb filters, these allpass filters give the same gain at every frequency. But though they do not provide an accurate physical model of diffuse reflections, they succeed in expanding the single reflections into many reflections, which produces a similar qualitative result.

Schroeder recognized the need to separate the coloration of reverberation (changing the frequency content) from its duration and density aspects. In Schroeder's original work, and in much work that followed, allpass filters are arranged in series, as shown in Figure 11.3, thus maintaining the uniform magnitude response. Since all of the filters are linear and time invariant, the series allpass chain can go either before or after the parallel comb-filter bank.

Schroeder suggests a progression of allpass delay line lengths close to the following:

$$d_i \sim 100 \text{ m}/3^i, \quad i = 0,\ 1,\ 2,\ 3,\ 4 \qquad (11.11)$$

The delay line lengths d_i are typically mutually prime and spanning successive orders of magnitude. The 100 ms value was chosen so that when $g = 0.708$ in Equation (11.10), the time to decay 60 dB (T_{60}) would be 2 s. Thus, for $i = 0$, $T_{60} \sim 2$, and each successive allpass has an impulse response duration that is about a third of the previous one. Using five series allpass sections in this way yields an impulse response echo density of about 810 per second, which is close to the desired thousand per second.

A system using one of these reverberators simply adds its output, suitably scaled, to the reverberator input sample at the current time.

Moorer's Reverberator

The next important advance in algorithmic reverberators is attributable to James Moorer in 1979 [93]. Moorer's reverberator is the combination of a finite impulse response (FIR) filter that simulates the early impulse response activity of the room in cascade with a bank of low-pass and comb filters. The FIR filter coefficients were based on the simulation of a concert hall.

Although comb filters can be used for modeling the decaying impulse response, as in Schroeder's reverberator, they do not simulate the tendency