Fifty Years of Artificial Reverberation

Vesa Välimäki, *Senior Member, IEEE*, Julian D. Parker, Lauri Savioja, *Senior Member, IEEE*, Julius O. Smith, *Member, IEEE*, and Jonathan S. Abel, *Member, IEEE*

***Abstract*—The first artificial reverberation algorithms were proposed in the early 1960s, and new, improved algorithms are published regularly. These algorithms have been widely used in music production since the 1970s, and now find applications in new fields, such as game audio. This overview article provides a unified review of the various approaches to digital artificial reverberation. The three main categories have been delay networks, convolution-based algorithms, and physical room models. Delay-network and convolution techniques have been competing in popularity in the music technology field, and are often employed to produce a desired perceptual or artistic effect. In applications including virtual reality, predictive acoustic modeling, and computer-aided design of acoustic spaces, accuracy is desired, and physical models have been mainly used, although, due to their computational complexity, they are currently mainly used for simplified geometries or to generate reverberation impulse responses for use with a convolution method. With the increase of computing power, all these approaches will be available in real time. A recent trend in audio technology is the emulation of analog artificial reverberation units, such as spring reverberators, using signal processing algorithms. As a case study we present an improved parametric model for a spring reverberation unit.**

***Index Terms*—Acoustics, acoustic scattering, acoustic signal processing, architectural acoustics, convolution, infinite impule response (IIR) digital filters.**

I. Introduction

SCHROEDER introduced the idea of artificial reverberation based on digital signal processing fifty years ago as of this writing [1]. In this overview article, we trace the development of artificial reverberation, concentrating on computational methods, starting with Schroeder's introduction of the digital comb and allpass filters in 1961.

Reverberation refers to the prolonging of sound by the environment, which is essentially caused by the reflectivity of surfaces and by the slow speed of sound in air, only about 345 m/s at room temperature [2]–[4]. As sound radiates from a source, it interacts with the environment, carrying with it to the listener an imprint of the space, and a sense of the objects and architecture present. Everyday objects and building materials are highly reflective of acoustic energy, and in an enclosed space, a listener invariably experiences source sounds smeared over time.

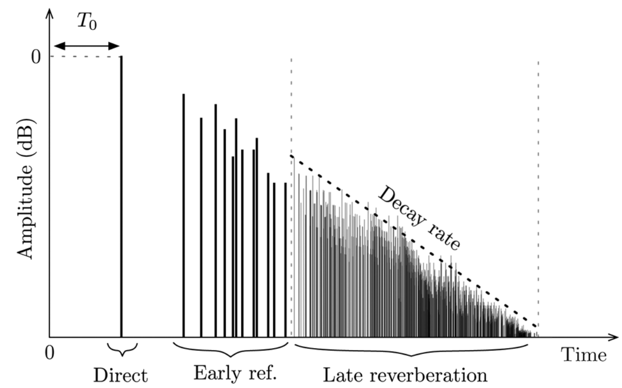


Fig. 1. Schematic example of a generic room impulse response.

Sound typically arrives at a listener in stages. Fig. 1 shows a simplified example of a room impulse response. A signal traveling along a *direct path* between the source and listener arrives after a delay of T_0 and is followed by *early reflections* from nearby objects and surfaces. While the direct path reveals the direction of the source, the early reflections convey a sense of the geometry and materials of the space. Reflections propagating in the environment themselves interact with features of the space to generate additional reflections, and therefore produce an increasing echo density. An example of the impulse response of a large room is given in Fig. 2.

Over time, isolated arrivals at the listener give way to a dense *late reverberation*, which forms the tail of the impulse response shown in Fig. 1. The tail is characterized by Gaussian statistics and an evolving power spectrum indicative of the size of the space and absorbing power of the materials present [5]. The decay time of the late reverberation, expressed in seconds per 60-dB decay, is called the *reverberation time*, denoted T_{60} [4].

Many environments produce distinctive reverberation: Consider sound created in a concert hall [4], [6], [7], a forest [8], a city street [9], [10], or on a mountain range [11]. The impression of a sound source is tied to the space in which it is heard, and as a result, artificial reverberation is widely used in music, film and virtual environment applications for artistic effect or to convey a sense of the space.

The need for artificial reverberation first arose in the context of music broadcasting and recording. The damped studio environment and close micing produced a "dry" sound that lacked the concert hall acoustics desired for music performance. As early as the 1920s, reverberation was artificially applied by sending the dry studio signal to a reverberant environment, often

Manuscript received September 29, 2011; revised February 08, 2012; accepted February 08, 2012. Date of publication February 29, 2012; date of current version March 21, 2012. The work of J. D. Parker was supported by the Graduate School of Electronics, Telecommunications, and Automation and by the Academy of Finland under Project 122815. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Walter L. Kellermann.

V. Välimäki and J. D. Parker are with the Department of Signal Processing and Acoustics, School of Electrical Engineering, Aalto University, Espoo FI-00076, Finland (e-mail: vesa.valimaki@aalto.fi).

L. Savioja is with the Department of Media Technology, School of Science, Aalto University, Espoo FI-00076, Finland.

J. O. Smith and J. S. Abel are with CCRMA, Stanford University, Stanford, CA 94304 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASL.2012.2189567

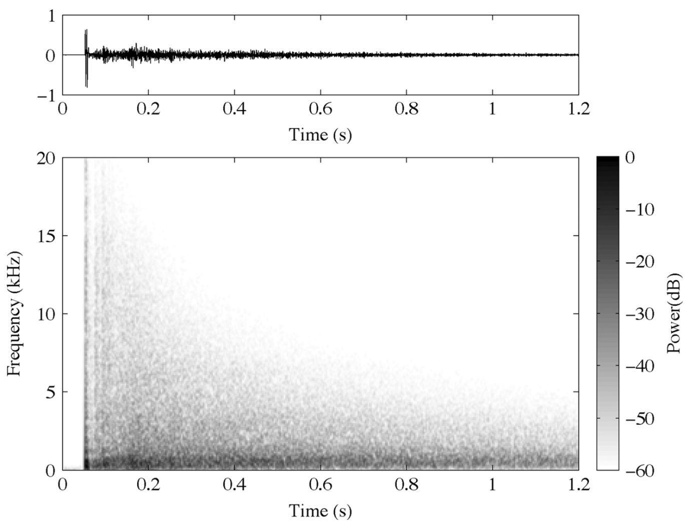


Fig. 2. Impulse response of the Memorial Church at Stanford University, measured using a balloon pop.

a specially constructed echo chamber [12], [13]. A number of electromechanical reverberation devices have been developed, including tape delays [14], [15], spring reverberation devices [16], and reverberation plates [17], [18].

While plates, chambers and the like produced high-quality reverberation, their physical nature generally limited their use to studio recording and broadcast settings. They were often difficult if not impossible to transport, were sensitive to external acoustic or mechanical disturbances, and required specialized knowledge to maintain and tune. Additionally, the reverberation produced would vary somewhat from unit to unit, and could change over time or with changing environmental conditions. By contrast, artificial reverberation produced by computational processes had the promise of convenience, portability and repeatability, and would allow automation (the recording, editing and play back of reverberator control parameter changes over time [19]). This motivated research into methods for digitally synthesizing high-quality reverberation.

Acoustic spaces, chambers and plates are distributed systems, well modeled by multidimensional wave equations. Disturbances propagating in these systems produce an increasingly dense set of reflections through boundary interactions, and maintain a correspondingly dense set of modes. The duration of typical room responses to applied sounds far exceeds the period of a 20-Hz sinusoid at the lower limit of human hearing. As a result, room responses are primarily heard temporally, as a sequence of events, rather than spectrally, as the presence or absence of modal frequencies. The structures available in a computational setting, however, are lumped elements, and the difficulty in developing a computational reverberation method lies in combining lumped elements to produce the effect of a distributed system.

The process of reverberation is approximately linear and time invariant, and computational methods for generating artificial reverberation attempt to reproduce the psychoacoustic impact of various reverberation impulse response features [20]–[23]. Reverberation algorithms generally fall into one of three categories (although hybrids exist):

1) *delay networks*, in which the input signal is delayed, filtered and fed back along a number of paths according to parametrized reverberation characteristics;
2) *convolutional*, wherein the input signal is simply convolved with a recorded or estimated impulse response of an acoustic space;
3) *computational acoustic*, wherein the input signal drives a simulation of acoustic energy propagation in the modeled geometry.

Which approach is used depends on the computational setting and the application. Delay-network and convolution techniques have been competing in popularity in the music technology field, and they are often employed to produce a desired perceptual or artistic effect.

Computational acoustic methods generally find application in acoustic design and analysis scenarios, such as predictive acoustic modeling and computer-aided design (CAD) of acoustic spaces, where accuracy is desired, but measurements are unavailable. The process of making room acoustic modeling or measurement results audible is called auralization [24]–[27]. The MPEG-4 standard includes a virtual acoustics modeling framework specifying many parameters for controlling artificial reverberation in synthetic multimedia presentations [28], [29].

There are many other applications of digital reverberation technology. In room acoustic enhancement, the acoustics of a hall are manipulated using loudspeakers to playback processed versions of sound captured at a set of microphones in the space [30]–[32], [6], [4]. In an extreme case, a small room can be converted to having the reverberation qualities of a large concert hall, which is useful for rehearsal purposes [33], [34]. In headphone audio, the addition of reverberation, particularly the early reflection part, helps to externalize the sound image, which otherwise is usually perceived inside the listener's head [35]–[38]. Artificial reverberation can also be used in the upmixing process to play stereo audio signals over multiple loudspeakers [39]. In speech processing research, artificial reverberation is applied to voice signals to evaluate its effect, because it is known that reverberation reduces intelligibility and degrades the performance of automatic speech recognition [40]–[43].

In sound synthesis applications, the resonator of a musical instrument, such as the body of a guitar [44], [45] or the soundboard of a keyboard instrument [46], [47] can be simulated using a reverberation algorithm, since vibration propagation in mechanical structures shares features with acoustic wave propagation in rooms. Similarly, the sustain pedal effects of the piano can be simulated as a special kind of reverberation [48]. A recent trend in music technology is to simulate "vintage" analog and electromechanical reverberation units by software [49]–[53].

The rest of this paper is organized as follows. Section II gives a brief tour of the historical development of artificial reverberation methods, both analog and digital. Section III focuses on artificial reverberation algorithms based on relatively sparse networks of delay lines and digital filters. Section IV tackles physical room models for virtual reality, gaming, and computer-aided design of concert halls. Section V discusses convolution techniques. Section VI is devoted to virtual analog reverberators, which simulate tape-based delays, bucket-brigade delay lines, reverberation plates, and spring reverberators. Section VII concludes this overview article.

II. Historical Developments

We begin with an overview of synthetic reverberation devices and methods, starting with early analog approaches and continuing with the digital techniques that are currently predominant.

A. Analog Methods

Artificial reverberation was first developed in the 1920s by RCA for use in broadcast applications [13]. The technique involved transmitting a sound into an empty acoustic space, and recording the response of the space via an appropriately positioned microphone [54]. M. T. "Bill" Putnam is credited with introducing artificial reverberation to studio recording [55]. He built a number of prized *echo chambers*, constructed to lack parallel surfaces so as to prevent flutter echo, and with a loudspeaker mounted in a corner to excite a large number of modes. Microphones could be positioned within the chamber to subtly vary the early portion of the reverberation, and wall hangings or other damping materials could be placed to provide some control over reverberation time. In use, the audio engineers would sum the original, *dry*, signal with the reverberated, *wet*, signal, controlling the relative mix for artistic or perceptual effect. As described above, this approach has the advantage of providing high-quality reverberation, but is costly and lacks portability and flexibility [56].

A number of electromechanical reverberators have been developed, the earliest being the *spring reverberator*, invented in the late 1920s by Hammond [16], [57] as a compact method for adding an acoustic quality to the dry sound of his electric organ. Hammond's device consisted of an electromechanical transducer which excited vibrations in helical springs according to an applied electrical signal. These vibrations were then received by a mechanico-electrical transducer, and mixed with the dry input signal. Each spring element produced a decaying series of echoes, providing some impression of an acoustic space. To generate an echo density profile more consistent with that of an acoustic space, spring elements were combined in series and parallel [58], [59]. To control the rate of energy decay, damping mechanisms, including partial submersion of the spring elements in oil, were employed [16], [60].

Several developments greatly improved the performance, cost and size of spring reverberators. The use of the torsional mode for wave propagation and mechanisms for holding the springs under tension, introduced in the late 1950s and early 1960s, greatly decreased their sensitivity to vibration and mechanical shock, and allowed the use of small-diameter springs [61]. This paved the way for their widespread use in guitar amplifiers and electronic musical instruments, and even in home stereo devices such as the Sansui RA-700.

Owing largely to their dispersive propagation [50], [52], spring reverberators have a distinctive sound. By placing impedance discontinuities along the spring, the echo density is increased, and the dispersion is less apparent. A number of mechanisms for doing this were developed in the 1970s and early 1980s, and found their way into the AKG reverberators, including the AKG BX20E [62], which has a very natural room-like sound.

In the late 1950s, the plate reverberator was introduced [56], [63] by German company Elektromesstechnik (EMT). The most widely used plate in high-end recording studios was the EMT 140, which featured a large, thin steel resonant plate held under tension. A single transducer near the plate center induced transverse vibrations on the plate, while a pair of sensors near the plate edges picked up the resulting mechanical disturbances. A damping plate coated with packed asbestos was positioned along side the resonant plate, with the distance between the plates determining the reverberation decay time. Due to the fast sound speed on the resonant plate relative to its size and somewhat dispersive propagation, the plate response is very quickly dense with arrivals, the high-frequencies slightly outrunning the low frequencies, thus giving the plate its characteristic whip-like onset [18], [64]. Later plate designs replaced the steel plate with gold foil [65]. The speed of sound is slower in gold than in steel, resulting in the effect of a larger and more room-like space.

Another type of effect used to add an impression of space in music production is the "delay" or "echo" effect. This differs from reverberation in that it generally concentrates on producing a series of repeating echoes which are clearly audible individually. Echo effects were achieved from the 1940s onwards via the use of tape-based techniques. A loop of tape could be made, which was then used in a tape deck set to both record and playback at the same time. The result was a decaying series of echoes. Initially, standard tape machines were used. Later, tape machines specifically designed for producing the echo effect were introduced, with multiple playback heads at different positions used to provide a more complex pattern of echoes. Examples of such units include the Watkins Copicat, the Meazzi Manager 666, the Echoplex, and the Roland Space Echo [51].

In the late 1960s, Philips developed the "Bucket-Brigade" Device (BBD) as a method of producing audio delay in an analog circuit [66]. The BBD delays a signal by sampling it, and storing the signal value in a capacitor as a charge. This charge is then passed along a long chain of capacitors through the opening and closing of interspersed MOS transistor switches controlled by an external clock signal. The BBD is therefore a sampled analog system, and produces a delay given by the length of the capacitor chain and the frequency of the clocking signal. Most BBDs were used for simple echo applications. However, some BBD chips such as the Panasonic MN3011 [67] were produced with multiple output taps in mutually prime positions, with the intention that they could be used to implement an analog electrical reverberation effect. This technique did not become popular, as it coincided with the availability of the first digital reverberators, but a number of BBD reverberation units were produced by companies such as Electro-Harmonix.

B. Digital Methods

The first digital reverberation algorithms were proposed in early 1960s by Manfred Schroeder and Ben Logan [1], [68]. In doing so, they introduced the digital allpass filter [1], which produced a series of decaying echoes, but maintained an overall "colorless" spectrum, irrespective of the decay rate. Also described was a nested allpass structure developed for the purpose of providing a means of controlling the reverberator wet/dry mix. A series combination of allpass filters was described as a means to achieve an increasing echo density, and a structure involving a parallel set of comb filters driving a cascade of allpass filters was suggested for providing independent control over the wet/dry mix, delay of the reverberated sound and decay rate. Schroeder later provided an additional structure for simulating early reflections using a sparse FIR filter [69].

Moorer published an important paper on reverberation algorithms in 1979 [20]. His work was more tutorial in nature than Schroeder's earlier work, and included example architectures with delay line and filter parameter values. Moorer enhanced the Schroeder structures by inserting one-pole filters into delay loops to control reverberation time as function of frequency.

Schroeder's work provided a basis for the first commercial digital reverberators, developed for music production in the mid and late 1970s. The first commercial device for digital audio processing was a digital delay unit, Lexicon Delta T-101, which appeared in 1971 [70], [71]. The first digital reverberators included the EMT 250 (introduced in 1976) [72], the Ursa Major Space Station (1978) [73], the Lexicon 224 (1978) [74], and the AMS RMX-16 (1981) [75]. Little memory and computational resources were available to the early units, making it challenging to create reverberation without unwanted resonances or irregularities in the decay. To overcome these limitations, Blesser and Bader in [76] (and likely used in the EMT 250) propose Schroeder-type structures in combination with decorrelating processes. Griesinger in the Lexicon 224 reportedly arranged a number of comb and allpass filters in a physically inspired architecture, and used time varying delay line lengths to ensure a smooth decay. The RMX-16 appears to use a cascade of allpass-like structures with carefully selected delay line lengths to provide a smooth decay, and with feedback to the input through a filter to control the decay time as a function of frequency. A similar architecture was later described in [77].

The image-source method was formulated for room acoustics by Allen and Berkeley in their 1979 article [78]. The idea was to replace a source in an enclosed space with a source and set of properly positioned virtual sources. The method was extended by Borish in [79] to include non-convex spaces. The image-source method is often applied in virtual acoustics systems to calculate the delays and directions for early reflections [80], [25].

Smith introduced the digital waveguide method as a physical modeling approach to artificial reverberation [81], [82]. The method defines a network of waveguides, connected via lossless scattering junctions. The waveguides include bidirectional delay lines, and can incorporate losses via scaling or filtering. Subject to numerical effects, these reverberators can be made to reverberate indefinitely, as all losses in the system may be eliminated. Two architectures were common, the "daisy" in which all waveguides emanated from and returned to a single scattering junction, and the "football" in which the waveguides conveyed energy between two scattering junctions.

Gerzon in 1971 [83] described multichannel extensions to the digital allpass filter, including a network of delay lines with a unitary feedback matrix. Control over the reverberation decay time as a function of frequency via feedback filters was discussed. In related work, a nested allpass structure was described in [84].

In 1991, Jot and Chaigne published their Feedback Delay Network (FDN) approach to digital reverberation [85], [86]. This work presented a novel, general design structure for reverberators which provided separate, independent control over the energy storage, damping and diffusion components of the reverberator. The form is similar to that described by Gerzon, but with a simple, elegant method for feedback filter design to control decay rate. The FDN remains even today a state-of-the-art reverberation method against which new solutions are compared. A short time later, two excellent reviews on reverberation algorithms appeared in 1997 [87], [21].

One of the first commercial reverberation products to use convolution was the Lake DSP Huron [88], designed for virtual reality applications. A room impulse response was convolved with an input signal, and the first 4096 taps were crossfaded from a table of impulse response onsets according to the listener position.

Convolution reverbs for music and film production first appeared in the late 1990s, with the introduction of the sampling digital reverb by Sony [89]. Yamaha soon followed with the Yamaha SREV-1. Waves IR1, released in 2004, provided interactive control over reverberation parameters, such as decay time. The popularity of convolutional reverberators was enabled by improved impulse response measurement techniques, including, e.g., Farina's swept sinusoid method [90], which led to the widespread availability of room impulse responses.

The newest milestone in the progress of artificial reverberation has been the appearance of new graphics processing units (GPUs), which are as powerful as supercomputers of the near past. In addition to image and video data processing, GPUs can be applied to audio processing, and this has finally enabled running physically based room simulations in real time [91]–[93].

III. Delay Network Methods

The earliest and still most efficient approaches to artificial reverberation are based on networks of delay lines and digital filters. This section reviews some of the main developments in chronological order.

A. Comb Filters

Comb filters were proposed for use in artificial reverberation by Schroeder and Logan [1]. Fig. 3 shows a block diagram of the feedback comb-filter structure proposed in [1] having transfer function[1]

$$H(z) = \frac{Y(z)}{X(z)} = \frac{z^{-M}}{1 - gz^{-M}}. \tag{1}$$

For stability, we require $|g| < 1$. The impulse response is then

$$h(n) = \delta(n-M) + g\delta(n-2M) + g^2\delta(n-3M) + \dots \tag{2}$$

[1]The corresponding diagram in [1] was for continuous time, but it was implemented digitally. Implementation using tape delay was also contemplated.

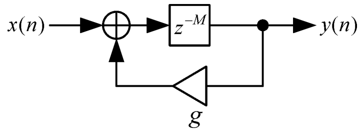


Fig. 3. Feedback comb filter block diagram.

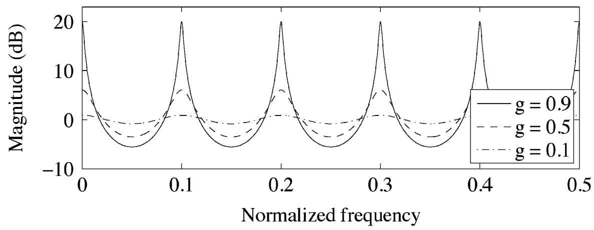


Fig. 4. Magnitude response of the feedback comb filter in Fig. 3 for feedback gains $g = 0.1, 0.5$, and 0.9 and delay $M = 10$ samples.

which can be described as a sparse exponential decay. Reverb designers often listen to impulse responses as an "acid test" of reverberator quality.

When M is sufficiently large, the impulse response of the feedback comb filter $h(n)$ is heard as discrete "echoes" of the input impulse. When M is below the time-resolution limit of the ear, it is more relevant to look at the magnitude frequency response

$$|H(e^{j\omega T})| = \frac{1}{1 - g\, e^{-j\omega MT}} \tag{3}$$

where T denotes the sampling interval in seconds. This is plotted in Fig. 4 for $g = 0.1, 0.5, 0.9$ and $M = 10$. From this figure it can be appreciated that it is the amplitude response of comb filters that gives them their name. It is noted in [1] that each local maximum in $|H(e^{j\omega T})|$ corresponds to a *normal mode* of vibration in the reverberator. In signal processing terms, a normal mode corresponds to a lightly damped conjugate-pole-pair in the system transfer function. A normal mode may be parametrized by its center-frequency, bandwidth, and amplitude. In acoustic spaces, the normal modes can be considered to be superpositions of the various possible *standing waves* [94].[2]

B. Allpass Filters

Also in [1], Schroeder and Logan introduced the use of *allpass filters* for artificial reverberation. By definition, an allpass filter passes all frequencies with equal gain. It is easy to show [95] that the general form of a real, causal, stable, single-input, single-output, digital allpass filter is given by

$$H(z) = z^{-K}\frac{\tilde{A}(z)}{A(z)} \tag{4}$$

where $K \geq 0$ is an integer pure-delay, $\tilde{A}(z) = z^{-N}A(z^{-1})$ is the "flip" (i.e., a mirrored version) of polynomial $A(z)$, and $A(z)$ is any minimum-phase polynomial:

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_N z^{-N}. \tag{5}$$

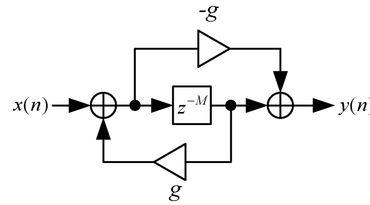


Fig. 5. Signal flow graph of a Schroeder allpass section in direct form II.

We say that $A(z)$ is minimum phase when all of its roots are inside the unit circle. This of course ensures stability of the allpass filter. Note that the zeros of $\tilde{A}(z)$ are reciprocals of the zeros of $A(z)$, which is how allpass filters can be characterized in a pole-zero diagram.

The allpass filters introduced for artificial reverberation by Schroeder and Logan [1] were of the form

$$H(z) = \frac{z^{-M} - g}{1 - gz^{-M}}. \tag{6}$$

A direct-form-II [95] implementation is shown in Fig. 5. The transfer function $H(z)$ can be viewed as a feedback comb filter $H_1(z) = 1/(1 - gz^{-M})$, as discussed in Section III.A, in series with the feedforward comb filter $H_2(z) = -g + z^{-M}$, i.e., $H(z) = H_1(z)H_2(z)$. An excellent benefit of allpass filters of this form is that (1) for large delays M, they sound virtually identical to feedback comb filters, thus providing almost identical echo qualities, while (2) for sinusoidal input signals, there is *no gain variation* across frequency. It is in this sense that allpass reverberators are said to be "colorless."

Free-software implementations of artificial reverberators along the lines outlined by Schroeder *et al.* may be found in the FAUST distribution in the file `effect.lib` [96], [97].[3] Additionally, a popular Schroeder reverb called "Freeverb" is given in the example `freeverb.dsp` within the FAUST distribution, and appears in many other distributions as well.[4]

C. Digital Waveguide Networks

A generalization of allpass filters to arbitrary closed networks of *digital waveguides* was proposed in [81], [82], [98], [99]. A digital waveguide, or bidirectional delay line, can be viewed as a discrete-time counterpart of an electric transmission line in which traveling waves are explicitly propagated. In the lossless, constant-wave-impedance[5] case, the traveling waves propagate without alteration, so that a simple delay line models the acoustic (or electric) medium in each direction. At discontinuities in the wave impedance, so-called *signal scattering* occurs, as shown in Fig. 6. (If the digital waveguides are taken to be one sample long, then the structure of the Kelly–Lochbaum vocal-tract model is obtained, from which ladder and lattice digital filters can be derived [100].[6]).

[2]https://ccrma.stanford.edu/~jos/pasp/-D_Boundary_Conditions.html.

[3]FAUST is a "Functional AUdio STreaming" language that compactly represents signal-processing block diagrams and compiles to C++.

[4]https://ccrma.stanford.edu/~jos/pasp/Freeverb.html.

[5]For transmission lines, the wave impedance is conventionally called the *characteristic impedance*.

[6]https://ccrma.stanford.edu/~jos/pasp/-Conventional_Ladder_Filters.html.

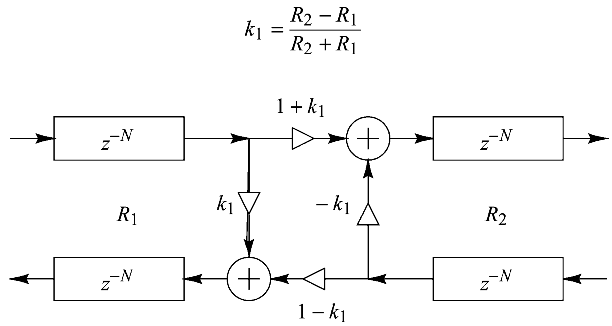


Fig. 6. Two equal-length digital waveguides joined by a scattering junction due to an impedance discontinuity. The wave impedances are R_1 on the left and R_2 on the right, resulting in reflection coefficient k_1 for pressure (or voltage) traveling waves (from [94]).

A key property of scattering junctions is that they are *lossless*. That is, they *conserve signal power*. Fig. 6 shows the scattering junction for two waveguides connected together. It is also straightforward to derive the lossless scattering junction for any number of intersecting waveguides [94]. The practical import is that one may take an arbitrary collection of digital waveguides, having arbitrary lengths and wave impedances, and combine them together in any network topology whatsoever, following the rules of lossless scattering at each waveguide junction. In this way, one always obtains a lossless ("marginally stable") digital network. Such networks are ideal starting points for artificial reverberation systems having long decay times, such as concert hall simulations.

Fig. 6 shows the scattering junction for traveling pressure waves. For velocity waves (the dual variable), simply replace k_1 by $-k_1$ everywhere it occurs [94]. Given this, we can calculate that the power reflection-coefficient seen from the left is $-k_1^2$ (where the minus sign is associated with the direction of travel) while the power transmission-coefficient from left-to-right is $(1+k_1)(1-k_1) = 1-k_1^2$. Thus, the signal power incident on the left side of the scattering junction is split into reflected and transmitted components, without loss or gain (neglecting round-off error). Similarly, the incoming power on the right sees the same reflection and transmission coefficients $-k_1^2$ and $1-k_1^2$. Thus, it is shown that signal-power is conserved by the scattering junction—the power in from the left and right equals the power out to the left and right. This lossless behavior occurs for any number of intersecting waveguides, and it follows simply from the physical constraints that pressure must be everywhere continuous (imposed by Newton's law $f = ma$) and flows must sum to zero (imposed by conservation of matter) [94].

When constructing artificial reverberation systems from digital waveguide networks, one typically first builds a *lossless prototype* using waveguides having lengths comparable to the reflection-free path-lengths in the desired physical space. The waveguide lengths are also typically taken to be mutually prime, as in the delay-line lengths of Schroeder reverberators. Finally, losses are inserted sparsely at arbitrary points in the network to obtain the desired reverberation time in each frequency band. Frequency-independent losses are obtained using constant coefficients $|g| < 1$, and frequency-dependent losses are obtained using filters having gains bounded by 1 at all frequencies (for stability/passivity). The loss filters can be inserted in any traveling-wave path within the network. Input signals may be summed into any points within the network (scattering junctions are commonly chosen), and any samples in the network may be taken as outputs (e.g., sums of scattering-junction incoming or outgoing waves—alternate scattering-junction forms provide such sums for "free" [94]). Karjalainen *et al.* have discussed the design of a waveguide network reverberator for simple room geometries [101].

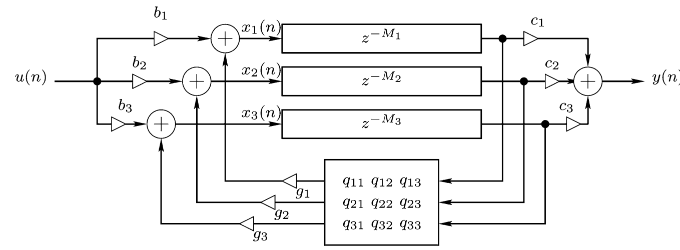


Fig. 7. Order 3 feedback delay network (from [94]).

D. Feedback Delay Networks (FDN)

A *feedback delay network* can be regarded as a "vectorized" feedback comb filter, as shown in Fig. 7 for order three. It can also be derived as a digital-waveguide network in which all waveguide endpoints meet at the same scattering junction [94]. FDNs are typically analyzed using adapted *state space* methods [85]. Similar structures were explored earlier by Gerzon [83] and Stautner and Puckette [102]. Jot and Chaigne developed the FDN approach essentially to its current level of application [85], [86]. The choice of orthogonal feedback matrix is a particularly interesting topic that strongly affects the quality of the reverberation obtained [85], [103]–[105]. Piirilä *et al.* have shown how to produce non-exponentially decaying reverberant responses using two FDNs with slightly different parameters or using other modified comb filter structures [106]. De Sena *et al.* have recently expanded the FDN concept further by incorporating frequency-dependent wall absorption and directivity of sources and receivers (microphones) [107].

Schroeder and Logan introduced an early example of *perceptual orthogonalization* in systems for artificial reverberation [1]. In particular, allpass filters were introduced to provide echo density without frequency-response coloration, allowing easier independent control of each. FDN reverberators introduced further levels of perceptual orthogonalization, such as separating decay time from total energy in each band [85].

Free-software implementations of FDN reverberators may be found in the FAUST distribution in the file `effect.lib` [96], [97]. Of particular note is the reverberator `zita_rev1` which combines Schroeder allpass and FDN reverberation techniques.[7] This high-quality reverberator is examined further in the following section.

E. Case Study: The Zita-Rev1 FDN/Schroeder Reverberator

As described online,[8] *zita-rev1* is a reworking of the artificial reverberator developed for *Aeolus*, a highly regarded pipe organ synthesizer. Both Aeolus and `zita-rev1` are free,

[7]https://ccrma.stanford.edu/~jos/pasp/Zita_Rev1.html.

[8]*http://kokkinizita.linuxaudio.org/linuxaudio/-zita-rev1-doc/quickguide.html.*

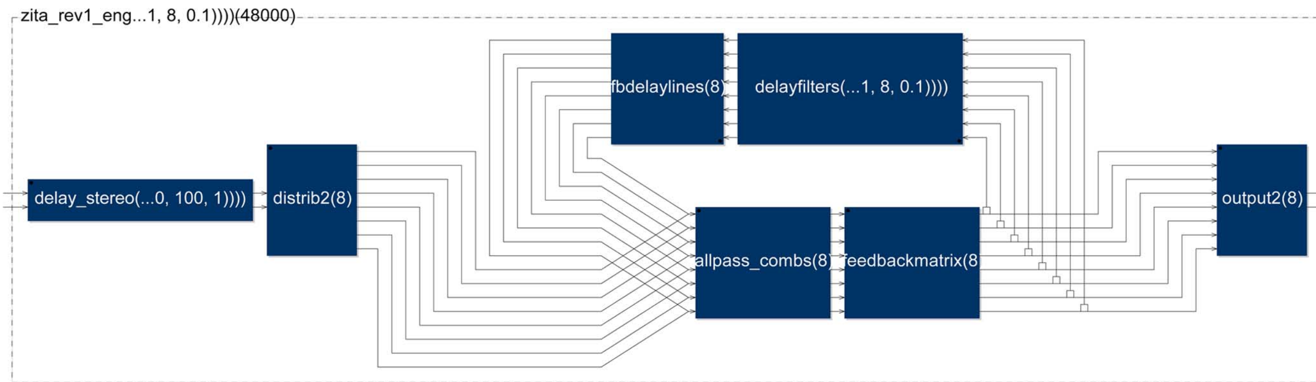


Fig. 8. High-level architecture of the `zita-rev1` reverberator, as drawn by the FAUST compiler using its `-svg` option [96]. Not shown are the dry/wet mix and the two second-order peaking equalizer sections [108] included with `zita-rev1`.

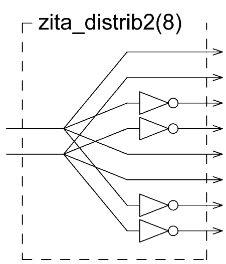


Fig. 9. Stereo to eight-channel distributor matrix `distrib2` used by the `zita_rev1` reverberator, as drawn by the FAUST compiler. The negation symbol here denotes sign inversion (as opposed to its more common use as boolean negation), as indicated by (7).

open-source projects by Fons Adriaensen. A high-level block diagram of `zita-rev1` appears in Fig. 8.

Referring to Fig. 8, a stereo input signal comes in on the left. Denote the stereo input signal by $\underline{x}(n) = [x_L(n), x_R(n)]^T$, where $x_L(n)$ and $x_R(n)$ denote the left and right stereo channel signals, respectively. The first processing block is a pair of equal delay lines that model the propagation delay between the source and listener. If this delay is denoted D in samples, then the delayed stereo output signal is $\underline{x}(n-D) = [x_L(n-D), x_R(n-D)]^T$. Next, the delay-line outputs are distributed among the eight channels of the FDN by the block labeled "`distrib2`" in Fig. 8. If its outputs are denoted $\underline{v}(n) = [v_1(n), \ldots, v_8(n)]^T$, then we have

$$\underline{v}(n) = \begin{bmatrix} \underline{x}(n-D) \\ -\underline{x}(n-D) \\ \underline{x}(n-D) \\ -\underline{x}(n-D) \end{bmatrix}. \tag{7}$$

This is diagrammed (again automatically by the FAUST compiler) in Fig. 9.

Referring again to Fig. 8, the eight-channel distributor output signal $\underline{v}(n)$ next sums with the eight FDN feedback signals. (Summation is implied when signals meet at a node.) The result of this summation goes to a parallel bank of eight Schroeder allpasses (the block labeled `allpass_combs` in Fig. 8), each as drawn in Fig. 5. The coefficient used is $g = \pm 0.6$ for all eight allpasses. (Historically, $g = 0.7$ is more commonly encountered.) The eight allpass delays M_i (see Fig. 5) are chosen to be $M_i = \text{round}\{f_s \cdot [0.020346, 0.024421, 0.031604, 0.027333, 0.022904, 0.029291, 0.013458, 0.019123]\}$ samples, where $f_s = 1/T$ denotes the sampling rate, and $\text{round}(x) = \lfloor x + 1/2 \rfloor$ denotes rounding to the nearest integer. It is interesting to note that `zita-rev1` does not try to preserve the mutually prime property of the M_i recommended by Schroeder and enforced in most other such reverberators.

The output of the parallel Schroeder-allpass bank is fed to an 8×8 orthogonal FDN feedback matrix Q. In general, Q can be any orthogonal matrix ($Q^T Q = I$). However, in `zita-rev1`, as in many other FDN reverberators, a normalized *Hadamard matrix* $Q_8 = H_8/\sqrt{8}$ is chosen. Hadamard matrices H_n, for n a power of 2, can be easily generated by Kronecker products with the basic 2×2 "butterfly" matrix

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{8}$$

Specifically, we have

$$H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix} \tag{9}$$

and

$$H_8 = \begin{bmatrix} H_4 & H_4 \\ H_4 & -H_4 \end{bmatrix}. \tag{10}$$

The resulting recursive butterfly structure is visible in the FAUST-generated block diagram shown in Fig. 10.

The `delayfilters` block consists of eight low-pass filters in parallel. For stability, each low-pass must have gain less than $1/\sqrt{8}$ at every frequency. (The Hadamard matrix normalization is moved into the branch filters.) The implementation chosen in `zita-rev1` consists of a first-order "low shelf" [95] in cascade with a first-order low-pass filter. These filters are controlled by the graphical user interface (GUI) so as to set the reverberation time in three frequency bands. A detailed derivation is given in [94].[9]

Finally, for the feedback loop, the block labeled `fbdelay` is simply a bank of eight delay lines in parallel. Their lengths L_i in samples are chosen so that $L_i + M_i = \text{round}\{f_s[0.153129, 0.210389, 0.127837, 0.256891, 0.174713, 0.192303, 0.125000, 0.219991]\}$, for $i = 1, 2, \ldots, 8$, where M_i denotes the delay used in the eight Schroeder allpasses described previously (block `allpass_combs` in Fig. 8).

[9]https://ccrma.stanford.edu/~jos/pasp/-Zita_Rev1_Delay_Line_Filters.html.

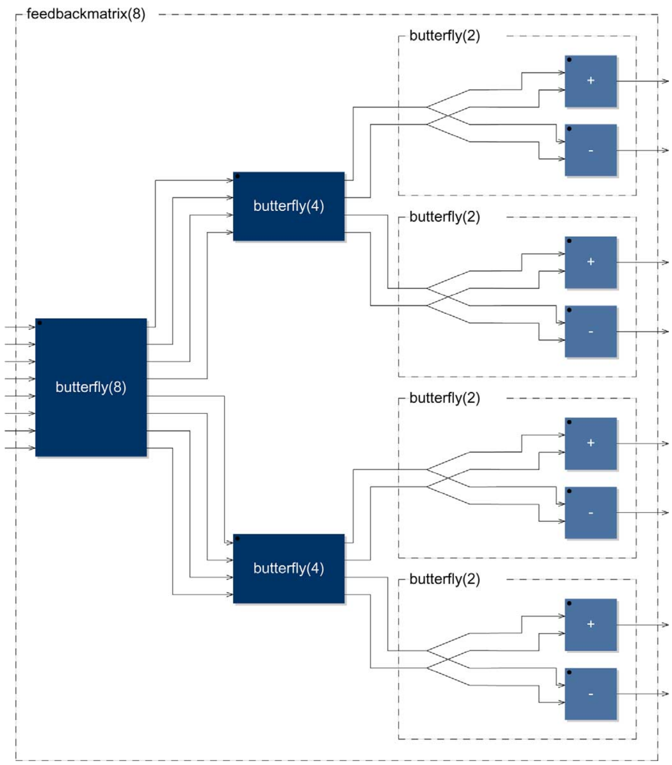


Fig. 10. Eight-by-eight Hadamard feedback matrix H_8 used by the `zita_rev1` reverberator, as drawn by the FAUST compiler.

The output signal $\underline{y}(n) = [y_L(n), y_R(n)]^T$ is extracted from the feedback loop using a so-called *mixing matrix*, in this case, an 8×2 matrix. Let $\underline{w}(n) = [w_1(n), \ldots, w_8(n)]^T$ denote the length 8 vector emerging from the feedback-matrix multiply in Fig. 8. Then the choice made by `zita-rev1` for stereo output is given by

$$\underline{y}(n) = H_2 \cdot \begin{bmatrix} w_2(n) \\ w_3(n) \end{bmatrix} = \begin{bmatrix} w_2(n) + w_3(n) \\ w_2(n) - w_3(n) \end{bmatrix} \tag{11}$$

where H_2 was defined in (8). That is, the stereo output is taken to be a "shuffle" [multiplication by the elementary butterfly H_2, which is in fact a length 2 discrete Fourier transform (DFT)] of channels 2 and 3 of the eight.

Note that the function `zita_rev_fdn` in the file `effect.lib` within the FAUST distribution brings out the internal 8×8 Schroeder/FDN engine separately, as shown in Fig. 11. That is, it is `zita_rev1` as shown in Fig. 8 minus its input delay, $2 \rightarrow 8$ distributor matrix, and $8 \rightarrow 2$ mixing matrix. For four- and eight-channel systems, `zita_rev_fdn` can provide a useful starting point for finding good-sounding distributor and mixing matrices.

F. Time-Varying Reverb Algorithms

It is well known that the acoustic characteristics of any room are under minor but continuous change because of temperature changes, air conditioning, or physical movement of persons or objects, which all affect the propagation delay of sound [109], [110], [13]. This logically leads one to think that reverberation algorithms may also include time-variant elements. In the past, it has been suggested to slowly modulate the length of some of the delay lines over time [31], [87], [111]. The modulation renders the modes wider in the frequency domain and breaks up the temporal regularity of the impulse response. If the delay-line length is simply stepped from one value to another, audible disturbances appear, however. To accurately implement the low-frequency delay modulation, a time-varying fractional-delay finite impulse response (FIR) filter, such as one based on fourth- or higher-order Lagrange interpolation, can be applied [112], [111].

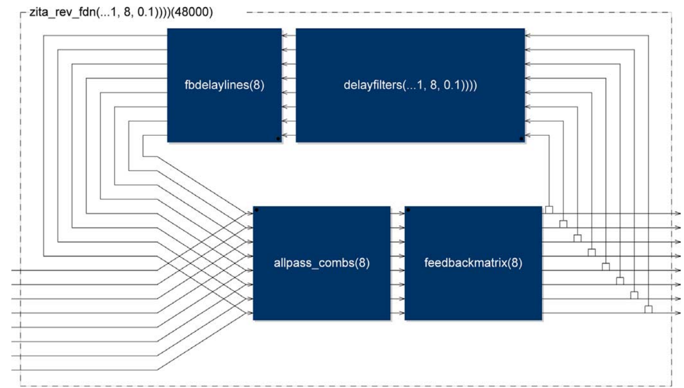


Fig. 11. Zita-rev1 internal eight-channel Schroeder-FDN reverb engine, useful for adapting to audio systems having other than two channels in and out.

An alternative method consists of modulating one or several parameters of the delay network at a low frequency [81], [113], [114]. Lokki and Hiipakka [113] and Choi *et al.* [114] implement part of each long delay line with a comb-allpass filter, as originally suggested by Väänänen *et al.* [115], so they can directly modulate a coefficient in the following difference equation to change the system characteristics with time

$$y(n) = -a(n)x(n) - x(n-N) + a(n)y(n-N) \tag{12}$$

where $x(n)$ and $y(n)$ are the input and output signal of the allpass filter, N is the allpass filter order (the same as the length of its embedded delay line), and $a(n)$ is the time-varying coefficient, which must remain $-1 < a(n) < 1$ for stability. Modulating the allpass filter parameter is technically easier to realize than delay modulation, but has practically the same effect when the modulation depth is small. Choi *et al.* report that this can also help save delay-line memory in FDN reverberators [114].

G. Pseudo-Random Late Reverb Algorithms

In 1998, Rubak and Johansen proposed a radically different approach to simulate late reverberation using a feedback structure: an FIR filter with pseudo-random (constant) coefficients in a feedback loop [116], [117]. A feedback coefficient, which can be replaced with a one-pole low-pass filter, brings about the temporal decay, and the random FIR coefficients must also follow a truncated exponential envelope to avoid discontinuities at repetition points. This filter has an impulse response consisting of exponentially decaying pseudo-random numbers but the underlying pseudo-random sequence repeats every L samples, which is the length of the FIR filter. In a computationally efficient version, the pseudo-random number sequence can be made sparse [117]. It turns out that 2000 to 4000 nonzero pulses per second is sufficient for high-quality reverberation.

May and Schobben [118] developed a related late reverberation algorithm, which is based on an FIR filter with randomly modulated coefficients, applying feedback. This technique first separates transient and steady-state audio signals and only uses the random filtering method for the transient part. The steady-state part is processed with an allpass-filter based reverberator.

Karjalainen and Järveläinen [119] expanded the pseudo-random late reverberation method by introducing the *velvet noise*, which is a very smooth-sounding sparse pseudo-random number sequence. In velvet noise, a -1 or $+1$ appears within each short interval of T_v samples while all others samples are zero. Two pseudo-random number sequences $r_1(n)$ and $r_2(n)$ are needed to produce the velvet noise samples $v(n)$. Both are uniformly distributed on the interval (0, 1). The locations of the pulses are determined as

$$k(m) = \text{round}\{T_v[m + r_1(n)]\} \tag{13}$$

where $m = 0, 1, 2, \ldots$ is an incrementing counter and T_v is the nominal (and average) distance of pulses. The pulse values (-1 or 1) are then drawn using the second sequence $r_2(n)$ as

$$v(n) = 2\text{round}\,[r_2(n)] - 1, \text{ when } n = k(m). \tag{14}$$

The rest of the samples $v(n)$ will be zero. The density of velvet noise is $N_v = f_s/T_v$ impulses per second, when f_s is the sample rate (such as $f_s = 44.1$ kHz).

Velvet noise samples are used as filter coefficients in a sparse FIR (SFIR) filter, which can then be implemented without multiplications [119]. A feedback loop can be used for repeating the noise. To avoid audible repetition, Karjalainen and Järveläinen use a library of velvet noise sequences and switch each SFIR coefficient using cross-fading. Vilkamo *et al.* have developed a subband technique, which contains a pseudo-random SFIR decorrelation filter after a feedback loop on each band [120]. Their decorrelation filters need not be variable with time, because the feedback loop delay is different for each narrow band, making repetitions inaudible.

Lee *et al.* [121] have further elaborated the switched random convolution ideas. For example, they have presented a structure in which a leaky integrator is used for switching only some pulses in the velvet noise sequence. The time constant of the integrator can be made frequency-dependent to avoid artifacts. A reverberation algorithm based on the switched pseudo-random sequences with a sound quality comparable to a 16th-order FDN requires 70% less operations and 90% less delay-line memory than the FDN structure [121].

IV. Physically Based Room Models

Physically based reverberation algorithms aim to reproduce the acoustics of a given real or virtual space. The physics behind sound propagation and reflections is well known, but in practice solving the wave equation for the whole audible bandwidth in large halls is beyond the capacity of current computers. Instead there are several different techniques to approximate that equation. Some of them are actually capable of real-time processing whereas some others only provide room impulse responses as offline computation and those responses are used in convolution reverbs.

Physically based reverberation algorithms are utilized especially in virtual reality systems ranging from fully immersive virtual reality systems to desktop and even mobile environments. The required accuracy of these models depends heavily on the application area. For acoustic design of concert halls or other acoustically challenging spaces the models should be as accurate as possible whereas in computer games reproducing plausible acoustics is sufficient.

In the following, we will give an overview of different room acoustic modeling techniques, with a special emphasis on techniques that are capable of real-time computation. In the end, there are descriptions of some complete virtual acoustics systems targeted mostly for research purposes. They use selected room acoustic modeling techniques and make the obtained modeling results audible, i.e., auralize them [24], [26].

A. A General View on Room Acoustic Modeling Techniques

The main target in room acoustic modeling is to obtain an impulse response between given source and receiver locations. There are several ways to classify room acoustic modeling techniques. The most traditional one is to divide the techniques into two groups. The first is formed by the methods that aim to numerically solve the wave equation and are called wave-based methods. The second is based on geometrical acoustics in which sound is supposed to propagate in rays similarly as light and those are called ray-based methods. However, from a reverberation standpoint, the type of modeling result is essential, and thus in the following we use that classification. There are three different types of modeling results:

- impulse responses that can be directly used in convolution;
- time-energy responses that needs to be converted to an impulse response before convolution;
- list of reflection paths that can be auralized by some filter structure or converted to an impulse response.

It turns out that these two classifications are close to each other such that the wave-based methods typically provide impulse responses directly whereas ray-based methods either produce time-energy responses or information of reflection paths depending on their treatment of diffuse reflections. In the following, different room acoustic modeling techniques are introduced in the three groups listed above. For an earlier review on different room acoustic modeling techniques, see, e.g., [122], [123].

B. Wave-Based Methods for Room Acoustic Modeling

The main advantage of wave-based methods is that they are physically accurate such that modeling of wave phenomena such as diffraction and interference is inherently included in the techniques whereas in geometrical acoustics those have to be modeled separately.

The wave equation can be solved analytically in simple cases such as in a shoebox with hard walls, but for any realistic space that is not possible and a numerical approximation has to be used. The wave-based methods can operate either in the time domain or in the frequency domain. The most common time-domain approach is the finite-difference time-domain (FDTD) technique and its variants whereas the finite element method (FEM) and boundary element method (BEM) are most commonly applied in the frequency domain. This means that in FDTD simulations time has to be discretized, whereas in FEM and BEM the frequency is discretized.

From the viewpoint of reverberation, the FDTD methods are the most interesting ones since they directly provide impulse responses, although FEM and BEM results are applicable in convolution as well if high enough frequency resolution is applied. It seems that obtaining wideband responses with FDTD is currently the favored technique, and thus in the following we concentrate only on the time-domain techniques. Both FDTD and FEM techniques require discretization of the space in addition to time/frequency discretization. This leads to steep increase in computational load such that every doubling of the frequency band induces 16-fold increase in the computational load. For this reason the wave-based methods are typically used only on low-frequency modeling of room acoustics. Until very recently these techniques have been strictly only for offline simulation but recent advances in computer technology have enabled even real-time wave-based simulations for a limited frequency band [91].

Finite-Difference Time-Domain Methods: There are various finite-difference time-domain solvers for the wave equation. The first 3-D FDTD solvers for room acoustics were presented in the mid 1990s [124], [125]. The more traditional approach is to use a staggered grid approach in which both sound pressure and particle velocity are explicitly present in the update equations [125], [126]. This approach has certain benefits over simpler one-variable solvers, mostly related to setting of boundary conditions.

In addition to those one- and two-variable techniques there are other approaches each having different backgrounds. The transmission-line model is one such variant that has been shown to be equivalent to the FDTD formulation [127]. Still another time-domain simulation technique applicable for wave-based acoustic simulation is the functional transformation technique, which has its background in wave digital filters [128], [129]. According to our knowledge there is no thorough comparison of the efficiency of these techniques when applied to room acoustic problems. However, the one-variable FDTD techniques have been of more interest in room acoustic simulations targeting to auralization, and thus they will be in our focus in the following.

The Digital Waveguide Mesh (DWM) is a one-variable FDTD technique [130]. It is based on interconnected digital waveguides that form a rectangular grid to fill the space under study. The original DWM was 2-D and meant for simulation of plates and membranes [131]. The basic structure of a 2-D DWM is illustrated in Fig. 12. There are two alternative implementations for this structure. In the so-called W-mesh (W standing for waves) the actual delay elements are modeled whereas in the K-mesh (K standing for Kirchhoff) the nodes are modeled. Out of those two, the K-mesh is more memory efficient and is thus often preferred although incorporating different boundaries is easier in the W-mesh. However, it is possible to use both approaches simultaneously and connect them via K-W converters [132].

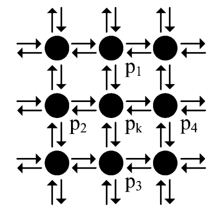


Fig. 12. In a 2-D digital waveguide mesh, each node is connected to four neighboring nodes by bidirectional unit delay elements.

The K-mesh is equivalent to an explicit FDTD scheme obtained by discretizing the wave equation using center-time center-space differences. The resulting update equation to govern the mesh behavior is

$$p_{l,m}^{n+1} = \frac{1}{2}\left(p_{l+1,m}^{n} + p_{l-1,m}^{n} + p_{l,m+1}^{n} + p_{l,m-1}^{n}\right) - p_{l,m}^{n-1} \tag{15}$$

where p is the sound pressure at time step n at location $[l, m]$ [131]. Expansion to three dimensions is straightforward and requires only adding one spatial dimension such that each node has six neighbors instead of four. This technique is efficient as such, but suffers from direction-dependent and frequency-dependent numerical dispersion. The error is most severe at high frequencies where waves propagating in axial directions (relative to the grid coordinates) are delayed whereas there is no distortion in sound propagating in purely diagonal direction [131]. How much dispersion is tolerated depends on application and is still under research [133], [134].

There have been several attempts to overcome the dispersion problem. The main strategies have been use of non-rectilinear mesh topologies [135] and use of larger stencils that use larger neighborhood in the update equation [136]–[138]. The interpolated wideband scheme introduced by Kowalczyk and van Walstijn presents the current state-of-the-art in obtaining uniform wave propagation characteristics in this type of FDTD methods [138]. However, frequency-dependent dispersion is still a problem that can be compensated to some degree by offline frequency warping [136], although warping could be incorporated into the mesh as well [139]. Setting the boundary conditions in a one-variable FDTD simulation has been a challenge for a long time [140], but recent advances have enabled use of digital filters on FDTD boundaries [141]. Even diffuse reflections can be modeled [142], [143]. Getting surround-sound responses from DWMs has been studied as well [144].

Several practical implementations of the DWM technique targeting room acoustic simulation have been presented [145]–[147]. The FDTD techniques are well suited for parallel computation, and for this reason recent development in GPUs have had a clear impact on acoustic FDTD research [148], [91], [149]–[152]. Another option is to use dedicated hardware for FDTD simulations [153].

When considering the frequency limits, it can be said that in any FDTD implementation there needs to be at least six nodes per minimum wavelength to be modeled. The actual required mesh density depends on the chosen scheme and on the amount of dispersion that can be tolerated. In practice this means that to model a room up to 1 kHz (corresponding to the wavelength of 35 cm) the volume should be discretized with a 5.8 cm, or even denser, grid. To reduce the computational load of full 3-D simulation, Kelloniemi *et al.* have proposed to use several interconnected 2-D DWMs that represent different slices of the modeled space [154]. By this means it is possible to have most important axial room modes, and some of the oblique modes, modeled. This means that the lowest end of the model is relatively accurate but higher in frequencies, the mode density is sparser than in reality. In addition, it is possible to combine results from 2-D and 3-D simulations [155]. An opposite proposal is to use hyper-dimensional DWM for reverberation [156]. Equation (15) can be expanded to even higher dimensions such that in a four-dimensional mesh each node would have eight neighbors. Naturally, this kind of model does not have a realizable physical counterpart. However, it enables the use of remarkably small mesh sizes to obtain a high mode density even at low frequencies.

Adaptive Rectangular Decomposition: One recent wave-based modeling technique is the adaptive rectangular decomposition method proposed by Raghuvanshi *et al.* [157]–[159]. It uses a rectilinear decomposition of the space but iterative sound propagation between nodes is performed such that there is no dispersion error at all. For this reason it outperforms all the FDTD techniques if no dispersion is tolerated. In practice, some dispersion can be allowed in modeling result, and what is the optimal modeling technique depends on required accuracy.

C. Case Study: TreeVerb

Many reverberation algorithms attempt to mimic aspects of a physical reverberant environment. One such example is the reverberator presented in [160], designed to simulate the acoustics of a stand of trees in a forest. Unlike typical rooms, in which sound propagates freely in the interior and is reflected at the boundaries, in a forest, reflections are generated in the interior by scattering from trees. As a result, the buildup and decay of energy can be very different than that in an enclosed space, following a more Gaussian rather than exponential envelope.

To model a forest, consider a two-dimensional geometry having a source, listener and a set of trees. The impulse response of the forest involves paths from the source, interacting with the trees, and arriving at the listener. In this model, there are a finite set of path segments traversed: between the source or listener and each of the trees, and between each pair of trees. The propagation between any two points may be modeled by a time delay and spreading loss. Assuming sufficiently separated cylindrical trees, the interaction with an intervening tree may be modeled using a filter representing the scattering from a hard cylinder. In this way, a network of bidirectional waveguides representing propagation between trees, connected via scattering junctions, as seen in Fig. 13, models the reverberation between a given source and listener.

The design of a specific reverberator then amounts to selecting the locations of the trees, source and listener, as well as the tree radii. Tree placement affects the energy envelope and echo density of the resulting impulse response.

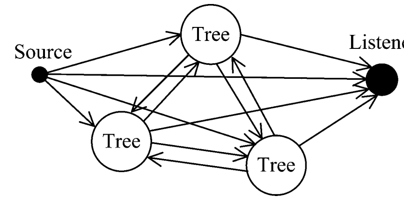


Fig. 13. Schematic overview of the TreeVerb model [160].

D. Geometrical Acoustics-Based Modeling Techniques

Geometrical acoustics (GA) presents quite crude approximation of the wave equation since it neglects the wave nature of sound and instead assumes a ray-like behavior of sound. Despite this, it is a valid approach, especially at high frequencies, where the wavelength of sound is small in comparison to dimensions of a typical modeled space.

All the GA modeling techniques are covered by the room acoustic rendering equation that presents a unifying mathematical framework for all of them [161]:

$$\ell(x', \Omega) = \ell_0(x', \Omega) + \int_{\mathcal{G}} R(x, x', \Omega)\ell\left(x, \frac{x'-x}{|x'-x|}\right) dx \tag{16}$$

where $\mathcal{G} \subset \mathbb{R}^3$ is the set of all surface points in the enclosure, $\ell(x', \Omega)$ is the outgoing time-dependent acoustic radiance from point x' in direction Ω, ℓ_0 is sound emission, and R is the reflection operator for sound coming from point x. The main component of R is the bidirectional reflectance distribution function (BRDF), which determines the reflection type. Basically the GA techniques differ from each other on how they sample the integral inside the equation, and what kind of BRDF is applied. Covering all the variations and improvements presented over the years is beyond the scope of this paper and thus in the following we limit ourselves only to those most relevant to the present topic.

Geometrical Acoustics and Specular Reflections: In any real closed space, the number of reflection paths between a source and receiver locations grows exponentially as a function of reflection order. In practice, this means that detecting all the reflection paths, even specular ones only, is typically not possible for the whole duration of an impulse response. Instead, the number of early specular reflections is limited and computing them in real time is possible. They are perceptually the most important reflections. In making accurate room acoustic modeling and auralization, it is essential to get the early reflections right. This knowledge forms the base of several auralization systems where early reflections and late reverberation are processed separately, as discussed in Section IV-G.

In computing specular reflection paths the reflection operator R of (16) implements the ideal mirror reflection, i.e., such that all the incoming energy is reflected to the specular reflection direction and there is no diffuse reflection component at all. The basic techniques for computing the early specular reflections are the image-source method [78], [79] and ray-tracing [162]–[164]. The image-source method is an accurate method such that it is guaranteed to find all the possible specular reflection paths between two given points whereas ray-tracing is based on Monte Carlo sampling of the reflection paths. Due to the exponential growth of the number of possible reflection paths the image-source method is practical only in finding very low-order paths. Ray-tracing does not suffer from this problem, and it can be used to obtain higher-order reflections. However, it should be noted that the accuracy of modeling decreases with increasing reflection order, as omission of valid reflection paths becomes more and more likely for a receiver of fixed size [165].

The beam-tracing techniques present an improvement over the image-source method by constructing a beam-tree to present possible reflection paths. This approach efficiently limits the growth in the number of reflection paths to be investigated in detail [166]. There are both conservative and approximate versions of this approach. One such conservative approach that guarantees to find all the paths has been presented by Laine *et al.* [167]. Their optimization techniques enable use of beam-tracing in real-time auralization with moving listener and static sound source up to fourth to sixth reflection order depending on the complexity of the space. Another optimization targeting at efficient image-source computation is use of visibility diagrams [168], [169]. Even better performance can be gained if an approximate solution, where there is no guarantee that all image sources are found, can be used [170]. However, it is not always necessary to compute the image sources up to very high order, as it is possible to get good estimates of reverberation based on early image source data only [171], [172]. A special case for the image-source technique is to compute the impulse response of a shoebox-shaped room, which can be implemented efficiently [173], and can be extended to higher dimensions than three [174].

One problem related to a large number of reflection paths is how to ensure they are audible. Performing the required signal processing separately for each path is a tedious process. Tsingos *et al.* have presented techniques to make perceptual clustering and culling for multiple moving sound sources to overcome this hindrance without sacrificing the quality of audio output [175].

Geometrical Acoustics and Diffuse Reflections: The models that have a diffuse component in the BRDF typically produce time-energy responses. There are two ways to utilize them in auralization. First, they can be converted to impulse responses and used in convolution [176], [177]. The second way to use the responses is to compute some acoustic attributes from them and to use them to parametrize some reverb algorithm.

In the radiosity methods the BRDF is ideally diffuse such that there is no specular component and the sound reflection does not depend on the incoming angle of sound [178], [179]. In practice, most of the reflections have both specular and diffuse components. Support for arbitrary BRDFs exists only in a couple of modeling techniques. The basic ray-tracing technique is easy to extend for arbitrary reflections as the traditional mirror reflection model can be replaced by any reflection model [164].

There are several variations of the ray-tracing principle such as sonel mapping [180] and phonon mapping [181]. Both of them are inspired by photon mapping that is a global illumination technique applied in computer graphics [182]. However, the idea of sound particles is even older and has been used, e.g., by Stephenson in his quantized pyramidal beam tracing method [183]. It is worth noting that the term “beam tracing” has two different meanings in the context of room acoustic modeling. In some computer graphics inspired work the term refers to a certain geometric algorithm, such as in [166] whereas in other research it has a wider meaning to describe that in acoustics the concept of rays is ill-advised and the rays should be considered to be volumetric objects such as cones [184] or pyramids [183].

The acoustic radiance transfer (ART) method is another technique that is able to handle arbitrary BRDFs [185]. It has been shown to be capable of real-time room acoustic simulation and auralization for a static sound source and a dynamic listener. Stavrakis *et al.* have presented the “Reverberation Graph” technique to handle complex spaces that contain several interconnected rooms with their own acoustic characteristic [186]. That technique has several similarities to the ART technique such that they both precompute the acoustic responses to a given grid of points and in real time those responses are processed for convolution.

Recently, computer graphics research has had a major impact on finding the most efficient ways to determine acoustic responses [187]–[189]. Some of these algorithms are designed for computer games having dynamic environments where both the sound source and listener can move and, moreover, the environment itself can change.

Diffraction Modeling in Geometrical Acoustics: Diffraction is a phenomenon that must be modeled separately in all geometric acoustic modeling techniques. It is an essential component in complex environments, especially, when there is no line-of-sight from the source to the receiver. There are several approximations. The most accurate are based on the Biot–Tolstoy–Medwin technique [190], but they are too complex for real-time processing and instead can be used to precompute diffraction filters that can be efficiently used in auralization [191]. To lessen the computational burden of diffraction some perceptual culling can be performed [192]. The uniform theory of diffraction is a computationally lighter approximation and it is well suited for geometric acoustic modeling techniques such as beam-tracing [193] and frustrum tracing [188]. Another approach investigated by Tsingos *et al.* is based on Kirchoff approximation of sound scattering from a surface [194]. Their technique is very efficient, but is limited only to first-order scattering and cannot model actual reverberation.

E. Hybrid Models

All the techniques mentioned above have their own strengths and weaknesses, and it seems that there is no single method that would be able to handle the whole audible frequency band for the duration of the whole impulse response. For this reason, many hybrid models have been introduced that combine the best parts of different models into one technique. One of the first hybrid models combined the image-source and ray-tracing methods [195]. The image-source method is typically the most suitable technique for modeling of the early reflections. For that reason it is often used in the hybrids where it can be accompanied for example by radiosity [178] or statistical reverberation [196].

Geometrical acoustics performs poorly in the lowest frequency range, and therefore it is necessary to have a wave-based model as a low-frequency simulator when the aim is accurate modeling. One such approach has been presented by Murphy *et al.* in their model that combines FDTD results and ray-tracing

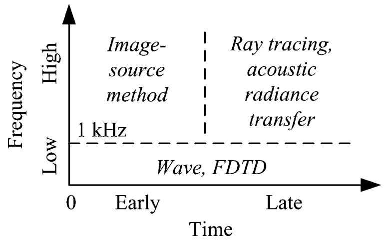


Fig. 14. An optimal hybrid model would have a wave-based model for the lowest frequencies and higher frequencies would be best modeled with a combination of different geometric acoustics techniques.

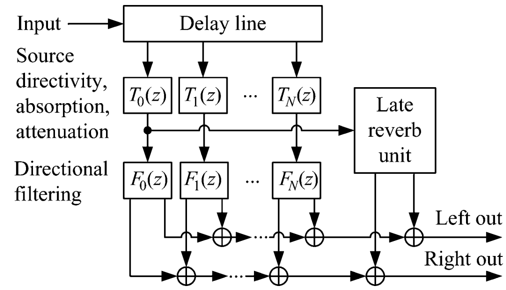


Fig. 15. Example of a filter structure used in an auralization system.

results [197]. However, an optimal hybrid model would make the division of the modeling task based on both time and frequency as illustrated in Fig. 14 [123], [198], [199].

F. Spatialization of Artificial Reverberation

In principle, each reflection in a reverberant acoustic space should be *spatialized*. That is, it should enter one's ear from the correct direction in 3-D space. Spatialized reflections were pursued by Kendall and Martens [200], as they present the basic principles for both headphone and loudspeaker reproduction. The same concepts were later used in 3-D virtual reality systems [201], [25]. More detailed work on spatialization using head-related transfer functions (HRTFs) has been presented by Jot *et al.* [202]. To create convincing headphone reproduction it is essential that the user's head is tracked such that the sound image stays stable even if the user turns his/her head [203]. Loudspeaker reproduction of artificial reverberation has been used, for example, by Gardner [32], in his work targeting to change the acoustics of an existing room. He suggests to record sound in a room and feed it in a real-time room acoustic simulator whose result is played back through loudspeakers.

G. Room Modeling Systems

In the early 1990s there was a boom in room acoustic modeling systems targeting acoustic design. At least the following pieces of software from that time are still commercially available or in inhouse use: EASE and associated auralization tool EARS [204], Odeon [205], CATT-Acoustic [206], Epidaure [207], [177], Ramsete and associated auralization tool Aurora [208]. They are all hybrid models based on geometrical acoustics. In addition to producing acoustic attributes they are able to auralize the simulation results for listening purposes.

One of the pioneers in real-time acoustic simulation has been NASA where spatial hearing has been a major research topic [201], [209]. In addition, there have been several more academic systems targeting real-time purposes whereas the acoustic design tools are mostly for offline simulation in static environments. One of the first such systems has been presented by Nakagawa *et al.* [210]. The DIVA auralization system developed at the Helsinki University of Technology (now part of the Aalto University), Finland, shares several similarities to their system [211], [25], [212]. In the DIVA system an image-source implementation provides a list of specular reflections, which are processed by an auralization module. It contains a separate set of filters for the direct sound $(T_0(z))$ and for each early reflection $(T_1(z)\ldots T_N(z))$ and an FDN type reverb structure, as illustrated in Fig. 15. The filters incorporate air absorption, distance attenuation, and reflection properties for each reflection path, which is individually spatialized $(F_0(z)\ldots F_N(z))$. Their new beam-tracing system, which is optimized for real-time walkthroughs and auralization, has been made publicly available [167], [213].

A similar approach has been utilized in the IKA-SIM system developed at Bochum University, Germany, as well [214]. Much research on real-time virtual auditory environments has been conducted at the RWTH, Aachen, Germany, where the RAVEN system has been developed [215]. Their latest works have investigated efficient acoustic simulation and auralization techniques [216], [217].

In current computers, real-time multichannel convolution is possible, and that opens up a possibility of a different approach to auralization. In the systems listed above, the early reflections and the late reverberation were treated separately, but that is no longer necessary if all the available resources can be used for auralization as is the case in acoustic design applications, whereas in computer games resources have to be shared with several tasks. This approach has been taken in the reverberation graphs [186] and in the frequency-domain acoustic radiance transfer method [185] presented above. In both techniques the whole responses are first computed for a set of predefined receiver points, and are then convolved with a dry excitation and HRTF filters in real time.

H. Recent Research Trends in Room Acoustic Modeling

Modern GPUs are massively parallel computing engines. This makes a huge difference in certain acoustic modeling and signal processing tasks [93], [218]. The trend started in early 2000s when GPUs were used to speed-up acoustic ray-tracing [219], [220]. Even larger performance gains have been achieved in FDTD simulations as mentioned earlier. In signal processing, GPUs can be used for efficient convolution and Fourier transforms [92].

The divergence of application areas for real-time acoustic simulation has created different needs for the required accuracy of the models. At one end, there are computer games in which it is sufficient to present a plausible approximation as long as it is computationally efficient and fits in the given computation budget, see, e.g., [221]. At the other end, there is acoustic design of concert halls where accuracy has crucial role. Recent studies have shown, e.g., that the role of phase in early reflections has been underestimated in current acoustic design [222]. This means that in the future the role of auralization will increase since those effects are something that traditional energy-based acoustic attributes are not able to capture, and the applied models need to be able to incorporate them.

V. Convolution Techniques

The implementation of a convolution reverb is equivalent to FIR filtering an audio signal with the impulse response of a space [223], [88]. The impulse response can be obtained by conducting a measurement in a real room or by taking it from a computer simulation, as described in the previous section. The impulse response values are then assigned to be the coefficients of the FIR filter, which may be a very high-order filter. In theory, this seems easy. The FIR filter does not even have any processing delay, as the output signal can be computed sample by sample as input samples are supplied.

However, an immediate drawback of the convolution reverberation technique is its computational load, which is huge in practice. For example, when the length of a room impulse response (RIR) is 1.0 s and the sample rate is 48 kHz, the convolution requires 48 000 multiplications and additions per output sample. This alone leads to about 5 GFLOPS (billion floating-point operations per second). Usually the two stereo channels have independent impulse responses, and it is becoming a standard to use more channels, such as five or six, as required in the multichannel surround-sound reproduction systems such as the "5.1" standard. The computational load increases linearly with the RIR length and number of audio channels, so in practice several dozen GFLOPS of computing power must be available. Luckily, processors have grown to be very fast, and for example current GPU processors can perform hundreds of GFLOPS that can be efficiently used in convolution [224], [92].

Reducing the computational load of the convolution technique has been an active research line during the past few decades. In the following we discuss the various aspects of obtaining and utilizing impulse responses for artificial reverberation.

A. Measuring Room Impulse Responses

Playing a unit impulse from a speaker and recording the response is not a practical way to obtain a concert hall's impulse response. The basic limitation of this simplistic approach is that there is not enough energy to excite a wide range of frequencies for having a good signal-to-noise ratio (SNR). Traditional practical methods include shooting a start pistol, clapping one's hands, and popping an air balloon on the stage, where sound sources are usually located [225], [226]. A drawback shared by impulses produced by all these methods is the lack of energy at low frequencies. A pistol produces a very short impulse, which must then be very loud to contain enough acoustic energy for good SNR [225]. Impulses produced by hand clapping may not include enough energy and they contain a resonance in the middle frequency range, usually between 500 Hz and 2 kHz, with a 3-dB bandwidth in the 100–300 Hz range [227].

To make the air balloon pop method repeatable, a fixed breaking mechanism must be constructed. Additionally, the balloons must be quite large, such as 40 cm in diameter, to produce sufficient low-frequency energy and to have a uniform directivity [226]. Abel *et al.* have proposed signal processing techniques to estimate the impulse response from an air balloon measurement [228]. With the current knowledge, balloon pops have become an easy and useful method to obtain the impulse response of a space. This method comes in handy where more sophisticated measurement equipment is unavailable or cannot be used, particularly in ancient ruins or caves.

Since the year 2000 the standard method for high-quality RIR measurements has been to use a sinusoidal sweep [225], [90], [229]. A logarithmic sweep of constant amplitude is played into the room and is recorded. The measured signal is then convolved with the time-reversed version of the sweep signal used during the measurement. The result will be the impulse response of sound propagation from the loudspeaker to the microphone, including all reflections and the full late reverberation, assuming that the recording time has been sufficiently long to capture it all. The SNR can be improved by either amplifying or by lengthening the sweep. Clipping caused by excessive signal amplitude may be detected by inspecting the spectrogram of the recorded signal: harmonic distortion brings about multiple ghost sweeps above the fundamental [90].

During the two decades before the prosperity of the sweep method, it was common to use the maximum-length sequence (MLS) method [230], [231]. It was introduced by Schroeder [232] and was based on a pseudo-random noise sequence, which was used as an excitation signal. A deconvolution algorithm based on number theory solves the impulse response from the measured signal. A reason for the decline of this technology was the observation that, under silent ambient noise conditions, the sweep method can reach an SNR which is 15 to 20 dB higher than that obtained with MLS [229], [233]. However, if the room is noisy during the measurement, the MLS method still is a competitive technique, as it is more robust against impulsive noise than the sweep technique [233].

B. Post-Processing of Room Impulse Responses

Measured RIRs often contain noise, which compromises high sound quality. The noise is caused by the ambient noise in the room, the time-varying nature of room acoustics, and—to a small extent—thermal noise from the microphones and amplifiers. The measured RIR sinks into the background noise, when it has decayed enough to reach the noise level, and after that the response does not seem to decay anymore but its level is saturated. For the convolution reverb application, it is necessary to post-process the RIR to avoid this effect.

Jot *et al.* have proposed a technique for extending the RIR after the point where the measurement noise begins to dominate [234]. The late part of the RIR is analyzed using the short-term Fourier transform and the energy decay relief is determined for each frequency bin signal. After the point at which the decay saturates, each subband response is replaced with exponentially decaying noise. Bryan and Abel [235] have developed a related algorithm, which uses a filter bank for RIR analysis and synthesizes similar noise for each subband to extrapolate the RIR. Ben-Hador and Neoran [236] describe a fade-out technique to extrapolate the measured RIR. Bryan and Abel [235] estimate the temporal energy profile of the RIR and carefully scale down the end part of the RIR to avoid the saturation effect caused by noise [235]. In another work, van Dorp Schuitman and de

Vries have proposed an algorithm for removing noise and position-dependent fluctuations from RIRs using psychoacoustic techniques [237].

It is possible to further process a recorded RIR to change its properties. Possibilities include time-stretching to shrink or extend the RIR duration, or applying an exponential envelope to affect the decay rate of the response [236]. To avoid destroying the early part of the RIR, it may be necessary to decompose the recorded RIR into an early and late part before manipulations [236].

C. Fast Convolution Techniques

The number of operations required by convolution in a von Neumann computing architecture[10] can be significantly reduced by employing a Fast Fourier Transform (FFT) algorithm, such as a split-radix power-of-2 Cooley-Tukey FFT. The Cooley–Tukey FFT algorithm was developed in the 1960s for computing efficiently the DFT of a signal without sacrificing accuracy. The basic FFT algorithm was apparently first devised by Gauss [238], [310]. The convolution theorem states that the convolution of two length N signals can be implemented by multiplying their DFTs (i.e., complex spectra) and by applying the inverse DFT to the result. The number of operations in the time-domain convolution is proportional to N^2, but the cost of the FFT-based convolution is proportional to $N \log_2 N$, where N is the length of the two sequences. This makes an enormous difference, particularly when the signals are long (i.e., N is large).

In the fast convolution algorithm proposed by Stockham, the convolution of an input signal and an impulse response is computed in blocks using the overlap-add method employing the FFT [239]. Unfortunately, a single-block FFT-based convolution causes too much latency to be practical in real-time audio processing. A buffer of N samples must be first filled and only then the FFT, the frequency-domain multiplication, and the inverse FFT (IFFT) can be executed. This leads to a minimum latency of N samples. Additionally, the computation takes some time, which adds to the latency. The optimal strategy minimizing the number of operations per sample spreads the operations equally over the next N sampling intervals. This leads to a total processing delay (latency) of $2N$ samples, which is prohibitive in any real-time application [223], [202], [240]. For example, when the RIR length is 1 s, then the latency is 2 s. To reduce the latency, the impulse response must be processed in segments, when an FFT is used.

Kulp [241] suggested to divide the RIR into several short segments of the same length and then convolve every input buffer with each RIR segment in the frequency domain. This reduces the latency and, as Wefers and Vorländer [242] have observed, also decreases the computational cost. The FFTs of the RIR blocks can be computed offline and stored. The real-time processing involves buffering the input signal, computing the FFT, delaying the FFTs of previous blocks using frequency-domain delay lines, multiplying the complex spectra, adding the result spectra, and computing the IFFT, as shown in Fig. 16. The FFT and IFFT must be twice as long as the buffer length (or larger, such as the next power-of-2) to avoid temporal aliasing. Zero-padding is used to extend the input signal buffer to the required length [243]. This processing will generate $2N$ samples of output data, which can be combined using the overlap-add or the overlap-save technique to produce the output signal [243].

[10]A *von Neumann architecture* refers to a sequentially executing CPU, unlike GPUs which spawn many parallel threads of execution. In a GPU, FFT algorithms are only needed for very large lengths [92].

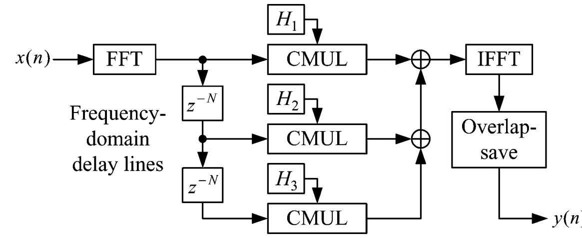


Fig. 16. Convolution of the input signal with a RIR partitioned into three equal-length segments whose spectra are denoted by H_1, H_2, and H_3 (adapted from [251]). The blocks "CMUL" refer to complex multiplication.

Gardner [223] and McGrath [244] proposed to divide the RIR into several segments of different lengths. This means that several FFTs and IFFTs of different length must be executed during real-time processing. The first two blocks in the beginning of the RIR must use the shortest block length N_1. This leads to a processing delay of $2N_1$ samples, which may be tolerable when N_1 is selected to be small. For example, at the 48-kHz sample rate, the choice $N_1 = 32$ leads to a latency of 13 ms (64 samples).

In order to eliminate the latency, the very beginning of the RIR before the first block can be implemented as an FIR filter using the standard time-domain convolution [240], [223], [244]. This part can be made a short, sparse FIR filter [69], [119], which can have a reasonable computational cost. If the initial time gap can be extracted from the RIR, it can be implemented as a pure delay line.

Garcia [245], [246] has pointed out that it pays to use a small number of different block lengths in a block FFT convolution implementations, because this reduces the number of inverse FFTs. All spectra of the same length can be combined in the frequency domain and be converted to the time domain using a single IFFT. For example, just two different block lengths may be selected, the first sufficiently small to ensure a short latency and the second as large as possible to maximize the saving offered by the FFT. Only two IFFTs of different length are needed in this case. Although the algorithms using multiple FFT block sizes are theoretically more efficient than those using only one block size, Torger and Farina [247] have shown that in a standard PC, the one-block-size FFT implementation is the fastest one. A drawback is that it causes more latency than the more sophisticated algorithms, which employ multiple block sizes, first short blocks and then longer ones.

Reijnen *et al.* [240] have suggested another approach to reduce latency of the block FFT convolution: they allow blocks to overlap with the previous block so that only B new samples are taken in. This leads to a total latency of $2B$. Hurchalla has proposed to implement low-latency convolution by dividing the RIR into equal-length blocks and using a two-dimensional FFT [248]. Recently, Kuk *et al.* [249] have derived a fast block FFT method using a new quarter-DFT technique, which reduces the redundancy of the overlap-save method and uses 50% shorter transformations.

The block FFT methods can be extended for time-varying responses, which are needed in interactive systems, when the desired response changes over time. Pérez *et al.* have proposed a method to simulate a moving source in violin synthesis by dividing the violin body impulse responses into several segments, which can be crossfaded to other segments, when the angle to the listener changes [250]. The output signal is synthesized using the overlap-add method. The same technique can be applied to room acoustic simulation.

The best performance of the block-FFT convolution can be obtained by knowing the processor hardware and optimizing the number of different FFTs and the block sizes accordingly [247], [242], [251]. This requires practical benchmarking comparing different choices. Benchmarking tests using a GPU processor showed that the FFT-based convolution is faster than the time-domain convolution only when the RIR length is larger than 100 000 samples [92], although the FFT algorithm can benefit much from the GPU implementation [252]. In both CPU and GPU convolutions, memory access patterns affect the performance, and with optimized cache usage, remarkable performance gains can be achieved [253].

D. Using a Scale Model's Impulse Response

It has been the dream of acousticians since the beginning of the last century that one could obtain a "real" audible impression of a concert hall during the design phase. In 1934, Spandlöck in Munich, Germany, tried to realize this with the help of a fifth-scale model [254]. For this purpose, he used the gramophone technique available at the time. First an anechoic audio signal was recorded on a wax cylinder, and it was then reproduced in the scale model using a loudspeaker, but the rotation speed of the wax cylinder was now five times higher than that of the original recording. Inside the scale model the sound was picked up by a microphone and was recorded on another wax cylinder with the same high speed. Using an earphone and by playing the wax cylinder at normal speed, Spandlöck could listen to a sound which was equivalent to that of a much larger room with the same geometry. A BBC report from the early 1970s describes how to do the same with a special tape recorder [255]. Using digital signal processing techniques, it is now possible to measure the RIR in a scale model and use that for auralization [256].

Spratt and Abel have proposed a related real-time reverberation method employing a scale model of a room, such as a soundproof piece of plastic tube in which a tiny loudspeaker and a microphone have been installed [257]. Their method divides the input signal into short frames, which are played at an appropriate higher speed into the scale model, then recorded, and finally scaled back to normal speed and synthesized using the overlap-add method. This technique does not require FFT processing and does not produce much latency, because the output signal is ready to be played as soon as it has been picked up by the microphone.

E. Parametric Modeling of Room Impulse Responses

Another way to avoid the use of the FFT is to convert the large FIR filter required in convolution reverbs into an IIR system. There are several possibilities to obtain computational savings this way. First of all, various techniques exist to convert an FIR filter to IIR filters, such as linear prediction [100], Prony's method [258], the Steiglitz–McBride iteration [259], and balanced model truncation [260], [261]. It is typical that an IIR filter closely approximating an FIR filter can be implemented with a smaller number of operations than the original FIR filter. A limitation of most of these techniques is that they can only handle fairly short FIR filters, for example up to 1000 coefficients only.

One way to enable IIR modeling of long responses up to several seconds is to divide the RIR into subbands, which are optimally downsampled, and then separately model each of them [262]. When the subbands are narrow, the length of the RIR on each subband will be short, because it is scaled by the downsampling factor. This leads to a kind of filter bank approximation of the original RIR. The *Frequency-Zooming ARMA* (FZ-ARMA) method is one suitable algorithm [262].

In subband techniques, the saving comes from the fact that usually RIRs have a highly frequency-dependent decay rate: energy vanishes sooner from high frequencies than from low frequencies. It is most advantageous to apply the FZ-ARMA method for the lowest frequencies, such as below 200 Hz only. Then the frequencies below this cutoff will be modeled efficiently with an IIR filter, while the residual impulse response above the cutoff can still be an FIR filter [262]. The residual FIR filter will have a shorter length than originally, because the lowest frequencies decaying slowly have been suppressed.

The idea of approximating the impulse response on subbands can be taken to an extreme by using the filter bank decomposition, as proposed by several authors [263]–[266]. The RIR can be divided into subbands, for example using the wavelet decomposition [264], and then each subband RIR can be critically downsampled and truncated. This leads to great computational savings over the standard long convolution, even when the subband signals are chosen to be complex [265]. Marelli and Fu [267] have proposed an optimized subband technique, which they have applied to short impulse responses, such as head-related transfer functions, but which should be applicable to long RIRs as well. Their comparison with other techniques shows that the subband method is superior to block-FFT and parametric pole-zero models both in terms of latency and computational cost [267]. Menzer and Faller have studied the synthesis of binaural RIRs and the replacement of the reverberation tail with Gaussian noise matching the energy-decay relief and interaural coherence [268].

Bai *et al.* have developed a subband technique in which each subband contains an IIR filter structure similar to those used in parametric artificial reverberators (see Section III) [269]. A genetic algorithm is used for optimizing the parameters so that the best approximation of the target response, a recorded RIR, is obtained. Subband methods offer the additional advantage that the decay rate of each subband can be parameterized independently by multiplying the subband signal in the synthesis stage with an exponential envelope [265]. Vickers *et al.* proposed a related algorithm in which the reverberation process is approximated by accumulating and attenuating the magnitude spectra of input signal frames [270].

F. Combining Convolution and Parametric Methods

Recorded RIRs are known to contain the perceptually important details in their early part but the late part resembles exponentially decaying colored noise [20]. However, Jot [271], [234] has shown that the FDN reverberator can be calibrated to approximate the late decay part of RIRs in a convincing manner: The general time–frequency characteristics of the tail are preserved this way, although details of RIR are not reproduced. Stewart and Murphy [272] proposed a hybrid structure in which a short FIR section exactly reproduces the reverberation onset, while an efficient FDN structure generates the late reverberation. The convolutional section may be taken directly from the measured impulse response onset, and the equalization and decay rates of the FDN designed to match those of the measured RIR tail.

The design of the hybrid structure involves decisions on the duration of the convolutional section and the cross-fading method between the convolutional and FDN sections. Greenblatt *et al.* [273] suggest to choose the convolutional onset duration as the time at which the measured and FDN impulse responses achieve the same normalized echo density, a quantity measuring closeness to Gaussian statistics, which has been shown to predict perceived echo density [274]. Greenblatt *et al.* [273] also show that it is advantageous to form the convolutional section by windowing the difference of the recorded and FDN impulse responses. Menzer recently published the idea of modeling the early and the late part of the RIR using two parallel FDNs [275].

G. Multi-Channel Convolution Reverberation

To convincingly capture the acoustics of a concert hall or other space requires a multichannel impulse response, including directional reflections [276]. The *Spatial Impulse Response Rendering* (SIRR) technique can be used for reproduction of measured RIRs with a multichannel sound system [277]–[280]. The RIR is measured using an ambisonics B-format microphone, which outputs four signals: an omnidirectional signal and three mutually orthogonal dipole signals. Using short-term FFT analysis, the intensity and angle of the directional components of the RIR are computed. The nondirectional part of the RIR is assumed to be diffuse, i.e., not coming from any particular direction but rather from all directions.

In the synthesis stage of the SIRR method, the omnidirectional RIR is decomposed and rendered to the different output channels based on the analysis data. The diffuse part of the response can be processed by using FFT-based phase randomization, i.e., having a random phase but the magnitude response of the original RIR, or by using decorrelation filters [279]. A different decorrelated version of the RIR is then used for each channel to produce the impression of a diffuse field. The impulse responses synthesized using the SIRR method can be used in a multichannel convolution reverb to synthesize the spatialized reverberation of a concert hall, where the RIR measurement was conducted [236].

Hulsebos *et al.* have studied capturing of multichannel RIRs using linear, cross, and circular microphone arrays [281]. Li *et al.* suggested simultaneous measurement of multiple RIRs using an array of seven directive microphones [282]. They then separate the early and late parts and use the FFT to run the required convolutions for five output channels. Kuster has proposed processing the different channels of a multichannel RIR with filters designed to produce a prescribed coherence [283]. The method is applicable to the late part of the RIR only.

VI. Virtual Analog Reverberation

Before the easy availability of digital signal processing-based artificial reverberation, studios and performers relied on methods of producing a reverberation-like effect via electrical or electromechanical means (see Section II-A). These primitive types of artificial reverberation had a quality that was distinct from the sound of a standard acoustic space, and were deemed preferable for some applications. Research towards understanding the special sound characteristics of these effects and reproducing them via digital signal processing is therefore desirable. For similar reasons, analogous work has been produced dealing with vintage analog sound processing and generation equipment such as analog synthesizers [284]–[287], compressors [288], [289], and guitar distortion/overdrive circuits [290]–[292]. This area is generally referred to as "virtual analog" modeling [293], [294] and therefore we adopt this moniker also for the study of early artificial reverberation devices.

In the following section, we examine the main early methods of generating artificial reverberation, and how these methods can be reproduced with digital signal processing. Firstly we deal with electromechanical reverberators—those which produce a reverberant effect by exciting vibrations in a resonant object (almost exclusively metal plates and springs) and receiving these vibrations at another point via a transducer. Second, we deal with methods of producing echoes or reverberant effects via electrical means.

A. Electro-Mechanical Reverberation

Electromechanical reverberators can be thought of as behaving similarly to acoustic spaces, with two important distinctions.

- The vibrations propagate through a medium with different properties than air. The common electromechanical reverberators are made of metal, resulting in a speed of sound which varies with frequency and is generally much faster than that in air. This frequency-dependent-speed property is known as dispersion, and alters the character of the individual reflections that make up the reverberant response.
- The geometry of the medium in which the vibrations travel is generally unlike the geometry of an acoustic space, and hence produces a different pattern of reflections. The common geometries are a metal plate or foil, or a helical wire spring. These geometries behave as a lower dimensional space than a standard acoustic cavity (albeit with

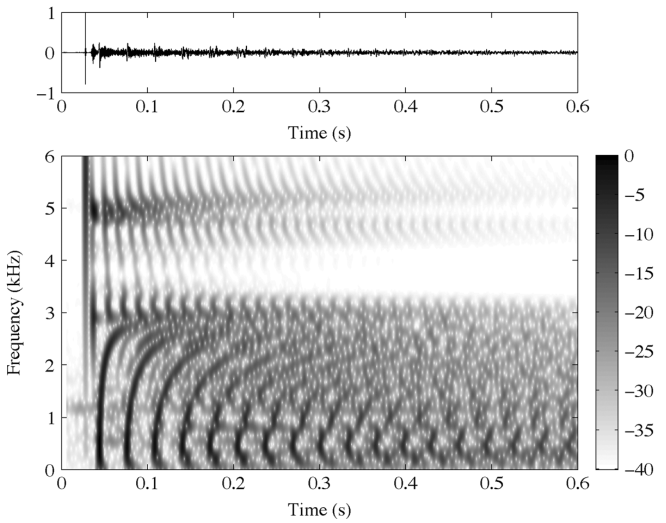


Fig. 17. Impulse response of a single spring from a Belton MB3BB2C1B reverberation tank.

non-Euclidean geometry in the case of the spring), and hence produce a simpler pattern of reflections.

These differences are responsible for the shortcomings of electromechanical reverberators when used as a substitute for an acoustic space, but also define the special sound which makes these reverberation techniques desirable in particular aesthetic applications.

1) Spring Reverb: The response of a spring reverberator to an input signal is characterized by the presence of a number of dispersive echoes, or chirps. Fig. 17 shows a spectrogram of such a response. These echoes can be divided broadly into two groups—first, a set of low-frequency echoes having a minimum delay at low frequency and a maximum delay at a point termed the transition frequency, F_c [52], [295], [296]. This frequency generally lies somewhere in the range of 2–5 kHz. Second, there is a group of wide-band echoes which cover the entire audible frequency range. These echoes are of a much lower amplitude, and possess a minimum delay at high frequency and a maximum at low frequency. The low-frequency echoes are separated by a constant delay, as are the wide-band echoes. The delay separating the low-frequency echoes at dc is termed T_D [52]. The low-frequency echoes seem to dominate the perceived sound of spring reverberation, and the parameters F_c and T_D describe much of the sound of a particular spring.

Analysis of mathematical models of spring vibration [297], [298] has allowed these perceptual parameters to be derived in terms of the physical parameters of a particular spring [52]:

$$T_D \approx \frac{4LR}{r\sqrt{\frac{E}{\rho}}} \tag{17}$$

where L is the uncurled length of the spring wire, R is the radius of the helix, r is the radius of the wire, ρ is the density of the material and E is Young's modulus of the material. The parameter T_D gives a measure of the delay time of the spring, specifically the time taken for a wave to propagate the length of the spring twice. An expression for the transition frequency F_c can also be derived:

$$F_c \approx \frac{3r\sqrt{\frac{E}{\rho}}}{16\sqrt{5}\pi R^2}. \tag{18}$$

Both sets of echoes appear to become more diffuse as they repeat, contributing strongly to the reverberant sound of the spring. More subtle details are visible in this response, and in others—notably the presence of faint pre-echoes before the main low-frequency chirps [299] and breaks or kinks in the echo that correspond to a transition point between different sets of modes [300].

Like the emulation of acoustics spaces, emulation of the response of a spring reverberator can be approached in a number of ways. The simplest approach is to measure the impulse response of a particular unit, and apply it to signals using convolution methods like those discussed in Section V. This method is perhaps more applicable in the case of spring reverberation than room reverberation, as the input and output points of the system are fixed and do not need to be varied. However, convolution methods offer little flexibility in tuning the sound of the effect to suit the application. Methods which allow some parametric variation are therefore desirable.

Abel *et al.* [50] presented the first model of spring reverberation, based on digital waveguide methods (as described in Section III-C). Dispersion was modeled by measuring the response of a real spring reverberation unit, isolating a single reflection, and optimizing a high-order allpass filter to fit the required dispersion curve. This method allows some flexibility with respect to the time between individual echoes, but dispersion characteristics are fixed to match those of the measured reverberation unit.

Bilbao and Parker [295], [52] proposed a model which is based on discretization of a simplified version of mathematical models of spring vibration. The system of partial differential equations is turned into an implicit finite difference scheme. The result accurately models the dispersion of the spring, along with the pattern of echoes, based on physical parameter values. However, the model does not reproduce the increasing diffuse quality of echoes with time seen in real springs. This work shows promise, but is currently not ideal for real-time application due to the computational cost of the methods employed.

Recently, Välimäki *et al.* [299] produced a spring-reverberation effect within the paradigm of traditional delay-network reverberation techniques. The structure consists of two parallel delay lines, with the addition of dispersion-generating allpass filter cascades [301] (known as spectral delay filters) and feedback. Modulation of the delay lines with a filtered noise signal is used to produce progressive diffusion of echoes. The allpass cascades employed can be computationally heavy, but efficiency can be improved by downsampling, altering the chirp-generation structure, or implementing the allpass sections in their most efficient form [302], [303]. Gamper *et al.* have investigated automatic calibration of this kind of model, based on measured

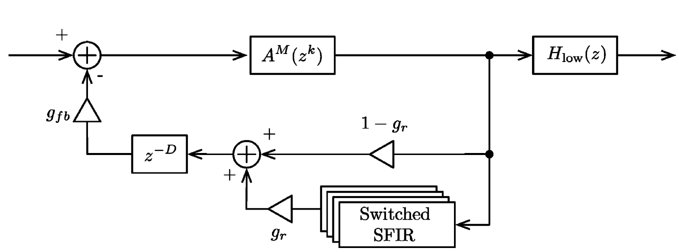


Fig. 18. Block diagram of the new spring reverberation algorithm.

spring reverb responses [304]. Hypothetical reverberators constructed using networks of large springs have also been investigated [305].

2) Case Study: A Parametric Spring Reverberation Algorithm Utilizing Sparse Noise: Generalizing from the structure described by Välimäki *et al.* [299], we can envision a generic form for digital spring-reverberation effects. As in the work of Välimäki *et al.*, this structure consists of two parallel structures, one of which produces the required low-frequency echoes below F_c and one which produces the faster wide-band echoes. Inside each of the structures is a dispersion-generating allpass filter cascade, some form of standard reverberation structure, and a delay line. The allpass filter generates the required chirp shape, the delay-line produces the required delay between echoes, and the reverberation structure produces the progressive blurring or diffusing of the echoes as they recirculate through the system. Any form of reverberation structure can be used to perform this task, and the aesthetic quality of the spring reverberation effect will vary accordingly. The delay-line modulation employed by Välimäki *et al.* can be thought of as a basic form of reverberation, similar to that employed by the Ursa-Major Space Station [73].

As an illustrative example, we propose a new spring reverberation effect designed to model only the behavior of the low-frequency chirp sequence below F_c. Neglecting the wide-band chirp sequence can be justified on the basis that it is present at a much lower level than the low-frequency chirp sequence, so that it has lesser perceptual importance [299]. The effect consists of a feedback structure containing a dispersion-generating allpass chain as described by Välimäki *et al.*, an allpass interpolated fractional delay line, and a SFIR structure with pseudo-random coefficients calculated according to the "velvet noise" technique [119]. Velvet noise is discussed in more depth in Section III-G. The SFIR is updated with new coefficients periodically, with these sets of coefficients being taken from a precompiled library. Each set of pseudo-random coefficients in the library also contains a precomputed scaling coefficient, so that the overall gain of the feedback structure does not vary. The feedback structure is followed by a tenth-order elliptic low-pass filter, designed to suppress activity above F_c. A block-diagram of the complete effect is shown in Fig. 18.

Varying the length of the velvet noise sequence and its relative gain compared to the dry feedback path allows the user to tweak how quickly the echoes become diffuse. The perceptual parameter F_c is varied by altering the dispersion-generating allpass chain $A^M(z^k)$. The parameter T_D is varied by changing the length of the delay line z^{-D}, after taking into account the inherent delay-time of the allpass-chain. Fig. 19 shows an example impulse response produced by this structure, with F_c set to 3 kHz and T_D set to 50 ms. The SFIR is using a library of velvet noise sequences of 1-ms length with a pulse density of 4000. A new sequence is chosen for the SFIR every 256 samples. The parameter g_r is set to 0.005, and g_{fb} is varied to change the reverberation time of the effect. The value of g_{fb} is generally negative, in order to match the reversal of the first echo seen in real spring reverberation units [299]. It need not be constrained to values below 1 as might be expected, as the diffusion introduced by the SFIR diminishes the total loop gain. In this example $g_{fb} = -1.4$.

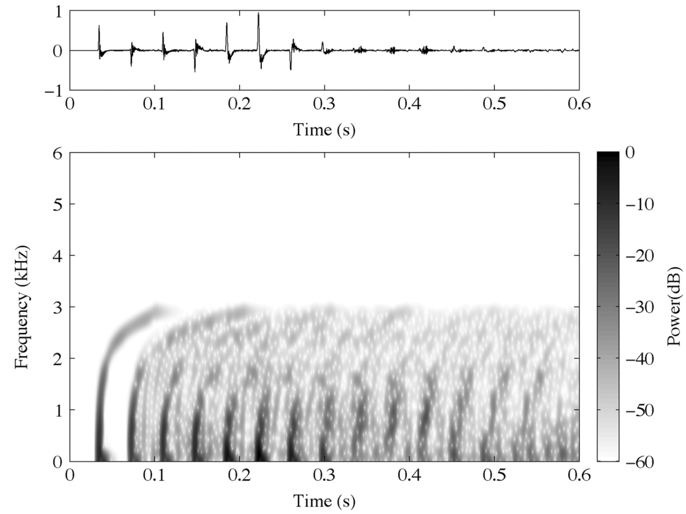


Fig. 19. Impulse response of the spring reverberation model.

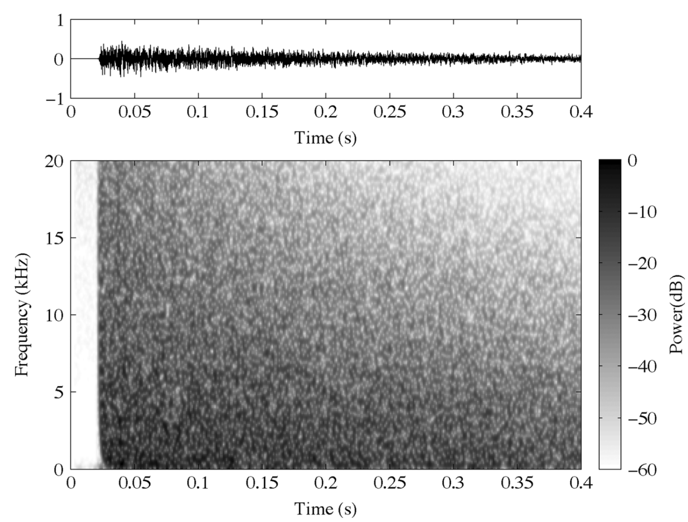


Fig. 20. Impulse response of an EMT 140 reverberation plate, with the damping control set to a medium value.

3) Plate Reverb: An example of an impulse response of a plate reverberator (in this case the EMT 140), is shown in Fig. 20. The response of a plate reverberator is characterized by a number of major qualities. First, it lacks any distinct individual echoes. This is caused by the dispersive nature of wave propagation on the thin metal plate, which spreads individual transients out into longer chirps [18]. Second, it exhibits very fast onset of dense reverberation, described as "whip-like" [64]. Another interesting quality of the onset of the response is that the arrival of the first reverberation components is faster at higher frequencies than at low frequencies. This is due again to dispersive wave propagation on the metal plate, and can be seen in Fig. 20 as a small curve in the onset below around 500 Hz.

Programs claiming to emulate plate reverberation have been present in digital reverberators since the earliest commercial products. The definition of a "plate reverb" in this context has generally been rather flexible, with any reverberation program that exhibits no early reflections and a fast onset of a diffuse reverberant tail being described as a "plate." Recently, work has begun on more rigorously emulating the sound of a plate reverberator.

Bilbao [306] produced a model based on numerical solution of the partial differential equations governing the vibration of the metal plate, via the application of finite difference methods. This approach can accurately replicate the majority of the behavior of the reverberation plate, but has two main drawbacks. First, it can be computationally quite heavy. Second, the mathematical modeling of the real damping mechanisms within the plate is extremely difficult, and hence abstract damping factors are instead used. This results in difficulty modeling the behavior of, for example, the damping control of the EMT 140 reverberation plate.

A different approach is taken by Greenblatt *et al.* [64], who construct an emulation of the EMT 140 plate based on a hybrid reverberation structure consisting of a short convolution section and an FDN. The convolution section is used to model the onset of the plate response, which does not vary as the damping control of the EMT 140 is varied. A part of a measured impulse response from an EMT 140 is used for the onset. The FDN is used to model the late portion of the reverb, and is constructed to match the late reverberation characteristics of the real unit. The reverberation time of the FDN can be varied, allowing some reproduction of the behavior of the damping control of the EMT 140.

B. Electrical Reverberation and Echo

Electrical methods are usually applied to producing the simpler echo effect rather than reverberation, as in contrast to electromechanical techniques there are no natural reflections produced within the system. Reverberation can be emulated using electrical techniques, but it requires the addition of multiple delay taps—in the case of magnetic tapes these consist of extra playback heads, and in the case of bucket-brigade devices (BBDs) these are separate connections to intermediary parts of the capacitor chain which must be specifically designed into the integrated circuit.

1) Tape-Based Echo: The sound of tape-based echo is characterized by several qualities [292]. First, there is the saturation and noise introduced by the medium of the tape itself. This can become very prominent once an echo has recirculated through the system many times. Second, there is a fairly strong (depending on the specific unit) modulation of the speed of the tape playback (and hence the pitch of the output signal), caused by the mechanics of circulating the magnetic tape around the system and past the playback and read heads.

Arnardottir *et al.* [51] modeled the behavior of one specific tape-echo unit, the Echoplex. A structure consisting of an interpolated delay-line with feedback is used. A saturating nonlinearity in the feedback loop is used to emulate the characteristics of the tape. The complexity of the model is in the way in which the length of the delay-line is constantly varied. The pitch fluctuations are separated into several components. The broadly periodic variations introduced by the capstan and the pinch wheel are reproduced using a sum of sinusoidal components, the frequencies and amplitudes of which are matched to measurements from the real unit. The phases of the sinusoidal components are modulated slightly by low-passed zero-mean noise. The broader and slower drift of playback speed, possibly due to the motor driving the system, is modeled using low-pass filtered noise. Finally, the user controllable mechanical "delay handle," which controls the distance between the recording and playback heads, is modeled by low-passing a user controllable parameter. These different sources of pitch fluctuation are summed, and passed through a comb-filter designed to emulate the propagation of mechanical disturbances around the system. This signal is then used to vary the output position from the delay line.

Later work has generalized this model of tape-echo further [292], with the addition of more rigorous modeling of the tape saturation characteristics, and of the tone-control present in the Echoplex as well as other units such as the Roland Space Echo.

2) Bucket-Brigade Devices: Huovilainen [307] and later Raffel and Smith [308] investigate the properties of the sound produced by BBD-based effects. The sound of BBD-based echo and reverberation effects is strongly influenced by the limitations of the BBD itself. First, the BBD is a sampled system, and therefore aliasing is produced based on the clock rate (i.e., sampling frequency) use. Also, the BBD itself is a fairly nonlinear system, and produces distortion and noise, especially when sound is recirculated through it. This distortion is broadly related to how many capacitor stages the signal is passed along, and therefore is more prevalent in chips with many stages. BBD echo circuits often employ BBDs with the maximum available number of stages (4096). To produce an acceptable delay-time for the echo even with this length of chain requires a fairly low clock speed (and hence sampling rate). Since BBD echo circuits also recirculate the signal many times, they generally exhibit the limitations of the BBD device more strongly than other effects. These problems are counteracted somewhat by the addition of a compander surrounding the BBD device, an anti-aliasing filter preceding the BBD, and a reconstruction filter following the BBD. In exchange for reducing the problems with aliasing, noise and distortion, this new system instead has limited bandwidth (caused by the filtering) and also alters the dynamics of signals it processes.

Raffel and Smith [308] then model the BBD echo device on a functional block level. The BBD itself is modeled as a delay-line with a number of samples equivalent to the number of capacitor stages in the BBD being modeled. This delay-line is run at a rate slower than the sampling frequency using an artificial clock signal. This configuration then naturally reproduces the aliasing and variation of pitch when delay time is varied that are present in the real BBD. The delay-line core is then surrounded by modeled versions of the compander, the nonlinearity of the BBD chip and the anti-aliasing and reconstruction filters. A small amount of filtered noise is also injected into the system.

BBD-based reverberation effects, often based on chips such as the Panasonic MN3011 [67], require a shorter overall delay time than echo effects. This allows the use of a high clock rate, and greatly reduces many of the problems with distortion, noise and limited bandwidth exhibited in the echo effect [308].

VII. Conclusion

This overview article has discussed artificial reverberation algorithms, which imitate the acoustic reverberation process. Delay networks tend to be an efficient and flexible approach, but provide control over only gross reverberation features. As such they are well suited for use in game and virtual environment applications, in which the acoustics are interactively adjusted or in musical applications where matching of a particular real acoustic response is unimportant. Their efficiency makes them useful in mobile and home audio playback devices, where memory and computation are at a premium. Schroeder architectures and feedback delay networks appear to be the most widely used, while the novel random-sequence based algorithms appear to be promising competitors.

Computational acoustic methods require significant computational resources, and they are presently mainly used for simplified geometries or to generate reverberation impulse responses. The introduction of fast computing hardware, such as GPUs and other massively parallel architectures, will make computational acoustic simulations run in real time. Convolution-based methods are also computationally demanding, but have the advantage of providing an exact match to a measured space. They are popular in music production and film audio postproduction, where the acoustics of specific locations, such as a famous concert hall or recording studio, are desired.

Virtual analog reverberation algorithms go some way towards allowing the replication of the sound qualities heard in historical recordings. Besides improving the current approaches to modeling electromechanical and electrical reverberation and echo, the next logical step is to extend this virtual analog approach also to vintage digital reverberators, replicating their algorithms and also the effect of their hardware design on the qualities of the sound produced. Famous early digital reverberators such as the Lexicon 224 or the EMT 250 would be excellent targets for this approach.

Artificial reverberation has come a long way since Schroeder's first experiments. With today's computing power, it is already possible to use high quality reverberators in real time on handheld devices. As parallel computing architectures continue to proliferate, even specific time-varying physical spaces can be simulated convincingly, useful for games and virtual presence applications. However, because acoustic spaces are fundamentally very large systems, it will be quite a while before we can afford to ignore psychoacoustic principles that enable large reductions in the computational expense for a given degree of spatial audio realism.

Acknowledgment

The authors would like to thank their colleagues, Dr. A. Franck and Prof. P. Svensson, for helpful discussions and comments. A recent survey article by Dr. T. van Waterschoot and Prof. M. Moonen is acknowledged as a source of inspiration [309].

References

[1] M. R. Schroeder and B. F. Logan, "Colorless artificial reverberation," *J. Audio Eng. Soc.*, vol. 9, no. 3, pp. 192–197, Jul. 1961.
[2] H. Kuttruff, *Room Acoustics*, 3rd ed. London, U.K.: Elsevier Applied Science, 1991.
[3] T. D. Rossing, R. F. Moore, and P. A. Wheeler, *The Science of Sound*, 3rd ed. London, U.K.: Addison-Wesley, 2002.
[4] A. C. Gade, "Acoustics in halls for speech and music," in *Handbook of Acoustics*, T. D. Rossing, Ed. New York: Springer, 2007, pp. 301–350.
[5] L. Beranek, "Analysis of Sabine and Eyring equations and their application to concert hall audience and chair absorption," *J. Acoust. Soc. Amer.*, vol. 120, no. 3, pp. 1399–1410, Sep. 2006.
[6] D. Griesinger, "Concert hall acoustics and audience perception," *IEEE Signal Process. Mag.*, vol. 24, no. 1, pp. 126–131, Mar. 2007.
[7] L. Beranek, *Concert and Opera Halls: How They Sound*. New York: Acoust. Soc. of Amer., 1996.
[8] H. Sakai, S. Sato, and Y. Ando, "Orthogonal acoustical factors of sound fields in a forest compared with those in a concert hall," *J. Acoust. Soc. Amer.*, vol. 104, no. 3, pt. 2, pp. 1491–1497, Sep. 1998.
[9] F. M. Wiener, C. I. Malme, and C. M. Gogos, "Sound propagation in urban areas," *J. Acoust. Soc. Amer.*, vol. 37, no. 4, pp. 738–747, Apr. 1965.
[10] J. Picaut, T. Le Pollés, P. L'Hermite, and V. Gary, "Experimental study of sound propagation in a street," *Appl. Acoust.*, vol. 66, no. 2, pp. 149–173, Feb. 2005.
[11] R. Pieren and J. M. Wunderli, "Sound reflections from cliffs and reverberation in an alpine valley," in *Proc. Forum Acusticum*, Aalborg, Denmark, Jul. 2011.
[12] M. Rettinger, "Reverberation chambers for broadcasting and recording studios," *J. Audio Eng. Soc.*, vol. 5, no. 1, pp. 18–22, Jan. 1957.
[13] B. Blesser, "An interdisciplinary synthesis of reverberation viewpoints," *J. Audio Eng. Soc.*, vol. 49, no. 10, pp. 867–903, Oct. 2001.
[14] L. S. Goodfriend and J. H. Beaumont, "The development and application of synthetic reverberation systems," *J. Audio Eng. Soc.*, vol. 7, no. 4, pp. 228–234, Oct. 1959.
[15] H. Korkes, "Reverberation facilities at CBS radio," in *Proc. 11th Audio Eng. Soc. Conv.*, New York, Oct. 1959, preprint no. 110.
[16] L. Hammond, "Electrical musical instrument," U.S. Patent 2,230,836, Feb. 1941.
[17] W. Kuhl, "Acoustic reverberation arrangements," U.S . Patent 2,923,369, Feb. 1960.
[18] K. Arcas and A. Chaigne, "On the quality of plate reverberation," *Appl. Acoust.*, vol. 71, no. 2, pp. 147–156, Feb. 2010.
[19] B. Owsinski and M. O'Brien, *The Mixing Engineer's Handbook*. Boston, MA: Thomson Course Technology, 2006.
[20] J. Moorer, "About this reverberation business," *Comput. Music J.*, vol. 3, no. 2, pp. 13–28, 1979.
[21] W. G. Gardner, "Reverberation algorithms," in *Applications of Digital Signal Processing to Audio and Acoustics*, M. Kahrs and K. Brandenburg, Eds. Boston, MA: Kluwer, 1997, pp. 85–131.
[22] U. Zölzer, *Digital Audio Signal Processing*, 2nd ed. Chichester, U.K.: Wiley, 2008.
[23] V. Pulkki, T. Lokki, and D. Rocchesso, "Spatial effects," in *DAFX: Digital Audio Effects*, U. Zölzer, Ed., 2nd ed. Chichester, U.K.: Wiley, 2011, pp. 139–183.
[24] M. Kleiner, B.-I. Dalenbäck, and P. Svensson, "Auralization—An overview," *J. Audio Eng. Soc.*, vol. 41, no. 11, pp. 861–875, Nov. 1993.
[25] L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen, "Creating interactive virtual acoustic environments," *J. Audio Eng. Soc.*, vol. 47, no. 9, pp. 675–705, Sep. 1999.
[26] M. Vorländer, *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms, and Acoustic Virtual Reality*. Berlin, Germany: Springer-Verlag, 2007.
[27] M. Vorländer, "Auralization of spaces," *Phys. Today*, vol. 62, no. 6, pp. 35–40, Jun. 2009.
[28] R. Väänänen and J. Huopaniemi, "SNHC audio and audio decomposition," in *The MPEG-4 Book*, F. Pereira and T. Ebrahimi, Eds. Upper Saddle River, NJ: Prentice-Hall, 2002, pp. 545–581.
[29] R. Väänänen and J. Huopaniemi, "Advanced AudioBIFS: Virtual acoustics modeling in MPEG-4 scene description," *IEEE Trans. Multimedia*, vol. 6, no. 5, pp. 661–675, Oct. 2004.

[30] R. Vermeulen, "Stereo-reverberation," *J. Audio Eng. Soc.*, vol. 6, no. 2, pp. 124–130, Apr. 1958.

[31] D. Griesinger, "Improving room acoustics through time-variant synthetic reverberation," in *Proc. Audio Eng. Soc. 90th Conv.*, Paris, France, Feb. 1991, paper no. 5679.

[32] W. G. Gardner, "A realtime multichannel room simulator," in *Proc. 124th Meeting Acoust. Soc. Amer.*, New Orleans, LA, Nov. 1992.

[33] J. Pätynen, "Virtual acoustics in practice rooms," M.S. thesis, Helsinki Univ. of Tech., Espoo, Finland, Oct. 2007.

[34] T. Lokki, J. Pätynen, T. Peltonen, and O. Salmensaari, "A rehearsal hall with virtual acoustics for symphony orchestras," in *Proc. Audio Eng. Soc. 126th Conv.*, Munich, Germany, May 2009, paper no. 7695.

[35] N. Sakamoto, T. Gotoh, and Y. Kimura, "On out-of-head localization in headphone listening," *J. Audio Eng. Soc.*, vol. 24, no. 9, pp. 710–716, Nov. 1976.

[36] D. R. Begault, E. M. Wenzel, and M. R. Anderson, "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source," *J. Audio Eng. Soc.*, vol. 49, no. 10, pp. 904–916, Oct. 2001.

[37] T. Liitola, "Headphone sound externalization," M.S. thesis, Helsinki Univ. of Tech., Espoo, Finland, Mar. 2006.

[38] V. Algazi and R. Duda, "Headphone-based spatial sound," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 33–42, Jan. 2011.

[39] M. R. Bai and C.-C. Lee, "Comparative study of design and implementation strategies of automotive virtual surround audio systems," *J. Audio Eng. Soc.*, vol. 58, no. 3, pp. 141–159, Mar. 2010.

[40] R. Petrick, K. Lohde, M. Wolff, and R. Hoffmann, "The harming part of room acoustics in automatic speech recognition," in *Proc. Interspeech*, Antwerp, Belgium, Aug. 2007, pp. 1094–1097.

[41] A. A. de Lima, F. P. Freeland, P. A. A. Esquef, L. W. P. Biscainho, B. C. Bispo, R. A. de Jesus, S. L. Netto, R. Schafer, A. Said, B. Lee, and A. Kalker, "Reverberation assessment in audioband speech signals for telepresence systems," in *Proc. Int. Conf. Signal Process. Multimedia Applicat.*, Porto, Portugal, Jul. 2008, pp. 257–262.

[42] R. Gomez and T. Kawahara, "Robust speech recognition based on dereverberation parameter optimization using acoustic model likelihood," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 7, pp. 1708–1716, Sep. 2010.

[43] S. H. Mallidi, S. Ganapathy, and H. Hermansky, "Modulation spectrum analysis for recognition of reverberant speech," in *Proc. Interspeech*, Florence, Italy, Aug. 2011, pp. 189–192.

[44] M. Karjalainen and J. Smith, "Body modeling techniques for string instrument synthesis," in *Proc. Int. Comput. Music Conf.*, Hong Kong, Aug. 1996, pp. 232–239.

[45] H. Penttinen, M. Karjalainen, T. Paatero, and H. Järveläinen, "New techniques to model reverberant instrument body responses," in *Proc. Int. Comput. Music Conf.*, Havana, Cuba, Sep. 2001, pp. 182–185.

[46] V. Välimäki, H. Penttinen, J. Knif, M. Laurson, and C. Erkut, "Sound synthesis of the harpsichord using a computationally efficient physical model," *EURASIP J. Appl. Signal Process.*, vol. 2004, no. 7, pp. 934–948, 2004.

[47] B. Bank, S. Zambon, and F. Fontana, "A modal-based real-time piano synthesizer," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 4, pp. 809–821, May 2010.

[48] H.-M. Lehtonen, H. Penttinen, J. Rauhala, and V. Välimäki, "Analysis and modeling of piano sustain-pedal effects," *J. Acoust. Soc. Amer.*, vol. 122, no. 3, pp. 1787–1797, Sep. 2007.

[49] S. Bilbao, K. Arcas, and A. Chaigne, "A physical model for plate reverberation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Toulouse, France, May 2006, vol. 5, pp. 165–165.

[50] J. Abel, D. Berners, S. Costello, and J. Smith, III, "Spring reverb emulation using dispersive allpass filters in a waveguide structure," in *Proc. AES 121st Int. Conv.*, San Francisco, CA, Oct. 2006, paper no. 6854.

[51] S. Arnardottir, J. Abel, and J. Smith, "A digital model of the Echoplex tape delay," in *Proc. 125th Audio Eng. Soc. Conv.*, San Francisco, CA, May 2008, paper no. 7649.

[52] J. Parker and S. Bilbao, "Spring reverberation: A physical perspective," in *Proc. 12th Int. Conf. Digital Audio Effects*, Como, Italy, Sep. 2009, pp. 416–421.

[53] V. Välimäki, F. Fontana, J. O. Smith, and U. Zölzer, "Introduction to the special issue on virtual analog audio effects and musical instruments," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 4, pp. 713–714, May 2010.

[54] H. J. Round, "Transmission and reproduction of sound," U.S. Patent 1,853,286, Apr. 1932.

[55] B. Swedien and Q. Jones, *Make Mine Music.*. Winona, MN: Hal Leonard Corp., 2009.

[56] B. Blesser and L. Salter, *Spaces Speak, Are You Listening?*. Cambridge, MA: MIT Press, 2006.

[57] H. Bode, "History of electronic sound modification," *J. Audio Eng. Soc.*, vol. 32, no. 10, pp. 730–739, Oct. 1984.

[58] H. Meinema, "Artificial reverberation apparatus," U.S. Patent 2,982,819, May 1961.

[59] W. Fidi and B. Weingartner, "Helical spring for producing artificial reverberation," U.S. Patent 3,662,292, May 1972.

[60] B. Weingartner and W. Fidi, "Arrangement for damping torsional oscillations induced in a helical spring, particularly in artificial reverberation devices," U.S. Patent 3,742,140, Jun. 1973.

[61] H. E. Meinema, "Folded-line reverberation device," U.S. Patent 3,363,202, Jan. 1968.

[62] Service Manual for the AKG BX20E.

[63] W. Kuhl, "The acoustical and technological properties of the reverberation plate," *EBU Review, Part A—Technical*, vol. 49, pp. 8–14, 1958.

[64] A. Greenblatt, J. Abel, and D. Berners, "An emulation of the EMT 140 plate reverberator using a hybrid reverberator structure," in *Proc. 127th Audio Eng. Soc. Conv.*, New York, Oct. 2009, paper no. 7928.

[65] K. Kuhl and J. Wieking, "Reverberation device," U.S. Patent 3,719,905, 1973.

[66] F. Sangster and K. Teer, "Bucket-brigade electronics: New possibilities for delay, time-axis conversion, and scanning," *IEEE J. Solid-State Circuits*, vol. 4, no. 3, pp. 131–136, Jun. 1969.

[67] Datasheet for MN3011 IC, Panasonic Semiconductor.

[68] M. R. Schroeder, "Natural-sounding artificial reverberation," *J. Audio Eng. Soc.*, vol. 10, no. 3, pp. 219–223, Jul. 1962.

[69] M. R. Schroeder, "Digital simulation of sound transmission in reverberant spaces," *J. Acoust. Soc. Amer.*, vol. 47, no. 2, pt. 1, pp. 424–431, Feb. 1970.

[70] *Lexicon Celebrates 35 Years of Digital Audio.* Jan. 2007, Clyne Media, Inc., press release.

[71] B. Blesser and F. F. Lee, "An audio delay system using digital technology," *J. Audio Eng. Soc.*, vol. 19, no. 5, pp. 393–397, May 1971.

[72] "EMT 250 electronic reverberation unit instruction manual, after serial no. 162," Nov. 1978, User Manual, Electronik Mess- Und Tonstudiotechnik.

[73] H. Robjohns, "Ursa Major Space Station," *Sound on Sound*, vol. 21, May 1985.

[74] Lexicon, Inc., Little Rock, AR, "Lexicon model 224 digital reverb," Jul. 1980, Service Manual.

[75] *Advanced Music Systems RMX16 Digital Reverberation System User Manual*. Burnley, U.K.: AMS Industries, plc, 1981, User Manual.

[76] B. A. Blesser and K.-O. Bader, "Electric reverberation apparatus," U.S. Patent 4,181,820, Jan. 1980.

[77] L. Dahl and J.-M. Jot, "A reverberator based on absorbent all-pass filters," in *Proc. Int. Conf. Digital Audio Effects*, Verona, Italy, Dec. 2000, pp. 67–72.

[78] J. Allen and D. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, no. 4, pp. 943–950, Apr. 1979.

[79] J. Borish, "Extension of the image model to arbitrary polyhedra," *J. Acoust. Soc. Amer.*, vol. 75, no. 6, pp. 1827–1836, Jan. 1984.

[80] G. Kendall and W. Martens, "Spatial reverberator," U.S. Patent 4,731,848, Mar. 1988.

[81] J. O. Smith, "A new approach to digital reverberation using closed waveguide networks," in *Proc. Int. Comput. Music Conf.*, Vancouver, BC, Canada, Aug. 1985, pp. 47–53.

[82] J. O. Smith, "Digital signal processing using waveguide networks," U.S. Patent 4,984,276, Jan. 1991.

[83] M. A. Gerzon, "Synthetic stereo reverberation, parts I and II," *Studio Sound*, vol. 13(I), 14(II), pp. 632–635(I), 24–28(II), 1971(I), 1972(II).

[84] B. Vercoe and M. Puckette, "Synthetic rehearsal: Training the synthetic performer," in *Proc. Int. Comput. Music Conf.*, Vancouver, BC, Canada, Aug. 1985, pp. 275–278.

[85] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *Proc. 90th Conv. Audio Eng. Soc.*, Paris, France, Jan. 1991, preprint no. 3030.

[86] J.-M. Jot and A. Chaigne, "Method and system for artificial spatialisation of digital audio signals," U.S. Patent 5,491,754, Feb. 1996.

[87] J. Dattorro, "Effects design, part 1: Reverberator and other filters," *J. Audio Eng. Soc.*, vol. 45, no. 9, pp. 660–683, Sep. 1997.

[88] A. Reilly and D. McGrath, "Convolution processing for realistic reverberation," in *Proc. 98th Audio Eng. Soc. Conv.*, Paris, France, Feb. 1995, preprint no. 3977.

[89] H. Robjohns, "Sony DRE S777 sampling digital reverb," *Sound on Sound*, vol. 15, Dec. 1999.

[90] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Proc. AES 108th Conv.*, Paris, France, Feb. 2000, paper no. 5093.
[91] L. Savioja, "Real-time 3D finite-difference time-domain simulation of low- and mid-frequency room acoustics," in *Proc. Int. Conf. Digital Audio Effects*, Graz, Austria, Sep. 2010, pp. 77–84.
[92] L. Savioja, V. Välimäki, and J. O. Smith, "Audio signal processing using graphics processing units," *J. Audio Eng. Soc.*, vol. 59, no. 1/2, pp. 3–19, Jan./Feb. 2011.
[93] N. Tsingos, W. Jian, and I. Williams, "Using programmable graphics hardware for acoustics and audio rendering," *J. Audio Eng. Soc.*, vol. 59, no. 9, pp. 628–646, Sep. 2011.
[94] J. O. Smith, *Physical Audio Signal Processing*, Dec. 2010 [Online]. Available: https://ccrma.stanford.edu/~jos/pasp/, online book
[95] J. O. Smith, *Introduction to Digital Filters*, Sep. 2007 [Online]. Available: https://ccrma.stanford.edu/~jos/filters/, online book
[96] GRAME, "Faust," [Online]. Available: http://faust.grame.fr/
[97] J. O. Smith, Audio Signal Processing in Faust, [Online]. Available: https://ccrma.stanford.edu/~jos/aspf/ online tutorial
[98] J. O. Smith, "Physical modeling using digital waveguides," *Comput. Music J.*, vol. 16, no. 4, pp. 74–87, Winter 1992.
[99] D. Rocchesso and J. Smith, "Generalized digital waveguide networks," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 3, pp. 242–254, May 2003.
[100] J. D. Markel and A. H. Gray, *Linear Prediction of Speech*. New York: Springer Verlag, 1976.
[101] M. Karjalainen, P. Huang, and J. O. Smith, "Digital waveguide networks for room response modeling and synthesis," in *Proc. 118th Audio Eng. Soc. Conv.*, Barcelona, Spain, May 2005, preprint no. 6394.
[102] J. Stautner and M. Puckette, "Designing multi-channel reverberators," *Comput. Music J.*, vol. 6, no. 1, pp. 569–579, 1982.
[103] D. Rocchesso and J. O. Smith, "Circulant and elliptic feedback delay networks for artificial reverberation," *IEEE Trans. Speech Audio Process.*, vol. 5, no. 1, pp. 51–63, Jan. 1997.
[104] D. Rocchesso, "Maximally diffusive yet efficient feedback delay networks for artificial reverberation," *IEEE Signal Process. Lett.*, vol. 4, no. 9, pp. 252–255, Sep. 1997.
[105] F. Menzer and C. Faller, "Unitary matrix design for diffuse Jot reverberators," in *Proc. Audio Eng. Soc. 128th Conv.*, London, U.K., May 2010, paper no. 7984.
[106] E. Piirilä, T. Lokki, and V. Välimäki, "Digital signal processing techniques for non-exponentially decaying reverberation," in *Proc. 1st COST-G6 Workshop Digital Audio Effects*, Barcelona, Spain, Nov. 1998, pp. 21–24.
[107] E. De Sena, H. Hacihabiboğlu, and Z. Cvetković, "Scattering delay network: An interactive reverberator for computer games," in *Proc. AES 41st Int. Conf. Audio for Games*, London, U.K., Feb. 2011.
[108] P. A. Regalia, S. K. Mitra, and P. P. Vaidyanathan, "The digital all-pass filter: A versatile signal processing building block," *Proc. IEEE*, vol. 76, no. 1, pp. 19–37, Jan. 1988.
[109] D. Griesinger, "Practical processors and programs for digital reverberation," in *Proc. AES 7th Int. Conf.*, Toronto, ON, Canada, May 1989, pp. 187–195.
[110] P. Svensson and J. L. Nielsen, "Errors in MLS measurements caused by time variance in acoustic systems," *J. Audio Eng. Soc.*, vol. 47, no. 11, pp. 907–927, Nov. 1999.
[111] J. Frenette, "Reducing artificial reverberation requirements using time-variant feedback delay networks," M.S. thesis, Univ. of Miami, Coral Gables, FL, Dec. 2000.
[112] T. Laakso, V. Välimäki, M. Karjalainen, and U. Laine, "Splitting the unit delay: Tools for fractional delay filter design," *IEEE Signal Process. Mag.*, vol. 13, no. 1, pp. 30–60, Jan. 1996.
[113] T. Lokki and J. Hiipakka, "A time-variant reverberation algorithm for reverberation enhancement systems," in *Proc. Int. Conf. Digital Audio Effects*, Limerick, Ireland, Dec. 2001, pp. 28–32.
[114] T. Choi, Y.-C. Park, and D.-H. Youn, "Design of time-varying reverberators for low memory applications," *IEICE Trans. Inf. Syst.*, vol. E91-D, no. 2, pp. 379–382, Feb. 2008.
[115] R. Väänänen, V. Välimäki, J. Huopaniemi, and M. Karjalainen, "Efficient and parametric reverberator for room acoustics modeling," in *Proc. Int. Comput. Music Conf.*, Thessaloniki, Greece, Sep. 1997, pp. 200–203.
[116] P. Rubak and L. G. Johansen, "Artificial reverberation based on a pseudo-random impulse response, part I," in *Proc. 104th Audio Eng. Soc. Conv.*, Amsterdam, The Netherlands, May 1998, paper no. 4725.
[117] P. Rubak and L. G. Johansen, "Artificial reverberation based on a pseudo-random impulse response, part II," in *Proc. 106th Audio Eng. Soc. Conv.*, Munich, Germany, May 1999, paper no. 4900.
[118] T. May and D. Schobben, "Constant complexity reverberation for any reverberation time," in *Proc. 122nd Audio Eng. Soc. Conv.*, Vienna, Austria, May 2007, paper no. 7040.
[119] M. Karjalainen and H. Järveläinen, "Reverberation modeling using velvet noise," in *Proc. AES 30th Int. Conf. Intell. Audio Environ.*, Saariselkä, Finland, Mar. 2007.
[120] J. Vilkamo, B. Neugebauer, and J. Plogsties, "Sparse frequency-domain reverberator," *J. Audio Eng. Soc.*, vol. 59, no. 12, pp. 936–943, Dec. 2011.
[121] K.-S. Lee, J. S. Abel, V. Välimäki, T. Stilson, and D. P. Berners, "The switched convolution reverberator," *J. Audio Eng. Soc.*, vol. 60, 2012, to be published.
[122] H. Lehnert and J. Blauert, "Principles of binaural room simulation," *Appl. Acoust.*, vol. 36, no. 3–4, pp. 259–291, Jan. 1992.
[123] U. P. Svensson and U. Kristiansen, "Computational modelling and simulation of acoustic spaces," in *Proc. AES 22nd Conf. Virtual, Synth. Entertain. Audio*, Espoo, Finland, Jun. 2002, pp. 11–30.
[124] L. Savioja, T. Rinne, and T. Takala, "Simulation of room acoustics with a 3-D finite difference mesh," in *Proc. Int. Comput. Music Conf.*, Aarhus, Denmark, Sep. 1994, pp. 463–466.
[125] D. Botteldooren, "Finite-difference time-domain simulation of low-frequency room acoustic problems," *J. Acoust. Soc. Amer.*, vol. 98, no. 8, pp. 3302–3308, Jan. 1995.
[126] S. Sakamoto, H. Nagatomo, A. Ushiyama, and H. Tachibana, "Calculation of impulse responses and acoustic parameters in a hall by the finite-difference time-domain method," *Acoust. Sci. Technol.*, vol. 29, no. 4, pp. 256–265, Feb. 2008.
[127] Y. Kagawa, T. Tsuchiya, K. Fujioka, and M. Takeuchi, "Discrete Huygens' model approach to sound wave propagation-reverberation in a room, sound source identification and tomography in time reversal," *J. Sound Vib.*, vol. 225, no. 1, pp. 61–78, Sep. 1999.
[128] T. Schetelig and R. Rabenstein, "Simulation of three-dimensional sound propagation with multidimensional wave digital filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Seattle, WA, May 1998, vol. 6, pp. 3537–3540.
[129] S. Petrausch and R. Rabenstein, "Wave field simulation with the functional transformation method," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Toulouse, France, May 2006, vol. 5, pp. 317–320.
[130] D. Murphy, A. Kelloniemi, J. Mullen, and S. Shelley, "Acoustic modeling using the digital waveguide mesh," *IEEE Signal Process. Mag.*, vol. 24, no. 2, pp. 55–66, Mar. 2007.
[131] S. V. Duyne and J. Smith, III, "The 2-D digital waveguide mesh," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust.*, New Paltz, NY, Oct. 1993, pp. 177–180.
[132] M. Karjalainen, C. Erkut, and L. Savioja, "Compilation of unified physical models for efficient sound synthesis," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Hong Kong, China, Apr. 2003, vol. 5, pp. 433–436.
[133] M. Cobos, J. Escolano, J. Lopez, and B. Pueo, "Subjective effects of dispersion in the simulation of room acoustics using digital waveguide mesh," in *Proc. 124th Conv. Audio Eng. Soc.*, Amsterdam, The Netherlands, May 2008, paper no. 7471.
[134] A. Southern, T. Lokki, and L. Savioja, "The perceptual effects of dispersion error on room acoustic model auralization," in *Proc. Forum Acusticum*, Aalborg, Denmark, Jun. 2011.
[135] F. Fontana and D. Rocchesso, "Signal-theoretic characterization of waveguide mesh geometries for models of two-dimensional wave propagation in elastic media," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 2, pp. 152–161, Feb. 2001.
[136] L. Savioja and V. Välimäki, "Interpolated rectangular 3-D digital waveguide mesh algorithms with frequency warping," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, pp. 783–790, Nov. 2003.
[137] K. Kowalczyk and M. van Walstijn, "On-line simulation of 2D resonators with reduced dispersion error using compact implicit finite difference methods," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Honolulu, HI, Apr. 2007, vol. 1, pp. 285–288.
[138] K. Kowalczyk and M. van Walstijn, "Room acoustics simulation using 3-D compact explicit FDTD schemes," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 1, pp. 34–46, Jan. 2011.
[139] F. Fontana and D. Rocchesso, "Online correction of dispersion error in 2D waveguide meshes," in *Proc. Int. Comput. Music Conf.*, Berlin, Germany, Aug. 2000, pp. 78–81.

[140] J. Escolano and F. Jacobsen, "A note on the physical interpretation of frequency dependent boundary conditions in a digital waveguide mesh," *Acta Acustica United With Acustica*, vol. 93, no. 3, pp. 398–402, May/Jun. 2007.

[141] K. Kowalczyk and M. van Walstijn, "Modelling frequency-dependent boundaries as digital impedance filters in FDTD and K-DWM room acoustics simulations," *J. Audio Eng. Soc.*, vol. 56, no. 7/8, pp. 569–583, Jul. 2008.

[142] K. Kowalczyk, M. van Walstijn, and D. Murphy, "A phase grating approach to modeling surface diffusion in FDTD room acoustics simulations," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 3, pp. 528–537, Mar. 2011.

[143] J. Escolano, J. Navarro, D. Murphy, J. Wells, and J. Lopez, "A comparison of two diffuse boundary models based on finite differences schemes," in *Proc. 128th Audio Eng. Soc. Conv.*, London, U.K., May 2010, paper no. 8022.

[144] D. Murphy, G. Guilherme, and D. Howard, "Surround-sound reverberation using digital waveguide mesh modelling techniques," in *Proc. 19th Audio Eng. Soc. Int. Conf. Surround Sound—Techniques, Technol., and Percept.*, York, U.K., Jun. 2001.

[145] M. Beeson, A. Moore, D. Murphy, and S. Shelley, "RenderAIR—Room acoustics simulation using a hybrid digital waveguide mesh approach," in *Proc. 124th Audio Eng. Soc. Conv.*, Amsterdam, The Netherlands, May 2008, paper no. 7429.

[146] J. Escolano, J. Ló;pez, B. Pueo, and G. Ramos, "On the implementation of a room acoustics modeling software using finite-differences time-domain method," in *Proc. 122nd Conv. Audio Eng. Soc.*, Vienna, Austria, May 2007, paper no. 7090.

[147] J. Lopez, A. Gonzalez, and F. Orduña-Bustamante, "Simulation of complex and large rooms using a digital waveguide mesh," in *Proc. 123rd Conv. Audio Eng. Soc.*, New York, Oct. 2007, paper no. 7189.

[148] N. Röber, M. Spindler, and M. Masuch, "Waveguide-based room acoustics through graphics hardware," in *Proc. Int. Comput. Music Conf.*, New Orleans, LA, Nov. 2006.

[149] A. Southern, D. Murphy, G. Campos, and P. Dias, "Finite difference room acoustic modelling on a general purpose graphics processing unit," in *Proc. 128th Audio Eng. Soc. Conv.*, London, U.K., Mar. 2010, preprint no. 8028.

[150] C. Webb and S. Bilbao, "Virtual room acoustics: A comparison of techniques for computing 3D-FDTD schemes using CUDA," in *Proc. 130th Audio Eng. Soc. Conv.*, London, U.K., May 2011, preprint no. 8438.

[151] C. Webb and S. Bilbao, "Computing room acoustics with CUDA—3D FDTD schemes with boundary losses and viscosity," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Prague, Czech Republic, May 2011, pp. 317–320.

[152] R. Mehra, N. Raghuvanshi, L. Savioja, M. C. Lin, and D. Manocha, "An efficient GPU-based time domain solver for the acoustic wave equation," *Appl. Acoust.*, vol. 73, no. 2, pp. 83–94, Feb. 2012.

[153] G. Campos and S. Barros, "The Meshotron: A network of specialized hardware units for 3-D digital waveguide mesh acoustic model parallelization," in *Proc. 128th Audio Eng. Soc. Conv.*, London, U.K., May 2010, paper no. 8054.

[154] A. Kelloniemi, V. Välimäki, and L. Savioja, "Simulation of room acoustics using 2-D digital waveguide meshes," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Toulouse, France, May 2006, vol. 5, pp. 313–316.

[155] J. Wells, D. Murphy, and M. Beeson, "Temporal matching of 2-D and 3-D wave-based acoustic modeling for efficient and realistic simulation of rooms," in *Proc. 126th Audio Eng. Soc. Conv.*, Munich, Germany, May 2009.

[156] A. Kelloniemi, V. Välimäki, P. Huang, and L. Savioja, "Artificial reverberation using a hyper-dimensional FDTD mesh," in *Proc. Eur. Signal Process. Conf.*, Antalya, Turkey, Sep. 2005.

[157] N. Raghuvanshi, C. Lauterbach, A. Chandak, D. Manocha, and M. Lin, "Real-time sound synthesis and propagation for games," *Commun. ACM*, vol. 50, no. 7, pp. 66–73, Jul. 2007.

[158] N. Raghuvanshi, B. Lloyd, N. Govindaraju, and M. Lin, "Efficient numerical acoustic simulation on graphics processors using adaptive rectangular decomposition," in *Proc. EAA Symp. Auralization*, Espoo, Finland, Jun. 2009.

[159] N. Raghuvanshi, J. Snyder, R. Mehra, M. Lin, and N. Govindaraju, "Precomputed wave simulation for real-time sound propagation of dynamic sources in complex scenes," *ACM Trans. Graphics*, vol. 29, no. 4, pp. 68:1–68:11, Jul. 2010.

[160] K. Spratt and J. S. Abel, "A digital reverberator modeled after the scattering of acoustic waves by trees in a forest," in *Proc. 125th Audio Eng. Soc. Conv.*, San Francisco, CA, May 2008, preprint no. 7650.

[161] S. Siltanen, T. Lokki, S. Kiminki, and L. Savioja, "The room acoustic rendering equation," *J. Acoust. Soc. Amer.*, vol. 122, no. 3, pp. 1624–1635, Sep. 2007.

[162] A. Krokstad, S. Strøm, and S. Sørsdal, "Calculating the acoustical room response by the use of a ray tracing technique," *J. Sound Vibr.*, vol. 8, no. 1, pp. 118–125, Jan. 1968.

[163] A. Krokstad, S. Strøm, and S. Sørsdal, "Fifteen years' experience with computerized ray tracing," *Appl. Acoust.*, vol. 16, no. 4, pp. 291–312, Apr. 1983.

[164] A. Kulowski, "Algorithmic representation of the ray tracing technique," *Appl. Acoust.*, vol. 18, no. 6, pp. 449–469, 1985.

[165] H. Lehnert, "Systematic errors of the ray-tracing algorithm," *Appl. Acoust.*, vol. 38, no. 2–4, pp. 207–221, 1993.

[166] T. Funkhouser, I. Carlbom, G. Elko, G. Pingali, G. Pingali, M. Sondhi, and J. West, "A beam tracing approach to acoustic modeling for interactive virtual environments," in *Proc. Int. Conf. Comput. Graphics Interactive Tech. SIGGRAPH'98*, Jan. 1998, pp. 21–32.

[167] S. Laine, S. Siltanen, T. Lokki, and L. Savioja, "Accelerated beam tracing algorithm," *Appl. Acoust.*, vol. 70, no. 1, pp. 172–181, Jan. 2008.

[168] F. Antonacci, M. Foco, A. Sarti, and S. Tubaro, "Fast tracing of acoustic beams and paths through visibility lookup," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 4, pp. 812–824, Jan. 2008.

[169] F. Antonacci, A. Sarti, and S. Tubaro, "Two-dimensional beam tracing from visibility diagrams for real-time acoustic rendering," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. Article ID 642316, 2010.

[170] A. Chandak, C. Lauterbach, M. Taylor, Z. Ren, and D. Manocha, "AD-frustum: Adaptive frustum tracing for interactive sound propagation," *IEEE Trans. Visualiz. Comput. Graphics*, vol. 14, no. 6, pp. 1707–1722, Nov./Dec. 2008.

[171] U. Kristiansen, A. Krokstad, and T. Follestad, "Extending the image method to higher-order reflections," *Appl. Acoust.*, vol. 38, no. 2–4, pp. 195–206, Mar. 1993.

[172] E. Lehmann and A. Johansson, "Diffuse reverberation model for efficient image-source simulation of room impulse responses," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 6, pp. 1429–1439, Aug. 2010.

[173] S. G. McGovern, "Fast image method for impulse response calculations of box-shaped rooms," *Appl. Acoust.*, vol. 70, no. 1, pp. 182–189, Jan. 2009.

[174] S. McGovern, "The image-source reverberation model in an N-dimensional space," in *Proc. 14th Int. Conf. Digital Audio Effects*, Paris, France, Sep. 2011, pp. 11–18.

[175] N. Tsingos, E. Gallo, and G. Drettakis, "Perceptual audio rendering of complex virtual environments," *ACM Trans. Graphics*, vol. 23, Aug. 2004.

[176] K. Kuttruff, "Auralization of impulse responses modeled on the basis of ray-tracing results," *J. Audio Eng. Soc.*, vol. 41, no. 11, pp. 876–880, Nov. 1993.

[177] D. van Maercke and J. Martin, "The prediction of echograms and impulse responses within the Epidaure software," *Appl. Acoust.*, vol. 38, no. 2–4, pp. 93–114, Jan. 1993.

[178] T. Lewers, "A combined beam tracing and radiant exchange computer model of room acoustics," *Appl. Acoust.*, vol. 38, no. 2–4, pp. 161–178, Mar. 1993.

[179] E.-M. Nosal, M. Hodgson, and I. Ashdown, "Improved algorithms and methods for room sound-field prediction by acoustical radiosity in arbitrary polyhedral rooms," *J. Acoust. Soc. Amer.*, vol. 116, no. 2, pp. 970–980, Aug. 2004.

[180] B. Kapralos, M. Jenkin, and E. Milios, "Acoustical modeling with sonel mapping," in *Proc. 19th Int. Congr. Acoust.*, Madrid, Spain, Sep. 2007.

[181] M. Bertram, E. Deines, J. Mohring, J. Jegorovs, and H. Hagen, "Phonon tracing for auralization and visualization of sound," in *Proc. IEEE Visualiz.*, Minneapolis, MN, Oct. 2005, pp. 151–158.

[182] H. W. Jensen and N. J. Christensen, "Photon maps in bidirectional Monte Carlo ray tracing of complex objects," *Comput. Graphics*, vol. 19, no. 2, pp. 215–224, Mar. 1995.

[183] U. Stephenson, "Quantized pyramidal beam tracing—A new algorithm for room acoustics and noise immission prognosis," *Acustica United With Acta Acustica*, vol. 82, no. 3, pp. 517–525, May 1996.

[184] J. Martin, D. V. Maercke, and J.-P. Vian, "Binaural simulation of concert halls: A new approach for the binaural reverberation process," *J. Acoust. Soc. Amer.*, vol. 94, no. 6, pp. 3255–3264, Dec. 1993.

[185] S. Siltanen, T. Lokki, and L. Savioja, "Frequency domain acoustic radiance transfer for real-time auralization," *Acta Acustica United With Acustica*, vol. 95, no. 1, pp. 106–117, Jan. 2009.

[186] E. Stavrakis, N. Tsingos, and P. Calamia, "Topological sound propagation with reverberation graphs," *Acta Acustica United With Acustica*, vol. 94, no. 6, pp. 921–932, Nov. 2008.

[187] C. Lauterbach, A. Chandak, and D. Manocha, "Interactive sound rendering in complex and dynamic scenes using frustum tracing," *IEEE Trans. Visualiz. Comput. Graphics*, vol. 13, no. 6, pp. 1672–1679, Nov. 2007.

[188] M. Taylor, A. Chandak, Z. Ren, C. Lauterbach, and D. Manocha, "Fast edge-diffraction for sound propagation in complex virtual environments," in *Proc. EAA Auralization Symp.*, Espoo, Finland, Jun. 2009.

[189] C. Schissler and D. Manocha, "GSound: Interactive sound propagation for games," in *Proc. AES 41th Int. Conf.*, London, U.K., 2011.

[190] R. Torres, U. Svensson, and M. Kleiner, "Computation of edge diffraction for more accurate room acoustics auralization," *J. Acoust. Soc. Amer.*, vol. 109, no. 2, pp. 600–610, Feb. 2001.

[191] T. Lokki, U. Svensson, and L. Savioja, "An efficient auralization of edge diffraction," in *Proc. AES 21st Int. Conf. Architect. Acoust. Sound Reinforcement*, St. Petersburg, Russia, Jun. 2002, pp. 166–172.

[192] P. Calamia and U. Svensson, "Culling insignificant diffraction components for interactive acoustic simulations," in *Proc. 19th Int. Congr. Acoust.*, Madrid, Spain, Sep. 2007.

[193] N. Tsingos, T. Funkhouser, A. Ngan, and I. Carlbom, "Modeling acoustics in virtual environments using the uniform theory of diffraction," in *Proc. 28th Conf. Comput. Graphics Interact. Tech. (SIGGRAPH'01)*, Los Angeles, CA, Aug. 2001.

[194] N. Tsingos, C. Dachsbacher, and S. Lefebvre, "Instant sound scattering," in *Proc. 18th Eurographics Symp. Rendering*, Jun. 2007.

[195] M. Vorländer, "Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm," *J. Acoust. Soc. Amer.*, vol. 86, no. 1, pp. 172–178, Jul. 1989.

[196] R. Heinz, "Binaural room simulation based on an image source model with addition of statistical-methods to include the diffuse sound scattering of walls and to predict the reverberant tail," *Appl. Acoust.*, vol. 38, no. 2–4, pp. 145–159, Jan. 1993.

[197] D. Murphy, M. Beeson, S. Shelley, and A. Moore, "Hybrid room impulse response synthesis in digital waveguide mesh based room acoustics simulation," in *Proc. Int. Conf. Digital Audio Effects*, Espoo, Finland, Sep. 2008, pp. 129–136.

[198] M. Aretz, R. Nöthen, M. Vorländer, and D. Schröder, "Combined broadband impulse responses using FEM and hybrid ray-based methods," in *Proc. EAA Auralization Symp.*, Espoo, Finland, Jun. 2009.

[199] A. Southern, S. Siltanen, and L. Savioja, "Spatial room impulse responses with a hybrid modeling method," in *Proc. 130th Audio Eng. Soc. Conv.*, London, U.K., May 2011, preprint no. 8385.

[200] G. S. Kendall and W. L. Martens, "Simulating the cues of spatial hearing in natural environments," in *Proc. Int. Comput. Music Conf.*, Paris, France, Oct. 1984, pp. 111–125.

[201] S. Foster, E. Wenzel, and R. Taylor, "Real-time synthesis of complex acoustic environments," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust.*, New Paltz, NY, Oct. 1991, pp. 47–48.

[202] J.-M. Jot, V. Larcher, and O. Warusfel, "Digital signal processing issues in the context of binaural and transaural stereophony," in *Proc. 98th Audio Eng. Soc. Conv.*, Paris, France, Feb. 1995, preprint no. 3980.

[203] J. Abel and S. Foster, "Method and apparatus for efficient presentation of high-quality three-dimensional audio," U.S. Patent 5,596,644, Jan. 1997.

[204] W. Ahnert and R. Feistel, "EARS auralization software," *J. Audio Eng. Soc.*, vol. 41, no. 11, pp. 894–904, Nov. 1993.

[205] G. Naylor, "Computer modeling and auralisation of sound fields in rooms," *Appl. Acoust.*, vol. 38, no. 2–4, pp. 89–92, Jan. 1993.

[206] B.-I. Dalenbäck, "A new model for room acoustic prediction and auralization," Ph.D. dissertation, Chalmers Univ. of Tech., Gothenburg, Sweden, 1995.

[207] J. Vian and J. Martin, "Binaural room acoustics simulation: Practical uses and applications," *Appl. Acoust.*, vol. 36, no. 3–4, pp. 293–305, Mar. 1992.

[208] A. Farina, "Auralization software for the evaluation of the results obtained by a pyramid tracing code: Results of subjective listening tests," in *Proc. 15th Int. Congr. Acoust.*, Trondheim, Norway, Jun. 1995.

[209] E. M. Wenzel, J. D. Miller, and J. S. Abel, "Sound lab: A real-time, software-based system for the study of spatial hearing," in *Proc. 108th Audio Eng. Soc. Conv.*, Paris, France, Feb. 2000, preprint no. 5140.

[210] K. Nakagawa, T. Miyajima, and Y. Tahara, "An improved geometrical sound field analysis in rooms using scattered sound and an audible room acoustic simulator," *Appl. Acoust.*, vol. 38, no. 2–4, pp. 115–129, Jan. 1993.

[211] T. Takala, R. Hänninen, V. Välimäki, L. Savioja, J. Huopaniemi, T. Huotilainen, and M. Karjalainen, "An integrated system for virtual audio reality," in *Proc. 100th Audio Eng. Soc. Conv.*, Copenhagen, Denmark, May 1996, preprint no. 4229.

[212] J. Huopaniemi, L. Savioja, and M. Karjalainen, "Modeling of reflections and air absorption in acoustical spaces a digital filter design approach," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust.*, New Paltz, NY, Oct. 1997, pp. 1–4.

[213] M. Noisternig, B. F. G. Katz, S. Siltanen, and L. Savioja, "Framework for real-time auralization in architectural acoustics," *Acta Acustica United With Acustica*, vol. 94, no. 6, pp. 1000–1015, Nov. 2008.

[214] A. Silzle, P. Novo, and H. Strauss, "IKA-SIM: A system to generate auditory virtual environments," in *Proc. 116th Audio Eng. Soc. Conv.*, Jan. 2004, preprint no. 6016.

[215] T. Lentz, D. Schröder, M. Vorländer, and I. Assenmacher, "Virtual reality system with integrated sound field simulation and reproduction," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. Article ID 70540, Jan. 2007.

[216] D. Schröder and A. Pohl, "Real-time hybrid simulation method including edge diffraction," in *Proc. EAA Auralization Symp.*, Espoo, Finland, Jun. 2009, pp. 1–6.

[217] F. Wefers and D. Schröder, "Real-time auralization of coupled rooms," in *Proc. EAA Auralization Symp.*, Espoo, Finland, Jun. 2009, pp. 1–6.

[218] L. Savioja, D. Manocha, and M. Lin, "Use of GPUs in room acoustic modeling and auralization," in *Proc. Int. Symp. Room Acoust.*, Melbourne, Australia, Aug. 2010.

[219] M. Jedrzejewski and K. Marasek, "Computation of room acoustics using programmable video hardware," in *Proc. Int. Conf. Comput. Vis. Graph.*, Warsaw, Poland, Sep. 2004, pp. 587–592.

[220] N. Röber, U. Kaminski, and M. Masuch, "Ray acoustics using computer graphics technology," in *Proc. 10th Int. Conf. Digital Audio Effects*, Bordeaux, France, Sep. 2007, pp. 117–124.

[221] N. Tsingos, "Pre-computing geometry-based reverberation effects for games," in *Proc. AES 35th Int. Conf.*, London, U.K., 2009.

[222] T. Lokki, J. Pätynen, S. Tervo, S. Siltanen, and L. Savioja, "Engaging concert hall acoustics is made up of temporal envelope preserving reflections," *J. Acoust. Soc. Amer.*, vol. 129, no. 5, pp. EL223–EL228, Jun. 2011.

[223] W. G. Gardner, "Efficient convolution without input-output delay," *J. Audio Eng. Soc.*, vol. 43, no. 3, pp. 127–136, Mar. 1995.

[224] B. Cowan and B. Kapralos, "Real-time GPU-based convolution: A follow-up," in *Proc. ACM FuturePlay, GDC Canada, Int. Conf. Future of Game Design and Technology*, Vancouver, BC, Canada, May 2009, pp. 25–26.

[225] D. Griesinger, "Beyond MLS—Occupied hall measurement with FFT techniques," in *Proc. 101st Conv. Audio Eng. Soc.*, Los Angeles, CA, Nov. 1996, preprint no. 4403.

[226] J. Pätynen, B. F. G. Katz, and T. Lokki, "Investigations on the balloon as an impulse source," *J. Acoust. Soc. Amer.*, vol. 129, no. 1, pp. EL27–EL33, Jan. 2011.

[227] L. Peltola, C. Erkut, P. R. Cook, and V. Välimäki, "Synthesis of hand clapping sounds," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 3, pp. 1021–1029, Mar. 2007.

[228] J. S. Abel, N. J. Bryan, P. P. Huang, M. Kolar, and B. V. Pentcheva, "Estimating room impulse responses from recorded balloon pops," in *Proc. 129th Audio Eng. Soc. Conv.*, San Francisco, CA, Nov. 2010, paper no. 8171.

[229] S. Müller and P. Massarani, "Transfer-function measurement with sweeps," *J. Audio Eng. Soc.*, vol. 49, no. 6, pp. 443–471, Jun. 2001.

[230] J. Borish and J. B. Angell, "An efficient algorithm for measuring the impulse response using pseudorandom noise," *J. Audio Eng. Soc.*, vol. 31, no. 7/8, pp. 478–488, Jul./Aug. 1983.

[231] R. C. Rife and J. Vanderkooy, "Transfer-function measurements with maximum-length sequences," *J. Audio Eng. Soc.*, vol. 37, no. 6, pp. 419–444, Jun. 1989.

[232] M. R. Schroeder, "Integrated-impulse method measuring sound decay without using impulses," *J. Acoust. Soc. Amer.*, vol. 66, no. 2, pp. 497–500, Aug. 1979.

[233] G.-B. Stan, J. J. Embrechts, and D. Archambeau, "Comparison of different impulse response measurement techniques," *J. Audio Eng. Soc.*, vol. 50, no. 4, pp. 249–262, Apr. 1983.

[234] J.-M. Jot, L. Cerveau, and O. Warusfel, "Analysis and synthesis of room reverberation based on a statistical time-frequency model," in *Proc. 103rd Conv. Audio Eng. Soc.*, New York, Sep. 1997, paper no. 7150.

[235] N. J. Bryan and J. S. Abel, "Methods for extending room impulse responses beyond their noise floor," in *Proc. Audio Eng. Soc. 129th Conv.*, San Francisco, CA, Nov. 2010, paper no. 8167.

[236] R. Ben-Hador and I. Neoran, "Capturing manipulation and reproduction of sampled acoustic impulse responses," in *Proc. Audio Eng. Soc. 117th Conv.*, San Francisco, CA, Oct. 2004, paper no. 6232.

[237] J. van Dorp Schuitman and D. de Vries, "Applying cochlear modeling and psychoacoustics in room acoustics," in *Proc. 124th Conv. Audio Eng. Soc.*, Amsterdam, The Netherlands, May 2008, preprint no. 7469.

[238] M. Heideman, D. Johnson, and C. S. Burrus, "Gauss and the history of the FFT," *IEEE Signal Process. Mag.*, vol. 1, no. 3, pp. 14–21, Oct. 1984.

[239] T. G. Stockham, Jr., "High speed convolution and correlation," in *Proc. Spring Joint Comput. Conf.*, Boston, MA, Apr. 1966, pp. 229–233.

[240] A. J. Reijnen, J.-J. Sonke, and D. de Vries, "New developments in electro-acoustics reverberation technology," in *Proc. 98th Audio Eng. Soc. Conv.*, Paris, France, Feb. 1995, preprint no. 3978.

[241] B. D. Kulp, "Digital equalization using Fourier transform techniques," in *Proc. 85th Audio Eng. Soc. Conv.*, Los Angeles, CA, Nov. 1988, preprint no. 2694.

[242] F. Wefers and M. Vorländer, "Optimal filter partitions for real-time FIR filtering using uniformly-partitioned FFT-based convolution in the frequency-domain," in *Proc. 14th Int. Conf. Digital Audio Effects*, Paris, France, Sep. 2011, pp. 155–161.

[243] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 3rd ed. Upper Saddle River, NJ: Pearson, 2010.

[244] D. S. McGrath, "Method and apparatus for filtering an electronic environment with improved accuracy and efficiency and short flow-through delay," U.S. Patent 5,502,747, Mar. 1996.

[245] G. Garcia, "Optimal filter partition for efficient convolution with short input/output delay," in *Proc. Audio Eng. Soc. 113th Conv.*, New York, Oct. 2002, preprint no. 5660.

[246] G. Garcia, "System and method for signal processing using an improved convolution technique," U.S. Patent 6,625,629, Sep. 2003.

[247] A. Torger and A. Farina, "Real-time partitioned convolution for Ambiophonics surround sound," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, New Paltz, NY, Oct. 2001, pp. 195–198.

[248] J. Hurchalla, "Low latency convolution in one dimension via two dimensional convolutions: An intuitive approach," in *Proc. 125th AES Conv.*, San Francisco, CA, Oct. 2008, paper no. 7634.

[249] J. G. Kuk, S. Kim, and N. I. Cho, "A new overlap save algorithm for fast block convolution and its implementation using FFT," *J. Signal Process. Syst.*, vol. 63, no. 1, pp. 143–152, Mar. 2011.

[250] A. P. Carrillo, J. Bonada, J. Pätynen, and V. Välimäki, "Method for measuring violin sound radiation based on bowed glissandi and its application to sound synthesis," *J. Acoust. Soc. Amer.*, vol. 130, no. 2, pp. 1020–1029, Aug. 2011.

[251] E. Battenberg and R. Avižienis, "Implementing real-time partitioned convolution algorithms on conventional operating systems," in *Proc. 14th Int. Conf. Digital Audio Effects*, Paris, France, Sep. 2011, pp. 313–320.

[252] N. Govindaraju, B. Lloyd, Y. Dotsenko, B. Smith, and J. Manferdelli, "High performance discrete Fourier transforms on graphics processors," in *Proc. ACM/IEEE Conf. Supercomputing*, Nov. 2008.

[253] T. Ilmonen and T. Lokki, "Extreme filters—Cache-efficient implementation of long IIR and FIR filters," *IEEE Signal Process. Lett.*, vol. 13, no. 7, pp. 401–404, Jul. 2006.

[254] F. Spandöck, "Akustische modellversuche," *Ann. Physik V*, vol. 412, no. 4, pp. 345–360, 1934.

[255] N. F. Spring and C. L. S. Gilford, "Artificial reverberation," BBC Research Department Rep. no. 1972/19, May 1972.

[256] N. Xiang and J. Blauert, "Binaural scale modeling for auralisation and prediction of acoustics in auditoriums," *Appl. Acoust.*, vol. 38, no. 2–4, pp. 267–290, Jan. 1993.

[257] K. Spratt and J. S. Abel, "All natural room enhancement," in *Proc. Int. Computer Music Conf.*, Montreal, QC, Canada, Aug. 2009, pp. 231–234.

[258] J. Laroche, "A new analysis/synthesis system of musical signals using Prony's method—Application to heavily damped percussive sounds," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Glasgow, U.K., May 1989, vol. 3, pp. 2053–2056.

[259] K. Steiglitz and L. F. McBride, "A technique for the identification of linear system," *IEEE Trans. Autom. Control*, vol. AC-10, no. 4, pp. 461–464, Oct. 1965.

[260] B. Beliczynski, I. Kale, and G. Cain, "Approximation of FIR by IIR digital filters: An algorithm based on balanced model reduction," *IEEE Trans. Signal Process.*, vol. 40, no. 3, pp. 532–542, Mar. 1992.

[261] I. Kale, J. Mackenzie, and T. Laakso, "Motor car acoustic response modelling and order reduction via balanced model truncation," *Electron. Lett.*, vol. 32, no. 11, pp. 965–966, May 1996.

[262] M. Karjalainen, P. A. A. Esquef, P. Antsalo, A. Mäkivirta, and V. Välimäki, "Frequency-zooming ARMA modeling of resonant and reverberant systems," *J. Audio Eng. Soc.*, vol. 50, no. 12, pp. 1012–1029, Dec. 2002.

[263] U. Zölzer, N. Fliege, M. Schonle, and M. Schusdziara, "Multirate digital reverberation system," in *Proc. 89th Audio Eng. Soc. Conv.*, Los Angeles, CA, Sep. 1990, preprint no. 2968.

[264] M. Schoenle, N. Fliege, and U. Zölzer, "Parametric approximation of room impulse responses by multirate systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Minneapolis, MN, Apr. 1993, vol. 1, pp. 153–156.

[265] F. Jabloun and B. Champagne, "A fast subband room response simulator," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Istanbul, Turkey, Jun. 2000, vol. 2, pp. 925–928.

[266] J. M. Peterson and C. Kyriakakis, "Using subband filters to reduce the complexity of real-time signal processing," in *Proc. Audio Eng. Soc. 113th Conv.*, Los Angeles, CA, Oct. 2002, paper no. 5663.

[267] D. Marelli and M. Fu, "A recursive method for the approximation of LTI systems using subband processing," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1025–1034, Mar. 2010.

[268] F. Menzer and C. Faller, "Investigations on an early-reflection-free model for BRIRs," *J. Audio Eng. Soc.*, vol. 58, no. 9, pp. 709–723, Sep. 2010.

[269] M. R. Bai, K.-Y. Ou, and P. Zeung, "Multirate synthesis of reverberators using subband filtering," *J. Sound Vibr.*, vol. 321, no. 3–5, pp. 1090–1108, Apr. 2009.

[270] E. Vickers, J.-L. Wu, P. G. Krishnan, and R. N. K. Sadanandam, "Frequency domain artificial reverberation using spectral magnitude decay," in *Proc. 121st Audio Eng. Soc. Conv.*, San Francisco, CA, Oct. 2006, preprint no. 6926.

[271] J.-M. Jot, "An analysis/synthesis approach to real-time artificial reverberation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, San Francisco, CA, Mar. 1992, vol. 2, pp. 221–224.

[272] R. Stewart and D. Murphy, "A hybrid artificial reverberation algorithm," in *Proc. Audio Eng. Soc. 122nd Conv.*, Vienna, Austria, May 2007, paper no. 7021.

[273] A. Greenblatt, J. Abel, and D. Berners, "A hybrid reverberation crossfading technique," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Dallas, TX, Mar. 2010, pp. 429–432.

[274] J. S. Abel and P. Huang, "A simple, robust measure of reverberation echo density," in *Proc. Audio Eng. Soc. 121st Conv.*, San Francisco, CA, Oct. 2006, paper no. 6985.

[275] F. Menzer, "Binaural reverberation using two parallel feedback delay networks," in *Proc. Audio Eng. Soc. 40th Int. Conf.*, Tokyo, Japan, Oct. 2010, pp. 1–10.

[276] M. A. Gerzon, "Recording concert hall acoustics for posterity," *J. Audio Eng. Soc.*, vol. 23, no. 7, pp. 569–571, Sep. 1975.

[277] V. Pulkki, J. Merimaa, and T. Lokki, "Reproduction of reverberation with spatial impulse response rendering," in *Proc. Audio Eng. Soc. 116th Conv.*, Berlin, Germany, May 2004, paper no. 6057.

[278] J. Merimaa and V. Pulkki, "Spatial impulse response rendering I: Analysis and synthesis," *J. Audio Eng. Soc.*, vol. 53, no. 12, pp. 1115–1127, Dec. 2005.

[279] J. Merimaa, "Analysis, synthesis, and perception of spatial sound—binaural localization modeling and multichannel loudspeaker reproduction," Ph.D. dissertation, Helsinki Univ. of Technol., Espoo, Finland, Aug. 2006.

[280] T. Lokki, J. Merimaa, and V. Pulkki, "Method for reproducing natural or modified spatial impression in multichannel listening," U.S. Patent 7,787,638 B2, Aug. 2010.

[281] E. Hulsebos, D. de Vries, and E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by wave-field synthesis," *J. Audio Eng. Soc.*, vol. 50, no. 10, pp. 779–790, Oct. 2002.

[282] Y. Li, P. F. Driessen, G. Tzanetakis, and S. Bellamy, "Spatial sound rendering using measured room impulse responses," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol.*, Vancouver, BC, Canada, Aug. 2006, pp. 432–437.

[283] M. Kuster, "Multichannel room impulse response generation with coherence control," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 4, pp. 597–606, May 2009.

[284] V. Välimäki and A. Huovilainen, "Oscillator and filter algorithms for virtual analog synthesis," *Comput. Music J.*, vol. 30, no. 2, pp. 19–31, 2006.

[285] F. Fontana and M. Civolani, "Modeling of the EMS VCS3 voltage-controlled filter as a nonlinear filter network," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 4, pp. 760–772, May 2010.

[286] T. Hélie, "Volterra series and state transformation for real-time simulations of audio circuits including saturations: Application to the Moog ladder filter," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 4, pp. 747–759, May 2010.
[287] J. Parker, "A simple digital model of the diode-based ring-modulator," in *Proc. 14th Int. Conf. Digital Audio Effects*, Paris, France, Sep. 2011, pp. 163–166.
[288] J. S. Abel and D. P. Berners, "On peak-detecting and RMS feedback and feedforward compressors," in *Proc. Audio Eng. Soc. 115th Conv.*, New York, Oct. 2003, paper no. 5914.
[289] O. Kröning, K. Dempwolf, and U. Zölzer, "Analysis and simulation of an analog guitar compressor," in *Proc. 14th Int. Conf. Digital Audio Effects*, Paris, France, Sep. 2011, pp. 205–208.
[290] J. Pakarinen and D. T. Yeh, "A review of digital techniques for modeling vacuum-tube guitar amplifiers," *Comput. Music J.*, vol. 33, no. 2, pp. 85–100, 2009.
[291] D. Yeh, J. Abel, and J. Smith, "Automated physical modeling of nonlinear audio circuits for real-time audio effects; Part I: Theoretical development," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 4, pp. 728–737, May 2010.
[292] V. Välimäki, S. Bilbao, J. O. Smith, J. S. Abel, J. Pakarinen, and D. Berners, "Virtual analog effects," in *DAFX: Digital Audio Effects*, U. Zölzer, Ed., 2nd ed. Chichester, U.K.: Wiley, 2011, pp. 473–522.
[293] J. Pakarinen, V. Välimäki, F. Fontana, V. Lazzarini, and J. S. Abel, "Recent advances in real-time musical effects, synthesis, and virtual analog models," *EURASIP J. Adv. Signal Process.*, vol. 2011, pp. 1–15, 2011.
[294] G. De Sanctis and A. Sarti, "Virtual analog modeling in the wave-digital domain," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 4, pp. 715–727, May 2010.
[295] S. Bilbao and J. Parker, "A virtual model of spring reverberation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 4, pp. 799–808, May 2010.
[296] S. Bilbao, *Numerical Sound Synthesis*. Chichester, U.K.: Wiley, 2009.
[297] N. Fletcher, T. Tarnopolskaya, and F. de Hoog, "Wave propagation on helices and hyperhelices: A fractal regression," in *Proc. Math., Phys. Eng. Sci.*, 2001, pp. 33–43.
[298] W. Wittrick, "On elastic wave propagation in helical springs," *Int. J. Mech. Sci.*, vol. 8, no. 1, pp. 25–47, Jan. 1966.
[299] V. Välimäki, J. Parker, and J. S. Abel, "Parametric spring reverberation effect," *J. Audio Eng. Soc.*, vol. 58, no. 7/8, pp. 547–562, Jul./Aug. 2010.
[300] S. Bilbao and J. Parker, "Perceptual and numerical aspects of spring reverberation modeling," in *Proc. 20th Int. Symp. Music Acoust.*, Sydney, Australia, Aug. 2010.
[301] V. Välimäki, J. S. Abel, and J. O. Smith, "Spectral delay filters," *J. Audio Eng. Soc.*, vol. 57, no. 7/8, pp. 521–531, Jul./Aug. 2009.
[302] J. Parker, "Efficient dispersion generation structures for spring reverb emulation," *EURASIP J. Adv. Signal Process.*, vol. 2011, no. 646134, 2011.
[303] S. Mitra and K. Hirano, "Digital all-pass networks," *IEEE Trans. Circuits Syst.*, vol. 21, no. 5, pp. 688–700, Sep. 1974.
[304] H. Gamper, J. Parker, and V. Välimäki, "Automated calibration of a parametric spring reverb model," in *Proc. 14th Int. Conf. Digital Audio Effects*, Paris, France, Sep. 2011, pp. 37–44.
[305] J. D. Parker, H. Penttinen, S. Bilbao, and J. S. Abel, "Modeling methods for the highly dispersive Slinky spring: A novel musical toy," in *Proc. Int. Conf. Digital Audio Effects*, Graz, Austria, Sep. 2010, pp. 123–126.
[306] S. Bilbao, "A digital plate reverberation algorithm," *J. Audio Eng. Soc.*, vol. 55, no. 3, pp. 135–144, Mar. 2007.
[307] A. Huovilainen, "Enhanced digital models for analog modulation effects," in *Proc. Int. Conf. Digital Audio Effects*, Madrid, Spain, 2005, pp. 155–160.
[308] C. Raffel and J. Smith, "Practical modeling of bucket-brigade device circuits," in *Proc. Int. Conf. Digital Audio Effects*, Graz, Austria, Sep. 2010, pp. 50–56.
[309] T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: State of the art and future challenges," *Proc. IEEE*, vol. 99, no. 2, pp. 288–327, Feb. 2011.
[310] M. Heideman, D. Johnson, and C. S. Burrus, "Gauss and the history of the FFT," *Archive History Exact Sci.*, vol. 34, no. 3, pp. 265–277, 1985.

Vesa Välimäki (S'90–M'92–SM'99) received the M.S., the Licentiate of Science, and the Doctor of Science degrees in technology, all in electrical engineering, from the Helsinki University of Technology (TKK), Espoo, Finland, in 1992, 1994, and 1995, respectively. His doctoral dissertation dealt with fractional delay filters and physical modeling of musical instruments.

He was a Postdoctoral Research Fellow at the University of Westminster, London, U.K., in 1996. From 1997 to 2001, he was Senior Assistant (cf. Assistant Professor) at the TKK Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland. From 1998 to 2001, he was on leave as a Postdoctoral Researcher under a grant from the Academy of Finland. In 2001–2002, he was Professor of signal processing at the Pori unit of the Tampere University of Technology, Pori, Finland. Since 2002, he has been Professor of audio signal processing at TKK, which became part of Aalto University in 2010. He was appointed Docent (cf. Adjunct Professor) in signal processing at the Pori unit of the Tampere University of Technology in 2003. In 2006 and 2007, he was the Head of the TKK Laboratory of Acoustics and Audio Signal Processing. In 2008 to 2009, he was on sabbatical and was a Visiting Scholar at the Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA. His research interests include audio effects processing, digital filter design, sound synthesis techniques, and acoustics of musical instruments.

Prof. Välimäki is a member of the Audio Engineering Society and of the Finnish Musicological Society, and a life member of the Acoustical Society of Finland. He was President of the Finnish Musicological Society in 2003–2005. In 2004, he was a Guest Editor of the special issue of the *EURASIP Journal on Applied Signal Processing* on model-based sound synthesis. In 2008, he was the Chairman of DAFx-08, the 11th International Conference on Digital Audio Effects (Espoo, Finland). In 2011, he was a Guest Editor of the special issue of the *EURASIP Journal on Advances in Signal Processing* on musical applications of real-time signal processing. He is currently an Associate Editor of the *Journal of Electrical and Computer Engineering*. In 2000–2001, he was the Secretary of the IEEE Finland Section. From 2005 to 2009 he was an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS, and from 2007 to 2011 he was an Associate Editor of the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING. In 2007, he was the Lead Guest Editor of the special issue of the IEEE SIGNAL PROCESSING MAGAZINE on signal processing for sound synthesis, and in 2010, was the Lead Guest Editor of the special issue of the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING on virtual analog audio effects and musical instruments. He is a member of the Audio and Acoustic Signal Processing Technical Committee of the IEEE Signal Processing Society.

Julian D. Parker received the B.A. (Cantab) degree in natural sciences (astrophysics) from the University of Cambridge, Cambridge, U.K., in 2005 and the M.Sc. degree in acoustics and music technology from the University of Edinburgh, Edinburgh, U.K., in 2008, with a thesis that explored the modeling of spring reverberation using finite difference techniques. He is currently pursuing the Doctoral degree at Aalto University, Espoo, Finland, on the topic of modeling dispersive systems (particularly springs) for musical use.

His research interests are virtual analog modeling, sound synthesis, audio effects, and computer music, focusing on the modeling of complex or nonlinear dynamical systems for musical use. He is also involved in research on the effect of psychoacoustics on music theory. He has previously worked for QinetiQ plc. and for the British Antarctic Survey, and is currently a Researcher at Aalto University.

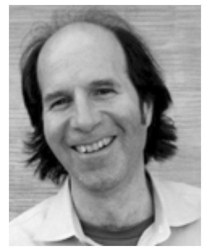
Lauri Savioja (M'00-SM'08) received the M.S., the Licentiate of Science, and the Doctor of Science in technology degrees in computer science from the Helsinki University of Technology (TKK), Espoo, Finland, in 1991, 1995, and 1999, respectively. The topic of his doctoral dissertation was room acoustic modeling.

He worked at the TKK Laboratory of Telecommunications Software, and Multimedia as a Researcher, Lecturer, and Professor from 1995 until the formation of the Aalto University, where he currently works as a Professor in the Department of Media Technology, School of Science, Espoo, Finland. The academic year 2009–2010, he spent as a Visiting Researcher at NVIDIA Research. His research interests include room acoustics, virtual reality, and parallel computation.

Prof. Savioja is a member of the Association for Computing Machinery, the Audio Engineering Society, and a life member of the Acoustical Society of Finland. Since 2010, he has been an Associate Editor for the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING.

Julius O. Smith (M'76) received the B.S.E.E. degree in control, circuits, and communication from Rice University, Houston, TX, in 1975 and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1978 and 1983, respectively. His Ph.D. research was devoted to improved methods for digital filter design and system identification applied to music and audio systems.

From 1975 to 1977, he worked in the Signal Processing Department at ESL, Sunnyvale, CA, on systems for digital communications. From 1982 to 1986, he was with the Adaptive Systems Department at Systems Control Technology, Palo Alto, CA, where he worked in the areas of adaptive filtering and spectral estimation. From 1986 to 1991, he was employed at NeXT Computer, Inc., responsible for sound, music, and signal processing software for the NeXT computer workstation. After NeXT, he became an Associate Professor at the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford, teaching courses and pursuing research related to signal processing techniques applied to music and audio systems. Continuing this work, he is currently a Professor of Music and Associate Professor of Electrical Engineering (by courtesy) at Stanford University. For more information, see http://ccrma.stanford.edu/~jos/.



Jonathan S. Abel (M'09) received the S.B. degree in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 1982, where he studied device physics and signal processing, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1984 and 1989, respectively, focusing his research efforts on statistical signal processing with applications to passive sonar and GPS.

He is currently a Consulting Professor at the Center for Computer Research in Music and Acoustics (CCRMA) in the Music Department, Stanford University. From 1999 to 2007, he was a Co-Founder and Chief Technology Officer of the Grammy Award-winning Universal Audio, Inc., where his research efforts concentrated on reverberation and digital emulation of vintage audio processors. He was a researcher at NASA/Ames Research Center, exploring topics in room acoustics and spatial hearing on a grant through the San Jose State University Foundation. He was also chief scientist of Crystal River Engineering, Inc., where he developed their positional audio technology, and a Lecturer in the Department of Electrical Engineering at Yale University. As an industry consultant, he has worked with Apple, FDNY, LSI Logic, NRL, SAIC, and Sennheiser, on projects in professional audio, GPS, medical imaging, passive sonar, and fire department resource allocation. His research interests include audio and music applications of signal and array processing, parameter estimation, and acoustics.

Prof. Abel is a Fellow of the Audio Engineering Society. In 2011, he was a Guest Editor of the special issue of the *EURASIP Journal on Advances in Signal Processing* on musical applications of real-time signal processing. He won the 1982 Ernst A. Guillemin Thesis Award for the best undergraduate thesis in the Department of Electrical Engineering and Computer Science, MIT.