Spectral Delay Filters

Article *in* Journal of the Audio Engineering Society · July 2009

CITATIONS
29

READS
1,679

3 authors:

Vesa Välimäki
Aalto University
396 PUBLICATIONS **8,298** CITATIONS
SEE PROFILE

Jonathan Abel
Stanford University
150 PUBLICATIONS **3,998** CITATIONS
SEE PROFILE

Julius Smith
Stanford University
355 PUBLICATIONS **9,377** CITATIONS
SEE PROFILE

Spectral Delay Filters

VESA VÄLIMÄKI,[1,2] *AES Member,* **JONATHAN S. ABEL,**[2] *AES Fellow,*
AND JULIUS O. SMITH,[2] *AES Fellow*

[1]*Department of Signal Processing and Acoustics, TKK - Helsinki University of Technology, Espoo, Finland*

[2]*CCRMA, Stanford University, Stanford, CA 94305, USA*

vesa.valimaki@tkk.fi, abel@ccrma.stanford.edu, jos@ccrma.stanford.edu

This paper discusses the implementation of spectral delay using filters comprising a cascade of many low-order allpass filters and an equalizing filter. The spectral delay filters have chirp-like impulse responses causing a large, frequency-dependent delay that is useful in audio effects processing. An equalizing filter design and a multirate technique, which stretches the allpass filters' impulse response, are introduced.

0 INTRODUCTION

Filtering an audio signal with an allpass filter does not usually have a major effect on the signal's timbre. The allpass filter does not change the frequency content of the signal, but only introduces a phase shift or delay. Audibility of the phase distortion caused by an allpass filter in a sound reproduction system has been a topic of many studies, see, e.g., [1], [2]. In this paper, we investigate audio effects processing using high-order allpass filters that consist of many cascaded low-order allpass filters. These filters have long chirp-like impulse responses. When audio and music signals are processed with such a filter, remarkable changes are obtained that are similar to the spectral delay effect [3], [4].

Allpass filters are commonly used in audio and music processing. The allpass filter was first used by Schroeder in a reverberation algorithm to introduce temporal smearing to its impulse response [5]. The Schroeder allpass filters consist of a combination of feedback and feedforward comb filters. Although their structure is similar to the first-order allpass filter, they have a long delay line where a first-order filter has a single unit delay. Griesinger has proposed to use a time-reversed allpass impulse response for room acoustic measurements [6].

First- and second-order allpass filters have become a standard part of phasers [7]. A fractional delay, which is often needed in audio processing to fine-tune the length of a delay-line, can be conveniently implemented using allpass filters [8], [9], [10]. Time-varying first-order allpass filters have also been used to implement a passive nonlinearity in string and percussion instruments models [11] and tension modulation in vibrating string models [12]. Recently, Pekonen suggested the use of a time-varying first-order allpass filter to simulate distortion [13], and Timoney *et al*. proposed an alternative implementation for the phase distortion synthesis technique using the time-varying allpass filter [14], [15]. Shelving filters and equalizers can be realized using a structure that contains a first or a second-order allpass filter [16], [17]. In warped filters, first-order allpass sections replace the unit delays of conventional digital filters. This improves the frequency resolution at low frequencies [18]-[21]. Warped filters can imitate the human hearing or facilitate modelling of acoustic responses. In one imaginative musical application, a warped filter model of guitar body resonances was modulated by varying the warping factor [22]. Warped linear prediction based on a first-order allpass filter chain suits well to modeling of wideband audio signals [23], [24]. Kautz filters use a chain of non-identical second-order allpass filters, achieving practically the same results as warped filters but with a smaller model order [25].

Other uses of the allpass filter in audio processing are based on the fact that they can introduce a highly dispersive (i.e., frequency-dependent) delay. Inharmonizing allpass filters have been used in waveguide synthesis of piano tones [26]-[31], spherical resonators [32], and carillons [33], [34]. Evangelista has shown how to apply time-varying frequency warping in real time to obtain audio effects, such as vibrato or increased inharmonicity [35]. Kleimola produced interesting effects by modulating the inharmonizing allpass filter parameters in a waveguide string algorithm at the audio rate [36]. Abel *et al.* have modelled the response of a spring reverberator using a high-order dispersive allpass filter [37]. Additionally, a high-order allpass filter can be used in adaptive notch filtering [38] or in implementing an accurate time delay needed in a multi-notch filter, which cancels harmonics of a musical signal [39], [40]. In audio coding, allpass filters have been used for decorrelating signals [41], [42]. A new way of selecting weights for a loudspeaker array from an allpass filter's impulse response was recently proposed [43].

In this paper, we introduce a novel use for high-order allpass filters in audio effects processing. We cascade these filters with an equalizing filter and call the overall system a spectral delay filter. A spectral delay filter can be seen as an alternative implementation to the FFT-based spectral delay: Spectral delay is an audio processing method in which different frequencies of a signal experience a different delay. Kim-Boyle has discussed FFT-based spectral delay for stereo and multi-channel effects [3]. He mentioned the idea of

spectral reverb where spectral delay is combined with feedback through a series of allpass filters. Amatriain considered in his PhD work a spectral delay algorithm [4].

Figure 1 shows an example of the spectral delay effect. The spectrograms of the original and processed recording are displayed in Figure 1(a) and Figure 1(b), respectively. In this case, the signal has been filtered by 2000 identical first-order allpass filters. Low frequencies are considerably delayed in comparison to higher frequencies. The delay varies smoothly as a function of frequency.

This paper is organized as follows. In Sec. 1, we discuss the group delay of a cascade of first-order allpass filters and its relation to the chirp-like impulse response of the spectral delay filter. Furthermore, a multirate method to stretch the impulse response of the spectral delay filter is proposed. Section 2 discusses the amplitude envelope of the impulse response and suggests a design method for the equalizing filter. Sec. 3 presents application examples using the spectral delay filter. Sec. 4 concludes this paper.

1 CHIRP-LIKE IMPULSE RESPONSES AND GROUP DELAY

In this section, we discuss the properties of a system that consists of a many cascaded first-order allpass filters. The transfer function of a first-order allpass filter is

$$A(z) = \frac{a_1 + z^{-1}}{1 + a_1 z^{-1}}, \tag{1}$$

where a_1 is the allpass filter coefficient. For stability it is required that $|a_1| < 1$. The transfer function of a cascade of M first-order allpass filters can be written as

$$H(z) = A^M(z) = \left(\frac{a_1 + z^{-1}}{1 + a_1 z^{-1}} \right)^M . \tag{2}$$

The cascade of allpass filters is also allpass, as their transfer functions multiply. We call the cascade of many allpass filters and an equalizing filter cascaded with them a spectral delay filter, shown in Figure 2.

We can first consider the properties of a single allpass filter section and then generalize for the cascade of many identical sections. The group delay of a filter is defined as the negative derivative of the filter's phase response $\phi(\omega)$ with respect to frequency:

$$\tau_g(\omega) = -\frac{d\phi(\omega)}{d\omega}. \tag{3}$$

The phase response of the first-order allpass filter is

$$\phi(\omega) = -\omega + 2\arctan\left(\frac{a_1 \sin\omega}{1 + a_1 \cos\omega}\right). \tag{4}$$

Thus, the group delay of the first-order allpass filter can be expressed in closed form:

$$\tau_g(\omega) = \frac{1 - a_1^2}{1 + 2a_1 \cos\omega + a_1^2}. \tag{5}$$

Figure 3(a) and Figure 3(b) show the group delay with various values of the coefficient a_1.

It is seen that the group delay of first-order allpass filters is typically small at all frequencies. For example, the largest delay observed in Figure 3(a) is obtained with $a_1 = -0.9$, but even in this case the delay is only 19 samples at low frequencies. This corresponds to about 0.4 ms at the 44-kHz sample rate. When an audio signal is filtered with this filter, it remains virtually unchanged. In fact, even when the same filtering is repeated a few times, the effect is hardly audible. However, when the same allpass filter is applied repeatedly many times, for example 100 times, the low and high frequencies in the output signal are separated in time, and a chirp-like effect is heard. One can also listen to the impulse response itself to observe this.

Figure 4(a) shows the impulse response of a cascade of 64 first-order allpass filters with $a_1 = -0.9$. It is seen that the impulse response is now quite long, about 28 ms. In fact, cascading M identical allpass filters produces a high-order allpass filter with M times the original phase and group delay. An impulse applied to the filter, because it is a broadband input, will be smeared into a chirp with frequencies appearing at times specified by the group delay. This is illustrated in Figure 4(b), which displays the spectrogram (linear magnitude) of the impulse response. Overlaid with it (dashed line), we show the group delay of the high-order allpass filter, which is obtained by multiplying the curve corresponding to $a_1 = -0.9$ in Figure 3(a) by 64 and plotting it with the frequency and time (delay) axes interchanged.

It helps to understand what kind of chirps can be generated with allpass filters, when the group delay curves are shown with frequency versus time, as in Figure 5. Cascading M identical first-order allpass sections produces chirps of the same shape, but the time axis must then be multiplied by M.

When the filter coefficient a_1 is negative [see Figure 5(a)] and many impulse responses are convolved, a chirp decreasing in frequency is heard. With a positive a_1 [see Figure 5(b)], a chirp increasing in frequency is produced. Contrary to naïve beliefs it is not always effective to use allpass filters with poles close to the unit circle. When coefficient a_1 is positive and close to 1.0, such as $a_1 = 0.9$, it is seen that the largest delay occurs at very high frequencies, which will be inaudible to humans at a standard sampling rate of 44.1 kHz. The most useful allpass filter designs bring about a large time delay difference between bass frequencies (100 Hz or so) and high treble (well below 10 kHz).

1.1 Estimating the Chirp Duration

It is useful to know the duration of the chirp, since the parameters can be designed in advance. The maximum value of the group delay is related to the length of the chirp. It can be obtained by evaluating the group delay (5) at $\omega = 0$ for non-positive a_1 and at $\omega = \pi$ for positive a_1, that is:

$$\tau_{g,\max} = \begin{cases} \tau_g(0) = \dfrac{1-a_1}{1+a_1}, \text{when } a_1 \le 0 \\ \tau_g(\pi) = \dfrac{1+a_1}{1-a_1}, \text{when } a_1 > 0. \end{cases} \tag{6}$$

Alternatively, it is possible to estimate how long the allpass chirp will last using previously introduced theory for the impulse-response length of IIR filters [44]. The effective duration of the impulse response of a first-order allpass filter that contains P% of the total energy is:

$$\delta_P = \frac{\log(1 - P/100) - \log(1 - a_1^2)}{\log(a_1^2)} - 1, \tag{7}$$

where P is the percentage of energy, e.g., 99.9. Notice that in Eq. (7) we skip the ceiling operation used in [44] to avoid rounding errors, when many allpass filters are cascaded. We also subtract 1 to get the duration in samples instead of the number of samples. For M cascaded allpass filters the effective duration is $M\delta_P$.

In Figure 6 we plot the maximum of the group delay according to Eq. (6) together with the duration of the impulse response that contains 99% and 99.9% of the total energy according to Eq. (7). It is seen that the maximum value of the group delay is between the δ_{99} and $\delta_{99.9}$ values, when $|a_1| > 0.3$. It approaches $\delta_{99.9}$, as $|a_1|$ is increased.

As an example, we can estimate the duration of the spectral delay filter impulse response shown in Figure 4 (M = 64, a_1 = –0.9). Using the maximum of the group delay function, Eq. (6), we get the length estimate of 1216 samples (19.00 samples for one allpass section), which corresponds to 27.57 ms. This is approximately right in comparison to what is seen in Figure 4. Using the energy-based formula with P = 99.9%, the estimated impulse response duration is 1593 samples ($\delta_{99.9}$ = 23.90 samples per filter), which corresponds to 34.69 ms. This estimate seems to be conservative. It appears that none of these impulse-response estimates is extremely accurate. To our knowledge the best, only slightly conservative estimate is $\delta_{99.9}$, although in some cases Eq. (6) can be more accurate.

There is a pure delay in the beginning of the allpass chirp, as is seen in Figure 4 and Figure 5. It may thus be of interest to remove the bulk delay from the duration estimate. The difference of the minimum and maximum of the group delay can approximately account for this. It can be written as:

$$\tau_{g,\mathrm{diff}} = \left| \frac{1-a_1}{1+a_1} - \frac{1+a_1}{1-a_1} \right| = \left| \frac{(1-a_1)^2 - (1+a_1)^2}{1-a_1^2} \right| = \frac{4|a_1|}{1-a_1^2}. \tag{8}$$

1.2 Stretched Allpass Filters

In this section, we introduce a multirate method to get a more dramatic effect with a small number of filters. By replacing the unit delay in each allpass filter with two or more delay elements, the chirp becomes slower. Let K be the number of unit delays in each allpass filter. We call this a stretched allpass filter.

Figure 7(a) shows the impulse response of a stretched spectral delay filter, when each allpass section contains a pair of unit delays (K = 2) instead of one. The impulse response is twice as long as in Figure 4(a), and its every second sample is zero. This method produces an image chirp at high frequencies, which propagates in the opposite direction. This is seen in the spectrogram display in Figure 7(b). Since one of the chirps now takes place in the lower half of the audio band (between 0 Hz and 11.025 kHz at 44.1 kHz), it is heard more easily than in the case of a single full-band chirp.

2 AMPLITUDE EQUALIZATION

The envelope of the chirp produced by a cascade of allpass filters is not constant over time. This is clearly seen in Figure 4(a). The amplitude envelope results from the fact that the impulse response contains an equal amount of energy at all frequencies, and therefore the chirp is smaller at frequencies where it lingers. Next we derive the envelope $\alpha(\omega)$ of an allpass chirp. This derivation has previously appeared in [45]. After that

we suggest an equalizing filter design method for the spectral delay filter.

2.1 Allpass Chirp Envelope

Denote by ω_+ and ω_- two nearby frequencies having difference Δ and mean ω, that is:

$$\Delta = \omega_+ - \omega_- , \tag{9}$$

$$\omega = (\omega_+ - \omega_-)/2 . \tag{10}$$

The allpass chirp has energy Δ/π in the frequency interval $[\omega_+, \omega_-]$. That energy is roughly equal to the impulse response energy in the time interval $[\tau_g(\omega_+),\tau_g(\omega_-)]$, during which the impulse response sweeps between frequencies ω_+ and ω_-. Denoting by $\alpha(\omega)$ the amplitude of the chirp while sweeping through frequency ω, we have

$$\Delta/\pi \approx \left|\tau_g(\omega_-) - \tau_g(\omega_+)\right| \alpha^2(\omega)/2 . \tag{11}$$

Taking the limit $\Delta \rightarrow 0$, gives

$$\alpha(\omega) = \left[\frac{\pi}{2}\left|\frac{d\tau_g(\omega)}{d\omega}\right|\right]^{-\frac{1}{2}} . \tag{12}$$

Thus, an allpass chirp may be filtered by an invertible normalization $\nu(\omega)$ with a magnitude chosen to approximate the inverse of its envelope $\alpha(\omega)$,

$$\left|\nu(\omega)\right| \approx \frac{1}{\alpha(\omega)} = \sqrt{\frac{\pi}{2}\left|\frac{d\tau_g(\omega)}{d\omega}\right|} \quad . \tag{13}$$

Provided the equalizing filter does not smear the phase of the signal too much, the resulting sequence will have a nearly constant envelope. One equalizing filter is sufficient for a cascade of many identical allpass filters. This is explained by the fact that the shape of the group delay function remains unchanged, when many allpass filters are cascaded.

In the case of the first-order filter, inverse envelope $1/\alpha(\omega)$ can be expressed in closed form:

$$\frac{1}{\alpha(\omega)} = \sqrt{\frac{\pi}{2}\left|\frac{d\tau_g(\omega)}{d\omega}\right|} = \frac{\sqrt{\left|\pi a_1(1-a_1^2)\sin\omega\right|}}{1+2a_1\cos\omega+a_1^2}. \tag{14}$$

The frequency derivative of the group delay in Eq. (13) is proportional to the number of filters in the cascade, M. So, when M identical first-order allpass filters are cascaded, we can simply replace $d\tau_g(\omega)/d\omega$ with $Md\tau_g(\omega)/d\omega$ in Eq. (13). This shows that the frequency response of Eq. (14) scaled by $\sqrt{M}$ will then be sufficient for equalization.

2.2 Allpass Chirp Equalization

The allpass chirp has an envelope which varies in time as a function of the group delay. An envelope equalization (EQ) filter for a cascade of first-order allpass filters can be designed by approximating Eq. (14). The motivation for this filter is that for longer chirps, a constant envelope allows slowly changing chirp frequencies to be more clearly heard. For real-time signal processing, we propose to use a standard second-order parametric EQ filter [16]. We choose the following transfer function, which has a nominal gain g_0 (instead of the more usual 1):

$$H_{\mathrm{EQ}}(z) = \frac{g_0}{2}\,\frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + (1+a)bz^{-1} + az^{-2}}, \tag{15a}$$

$$b_0 = 1 + a + g_K(1-a), \tag{15b}$$

$$b_1 = 2b(1+a), \tag{15c}$$

$$b_2 = [1 + a + g_K(a-1)], \tag{15d}$$

where $a = [1 - \tan(\Omega/2)]/[(1 + \tan(\Omega/2)]$, $b = -\cos(\omega_c)$, g_K is the peak gain, Ω is the bandwidth in radians, and ω_c is the center frequency in radians. This equalizing filter introduces a boost of $20\log(g_K)$ dB at the peak frequency ω_c. Its gain is chosen to be $20\log(g_0)$ dB at a selected low frequency, such as 1 Hz. The gain of the filter is also close to $20\log(g_0)$ dB at the Nyquist limit, when the center frequency is not very high.

The center frequency ω_c of the EQ filter can be selected to be the frequency at which the inverse envelope peaks. For the first-order allpass filter, the peak of the envelope $\alpha(\omega)$ can be found in closed form by differentiating Eq. (12) without the square root and setting the derivative equal to zero. This yields

$$\omega_c = \arccos\left(\frac{a_1^2 + 1 - \sqrt{a_1^4 + 34a_1^2 + 1}}{4a_1}\right). \tag{16}$$

The peak frequency is displayed in Hertz in Figure 8(a).

To approximate the shape of the required spectral envelope, the magnitude response of the equalizer is set equal to the gain of $1/\alpha(\omega)$ at two points, the peak frequency and a low frequency, such as 20 Hz. In addition, the –3-dB bandwidth of the equalizer is set approximately the same as that measured from the inverse envelope.

The EQ gain can be set to be the peak magnitude of the inverse envelope. To find the peak magnitude for a given value of coefficient a_1, evaluate Eq. (14) at ω_c obtained from Eq. (15). The peak gain as function of a_1 is shown in Figure 8(b).

The gain of the inverse envelope is zero at DC (0 Hz), because sin(0) = 0 in Eq. (14). However, it is easy to approximate the magnitude response of Eq. (14) at a low frequency, because $\sin(\omega) \approx \omega$ and $\cos(\omega) \approx 1$. This yields the following function:

$$\frac{1}{\alpha(\omega_{low})} = \sqrt{\frac{\pi a_1 (1 - a_1)\omega_{\text{low}}}{(1 + a_1)^3}}. \tag{17}$$

The –3-dB bandwidth for the equalizer can be measured from the ideal inverse envelope function Eq. (13). We give a third-order polynomial approximation for the bandwidth:

$$\omega_{\text{bw}} = 2.21 - 2.49|a_1| - 1.16a_1^2 + 1.48|a_1|^3. \tag{18}$$

The approximation is symmetric around $a_1 = 0$ and is based on the right half of the function shown in Figure 9. The approximation error remains small when a_1 is close to ±1, but is larger around $a_1 = 0$.

As an example, we equalize the impulse response of Figure 4(a). Figure 10(a) compares the inverse of the envelope of the allpass chirp of Figure 4(a) and the magnitude response of an EQ filter. The EQ filter response has been matched to the inverse envelope, as suggested above. In this case the parameters of the EQ filter are the following: The center frequency is 427 Hz, the bandwidth is 760 Hz, the peak gain g_K is 13.6, which corresponds to 22.6 dB, and we have chosen to have the responses meet at 1 Hz by choosing g_0 = 0.875.

Figure 10(b) shows the equalized impulse response of the spectral delay filter when the above EQ filter has been applied. It is seen that the chirp has an almost constant amplitude except at the very beginning of the chirp. The reason for this deviation is that the magnitude response of the second-order equalizing filter overshoots the inverse envelope at frequencies above 1.5 kHz, but it also undershoots the inverse envelope

below the center frequency, as seen in Figure 10(a). A more accurate equalization of the allpass chirp can be realized by designing a higher order IIR filter approximation of Eq. (14). In practice, the second-order IIR EQ filter used in this example may be sufficient.

An interpolated version of the EQ filter can be used for the stretched spectral delay filter, that is,

$$H_{\mathrm{EQ,s}}(z) = H_{\mathrm{EQ}}(z^K). \tag{19}$$

Alternatively, the EQ filter can be used to equalize only the first chirp (i.e., the lowest one in frequency). In this case, the center frequency and bandwidth parameters of the equalizing filter, Eq. (15), must be replaced with $\omega_{cs} = \omega_c/K$ and $\Omega_s = \Omega/K$, respectively, but otherwise the EQ filter parameters and structure are the same as in the regular case.

The use of the equalization filter is most important for chirps that are long enough so that their temporal details are heard. When the chirp is very short, less than about 20 ms, it resembles a click. In such cases, the equalization may not be necessary, because it introduces coloration but does not make the chirp more audible.

3 APPLICATIONS

The spectral delay filter can have many uses in audio effects processing. The chirp-like impulse responses of the spectral delay filter are reminiscent of early video game sounds, such as imagined laser gun sounds. The chirps can also be sampled and used as part of synthetic percussion. However, interesting applications of the spectral delay filter stem from the fact that it is a filter and can process any input signal. This leads to applications that Roads calls “cross-synthesis by convolution” [46], [47]: convolution of an audio signal and a short signal yields a convolution result that shares properties of both signals. When audio signals are filtered with the spectral delay filter, they can retain the original rhythm, but the chirps caused by the spectral delay filter are also audible.

As an example, a snare drum sample is processed with the spectral delay filter. Figure 11 shows the un-equalized impulse response of a spectral delay filter and a drum recording. When the drum sample is filtered with the spectral delay filter, its attack part changes, as seen in Figure 11(c). The spectrogram of the resulting sound, given in Figure 12(c), shares most of the temporal and spectral features of the original snare drum sound, but the spectral smearing caused by the allpass chirp appears clearly at frequencies below 1

kHz. This can sound similar to a synthetic drum sound or a sample filtered with a resonant filter, which has a sweeping cut-off frequency.

Unexpected results can be produced by a stretched spectral delay filter with K much larger than 2. Figure 13 shows the spectrogram of one such effect. In this example, each allpass filter has a delay line of 15 samples, and 150 stretched allpass filters are cascaded with an equalizer. The allpass filter coefficient is $a_1 = -0.9$ producing a basic chirp that increases in frequency. However, the stretching brings about 14 additional image chirps, and every second one of them decreases in frequency, as seen in Figure 13. The outcome does not resemble a drum sound, but rather a strange utterance.

An extreme case is obtained using a cascade of very many allpass filters. Figure 1(b) of this paper shows an example of the spectral delay effect caused by 2000 first-order allpass filters. This operation disturbs the rhythm of the recorded song, as transients are blurred into long sweeps, and low frequencies are delayed by more than half a second in comparison to high frequencies. We call this effect the shooting star. The impulse response of the spectral delay filter that produces this effect is shown in Figure 14.

In musical applications, it may be useful to mix the original and the processed sound and possibly apply equalization to both, as shown in Figure 15. The reason is that the timbre of processed sound may be quite different from the original, for example the middle frequencies may dominate, while the very low frequencies may be missing. According to our experiments, this happens often, when an equalized spectral delay filter is used, because the frequency range containing the slowest chirp is emphasized. The original signal may be used in the mix to bring back some of the original signal features.

Although the cascaded allpass filters are identical in the examples given in this paper, interesting results can also be obtained by using allpass filters with a different coefficient value. An equalizing filter is needed for each different allpass filter coefficient, if it is desired to equalize the overall impulse response. In general, the spectral delay filter can be used to bring electronic or synthesizer-like characteristic to audio signals. In this sense it is related to previous extended synthesis methods that can process arbitrary audio signals, such as the audio signal driven sound synthesis methods proposed by Poepel and Dannenberg [48], or the adaptive FM synthesis and adaptive split sideband synthesis methods introduced by Lazzarini *et al.* [49], [50].

4 CONCLUSION

The spectral delay filter was introduced as a new way to implement spectral delay. A digital filter consisting of a cascade a many allpass filters and an equalizing filter can produce unusual effects on an audio signal.

When the filter has an extended, chirp-like impulse response, the various frequency components contained in an audio signal travel through the filter at different speeds. A multirate algorithm, called the stretched spectral delay filter, was also proposed that improves the audibility of the chirp in two ways: It increases the chirp length and shifts the chirp to a frequency range where human hearing is most sensitive.

It was shown that the amplitude envelope of the allpass chirp is proportional to the inverse root time delay frequency derivative. Thus, the slowest portions of the chirp are also quietest. The allpass chirp can be equalized by applying a standard second-order parametric equalization filter, which approximates the inverse envelope. Since the group delay of the cascade of many identical allpass filters is proportional to that of a single allpass filter, the same equalization filter may be used, regardless of the order of the cascade. The transfer function of the equalizer must simply be scaled by the square root of M, the number of cascaded filters.

The spectral delay filter can be used to process arbitrary audio signals to transform their timbre. This is computationally more efficient than convolution, since the length of the impulse response of the spectral delay filter can be long. Alternatively, the impulse responses produced with spectral delay filters can be used as sound effects or synthetic percussion samples.

5 ACKNOWLEDGMENTS

This research was conducted in fall 2008 when Vesa Välimäki was a visiting scholar at CCRMA, Stanford University. His visit was financed by the Academy of Finland (project no. 126310). The authors would like to Dr. Henri Penttinen for his comments and for the snare drum sample used in this work.

6 REFERENCES

[1] D. Preis, "Phase Distortion and Phase Equalization in Audio Signal Processing—A Tutorial Review," *J. Audio Eng. Soc*, vol. 30, no. 11, pp. 774–794 (1982 Nov.).

[2] H. Møller, P. Minnaar, S. Krarup Olesen, F. Christensen, and J. Plogsties, "On the Audibility of All-Pass Phase in Electroacoustical Transfer Functions," *J. Audio Eng. Soc*, vol. 55, no. 3, pp. 115–134 (2007 Mar.).

[3] D. Kim-Boyle, "Spectral Delays with Frequency Domain Processing," in *Proc. Int. Conf. Digital Audio Effects* (DAFx-04) (Naples, Italy, 2004 Oct.), pp. 42–45.

[4] X. Amatriain, *An Object Oriented Metamodel for Digital Signal Processing with a Focus on Audio and Music*, PhD thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2004, pp. 184–188. See also: http://www.create.ucsb.edu/~xavier/Thesis/html/node127.html

[5] M. R. Schroeder, "Natural Sounding Reverberation," *J. Audio Eng. Soc.*, vol. 10, no. 3, pp. 219–223 (1962).

[6] D. Griesinger, "Impulse Response Measurements Using All-pass Deconvolution," presented at the *AES 11th Int. Conf. on Test and Measurement* (1992 May).

[7] J. O. Smith, "An Allpass Approach to Digital Phasing and Flanging," Rep. STAN-M-21, Center for Computer Research in Music and Acoustics (CCRMA), Dept. of Music, Stanford University, Stanford, CA (1984).

[8] D. A. Jaffe and J. O. Smith, "Extensions of the Karplus-Strong Plucked-String Algorithm," *Computer Music J.*, vol. 7, no. 2, pp. 76–87, 1983. Reprinted in C. Roads, ed. 1989. *The Music Machine*. Cambridge, MA: MIT Press.

[9] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the Unit Delay—Tools for Fractional Delay Filter Design," *IEEE Signal Processing Magazine*, vol. 13, no. 1, pp. 30–60 (1996 Jan.).

[10] J. Dattorro, "Effect Design—Part II: Delay-Line Modulation and Chorus," *J. Audio Eng. Soc.*, vol. 45, no. 10, pp. 764–788 (1997 Nov.).

[11] J. R. Pierce and S. A. Van Duyne, "A Passive Nonlinear Digital Filter Design which Facilitates Physics-Based Sound Synthesis of Highly Nonlinear Musical Instruments," *J. Acoust. Soc. Am.*, vol. 101, no. 2, pp. 1120–1126 (1997 Feb.).

[12] J. Pakarinen, V. Välimäki, and M. Karjalainen, "Physics-Based Methods for Modeling Nonlinear Vibrating Strings," *Acta Acustica united with Acustica*, vol. 91, pp. 312–325 (2005).

[13] J. Pekonen, "Coefficient-Modulated First-Order Allpass Filter as Distortion Effect," in *Proc. Int. Conf. Digital Audio Effects* (DAFx-08), Espoo, Finland, pp. 83–87 (2008 Sept.).

[14] J. Timoney, V. Lazzarini, J. Pekonen, and V. Välimäki, "Spectrally Rich Phase Distortion Sound Synthesis Using an Allpass Filter," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing* (Taipei, Taiwan, 2009 May), pp. 293–296.

[15] J. Timoney, V. Lazzarini, B. Carty, and J. Pekonen, "Phase and Amplitude Distortion Methods for Digital Synthesis of Classic Analogue Waveforms," presented at the 126th Convention of the Audio Engineering Society (Munich, Germany, 2009 May), convention paper 7792.

[16] P. A. Regalia and S. K. Mitra, "Tunable Digital Frequency Response Equalization Filters," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 35, no. 1, pp. 118–120 (1987 Jan.).

[17] F. Fontana and M. Karjalainen, "A Digital Bandpass/Bandstop Complementary Equalization Filter with Independent Tuning Characteristics," *IEEE Signal Processing Letters*, vol. 10, no. 4, pp. 119–122 (2003 April).

[18] M. Karjalainen and J. O. Smith, "Body Modeling Techniques for String Instrument Synthesis," in *Proc. Int. Computer Music Conf.* (ICMC'96) (Hong Kong, 1996 Aug.), pp. 232–239.

[19] J. O. Smith and J. S. Abel, "Bark and ERB Bilinear Transforms," *IEEE Trans. Speech and Audio*

Processing, vol.7, no. 6, pp. 697–708 (1999 Nov.).

[20] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi, "Frequency-Warped Signal Processing for Audio Applications," *J. Audio Eng. Soc.*, vol. 48, no. 11, pp. 1011–1031 (2000 Nov.).

[21] G. Evangelista, "Time and Frequency Warping Musical Signals," in U. Zölzer (ed.), *Digital Audio Effects*. Wiley, Chichester, UK, 2002, pp. 439–464.

[22] H. Penttinen, A. Härmä, and M. Karjalainen, "Digital Guitar Body Mode Modulation with One Driving Parameter," in *Proc. COST-G6 Conf. Digital Audio Effects (DAFx'00)* (Verona, Italy, 2000 Dec.), pp. 31–36.

[23] A. Härmä and U. K. Laine, "A Comparison of Warped and Conventional Linear Predictive Coding," *IEEE Trans. Speech and Audio Processing*, vol. 9, no. 5, pp. 579–588 (2001 July).

[24] T. van Waterschoot and M. Moonen, "Comparison of Linear Prediction Models for Audio Signals," *EURASIP J. Audio, Speech, and Music Processing*, vol. 2008, Article ID 706935, 24 pages, doi:10.1155/2008/706935 (2008 Dec.).

[25] T. Paatero and M. Karjalainen, "Kautz Filters and Generalized Frequency Resolution: Theory and Audio Applications," *J. Audio Eng. Soc.*, vol. 51, no. 1/2, pp. 27–44 (2003 Jan./Feb.).

[26] S. A. Van Duyne and J. O. Smith, "A Simplified Approach to Modeling Dispersion Caused by Stiffness in Strings and Plates," in *Proc. Int. Computer Music Conf.*, Aarhus, Denmark (1994 Sept.), pp. 407–410.

[27] D. Rocchesso and F. Scalcon, "Accurate Dispersion Simulation for Piano Strings," in *Proc. Nordic Acoustical Meeting*, Helsinki, Finland (1996 June), pp. 407–414.

[28] I. Testa, G. Evangelista, and S. Cavaliere, "Physically Inspired Models for the Synthesis of Stiff Strings with Dispersive Waveguides," *EURASIP J. Applied Signal Processing*, vol. 2004, no. 1, pp. 964–977 (2004).

[29] J. Rauhala and V. Välimäki, "Tunable Dispersion Filter Design for Piano Synthesis," *IEEE Signal Processing Letters*, vol. 13, no. 5, pp. 253–256 (2006 May).

[30] J. S. Abel and J. O. Smith, "Robust Design of Very High-Order Dispersive Allpass Filters," in *Proc. Int. Conf. Digital Audio Effects* (DAFx-06), Montreal, Canada (2006 Sept.), pp. 13–18.

[31] J. Rauhala, M. Laurson, V. Välimäki, V. Norilo, and H.-M. Lehtonen, "Parametric Piano Synthesizer," *Computer Music J.*, vol. 32, no. 4, pp. 17–30 (2008).

[32] D. Rocchesso and P. Dutilleux, "Generalization of a 3-D Acoustic Resonator Model for the Simulation of Spherical Enclosures," *EURASIP J. Applied Signal Processing*, vol. 1, no. 1, pp. 15–26 (2001).

[33] M. Karjalainen, V. Välimäki, and P. A. A. Esquef, "Efficient Modeling and Synthesis of Bell-Like Sounds," in *Proc. Int. Conf. Digital Audio Effects* (DAFx-02) (Hamburg, Germany, 2002 Sept.), pp. 181–

186.

[34] M. Karjalainen, P. A. A. Esquef, and V. Välimäki, "Making of a Computer Carillon," in *Proc. Stockholm Music Acoustics Conf.* (SMAC03) (Stockholm, Sweden, 2003 Aug.), pp. 339–342.

[35] G. Evangelista, "Real-Time Time-Varying Frequency Warping via Short-Time Laguerre Transform," in *Proc. COST G-6 Conf. Digital Audio Effects* (DAFX-00) (Verona, Italy, 2000 Dec.), pp. 7–12.

[36] J. Kleimola, "Dispersion Modulation Using Allpass Filters," in *Proc. Int. Conf. Digital Audio Effects* (DAFx-08) (Espoo, Finland, 2008 Sept.), pp. 193–197.

[37] J. S. Abel, D. P. Berners, S. Costello, and J. O. Smith, "Spring Reverb Emulation Using Dispersive Allpass Filters in a Waveguide Structure," presented at the AES 121st Convention (San Francisco, CA, 2006 Oct.), paper no. 6954.

[38] J. E. Cousseau, P. D. Donate, and Y. Liu, "Factorized All-pass IIR Adaptive Notch Filters," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, (2004 May), vol. 2, pp. 661–664.

[39] V. Välimäki, M. Ilmoniemi, and M. Huotilainen, "Decomposition and Modification of Musical Instrument Sounds Using a Fractional Delay Allpass Filter," in *Proc. Nordic Signal Processing Symposium* (NORSIG 2004) (Espoo, Finland, 2004 June), pp. 208–211.

[40] H.-M. Lehtonen, V. Välimäki, and T. I. Laakso, "Canceling and Selecting Partials from Musical Tones Using Fractional-Delay Filters," *Computer Music J.*, vol. 32, no. 2, pp. 43–56 (2008).

[41] J. Engdegård, H. Purnhagen, J. Rödén, and L. Liljeryd, "Synthetic Ambience in Parametric Stereo Coding," presented at the 116th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 52, no. 7/8, pp. 800–801 (2004 July/Aug.), convention paper 6074.

[42] J. Herre, K. Kjörling, J. Breebaart, C. Faller, S. Disch, H. Purnhagen, J. Koppens, J. Hilpert, J. Rödén, W. Oomen, K. Linzmeier, and K. S. Chong, "MPEG Surround—The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding," *J. Audio Eng. Soc.*, vol. 56, no. 11, pp. 932–955 (2008 Nov.).

[43] M. M. Goodwin, "All-Pass Linear Arrays," *J. Audio Eng. Soc.*, vol. 56, no. 12, pp. 1090–1101 (2008 Dec.).

[44] T. I. Laakso and V. Välimäki, "Energy-Based Effective Length of the Impulse Response of a Recursive Filter," *IEEE Trans. Instrumentation and Measurement*, vol. 48, no. 1, pp. 7–17 (1999 Feb.).

[45] J. S. Abel and D. P. Berners, "MUS424/EE367D: Signal Processing Techniques for Digital Audio Effects," Unpublished Course Notes, CCRMA, Stanford University, Stanford, CA (2005).

[46] C. Roads, "Musical Sound Transformation by Convolution," in *Proc. Int. Computer Music Conf.* (Tokyo, Japan, 1993), pp. 102–109.

[47] C. Roads, *The Computer Music Tutorial.* MIT Press, Cambridge, MA, 1996.

[48] C. Poepel and R. Dannenberg, "Audio Signal Driven Sound Synthesis," in *Proc. Int. Computer*

Music Conf. (Barcelona, Spain, 2005), pp. 391–394.

[49] V. Lazzarini, J. Timoney, and T. Lysaght, "The Generation of Natural-Synthetic Spectra by Means of Adaptive Frequency Modulation," *Computer Music J*., vol. 32, no. 2, pp. 9–22 (2008).

[50] V. Lazzarini, J. Timoney, and T. Lysaght, "Non-Linear Distortion Synthesis Using the Split Sideband Method, with Applications to Adaptive Signal Processing," *J. Audio Eng. Soc.*, vol. 56, no. 9, pp 684–695, (2008 Sept.).

7 FIGURE CAPTIONS

Figure 1. Spectrogram of a pop music excerpt: (a) original signal and (b) the same signal after filtering with 2000 first-order allpass filters ($a_1 = -0.9$). The spectrogram on a logarithmic frequency scale has been produced by evaluating the discrete-time Fourier transform of the signal in 2048-sample frames at 256 logarithmically spaced frequencies. The hop size is 256 samples.

Figure 2. The spectral delay filter consists of M allpass filters and an equalization filter.

Figure 3. Group delay of the first-order allpass filter with (a) non-positive ($a_1 \leq 0$) and (b) non-negative ($a_1 \geq 0$) coefficient values. The sampling rate is $f_s = 44.1$ kHz.

Figure 4. (a) Impulse response of the cascade of 64 first-order allpass filters ($a_1 = -0.9$) and (b) its spectrogram (linear magnitude of the discrete-time Fourier transform evaluated at 256 logarithmically located frequencies with the Blackman window and a hop size 1 sample), which shows a downwards chirp. The dashed line is the group delay of one allpass filter ($a_1 = -0.9$) multiplied by 64.

Figure 5. The group delay curves of Figure 3 displayed with interchanged time and frequency axes. These curves show the shape of the chirps produced by these allpass filters.

Figure 6. Three different estimates for the effective duration of a first-order allpass filter impulse response as a function of coefficient a_1: the group delay maximum and the energy-based effective duration for 99.0% and 99.9% of the total energy.

Figure 7. (a) Impulse response of a cascade of 64 stretched allpass filters ($a_1 = -0.9$) when each section contains two unit delays ($K = 2$) and (b) its spectrogram computed with the same parameters as in Figure 4. Now two chirps appear going in opposite directions in frequency.

Figure 8. The peak frequency [see Eq. (16)] and peak gain of the equalizer for a spectral delay filter as a function of allpass filter coefficient a_1. The peak frequency has been multiplied by $f_s/(2\pi)$, where is $f_s = 44.1$ kHz.

Figure 9. Measured bandwidth of the equalizer and its polynomial approximation [Eq. (18) multiplied by $f_s/(2\pi)$].

Figure 10. (a) The inverse envelope of an allpass chirp and an equalization filter's magnitude response. (b) The equalized impulse response [cf. Figure 4(a)].

Figure 11. (a) Impulse response of a spectral delay filter ($a_1 = -0.9$, $M = 100$, $K = 2$, non-equalized), (b) waveform of a snare drum sample, and (c) the snare drum sample after filtering with (a).

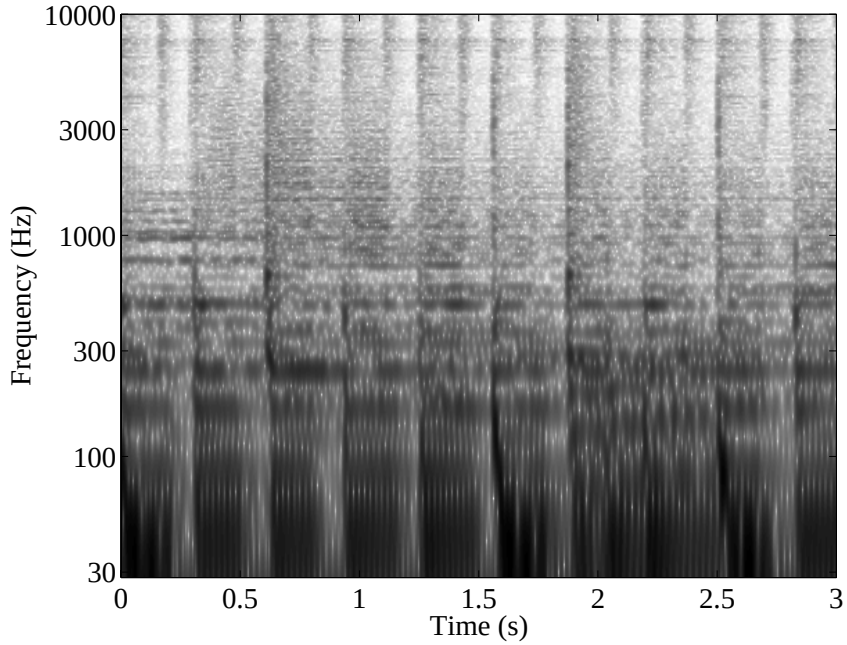
Figure 12. Spectrograms of signals in Figure 11: (a) spectral delay filter impulse response (shifted by 10 ms for visualization), (b) snare drum sample, and (c) snare drum sample filtered with the spectral delay filter.

Figure 13. Spectrogram of a special effect ("alien hiccup") that is produced by filtering a snare drum sample with a stretched spectral delay filter ($a_1 = 0.6$, $M = 150$, $K = 15$).

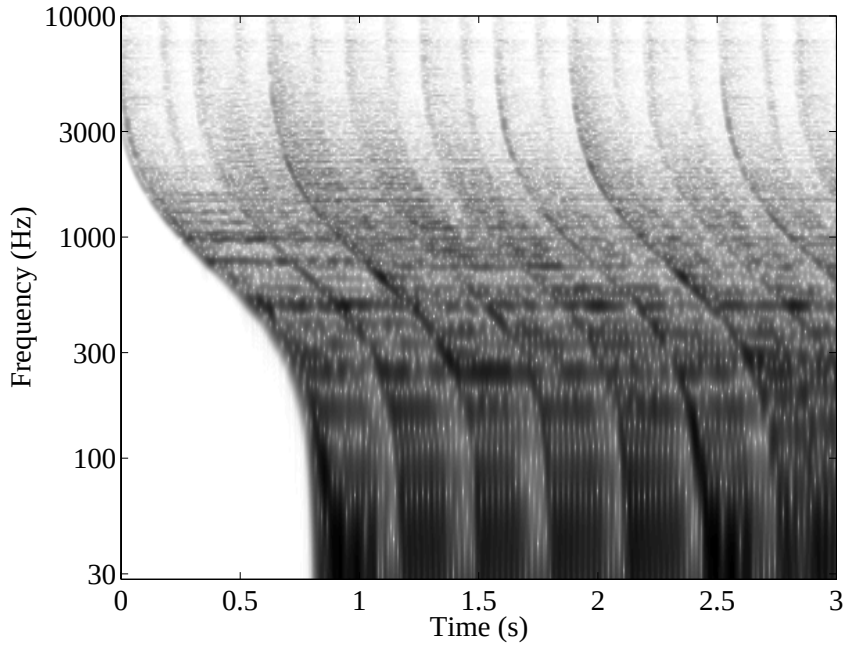
Figure 14. Impulse response of the spectral delay filter used for producing the shooting star effects displayed in Figure 1(b). The spectral delay filter consists of 2000 first-order allpass filters ($a_1 = -0.9$) and a second-order equalizer.

Figure 15. Mixing the spectral delay effect to an arbitrary audio signal $x(n)$. The tone control may be used to give room for the effect for example at middle frequencies. Coefficients g_1 and g_2 are adjusted to mix the dry and wet signal appropriately.

8 FIGURES



(a)



(b)

Figure 1. Spectrogram of a pop music excerpt: (a) original signal and (b) the same signal after filtering with 2000 first-order allpass filters ($a_1 = -0.9$). The spectrogram on a logarithmic frequency scale has been produced by evaluating the discrete-time Fourier transform of the signal in 2048-sample frames at 256 logarithmically spaced frequencies. The hop size is 256 samples.

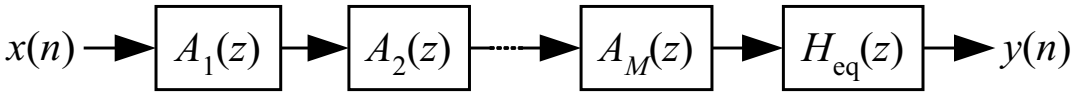
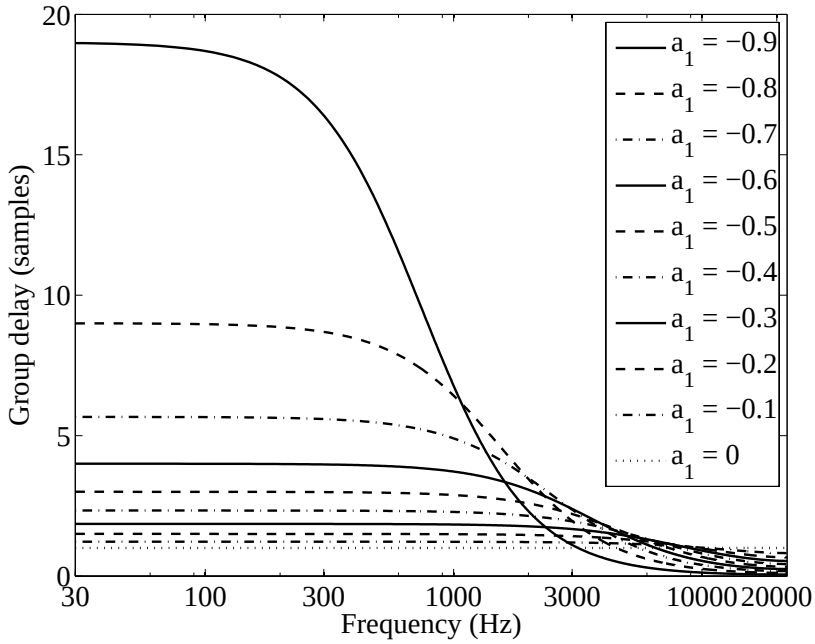
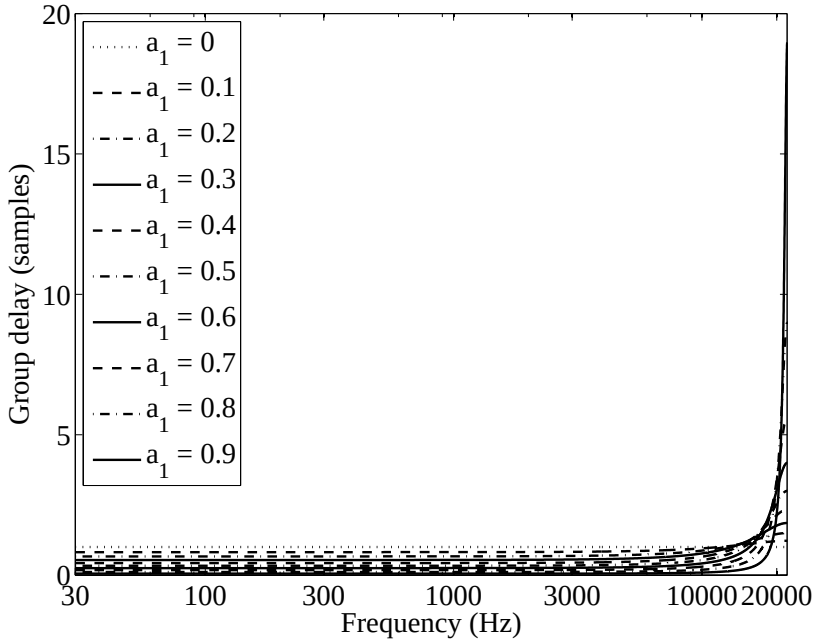


Figure 2. The spectral delay filter consists of *M* allpass filters and an equalization filter.



(a)



(b)

Figure 3. Group delay of the first-order allpass filter with (a) non-positive ($a_1 \leq 0$) and (b) non-negative ($a_1 \geq 0$) coefficient values. The sampling rate is $f_s = 44.1$ kHz.

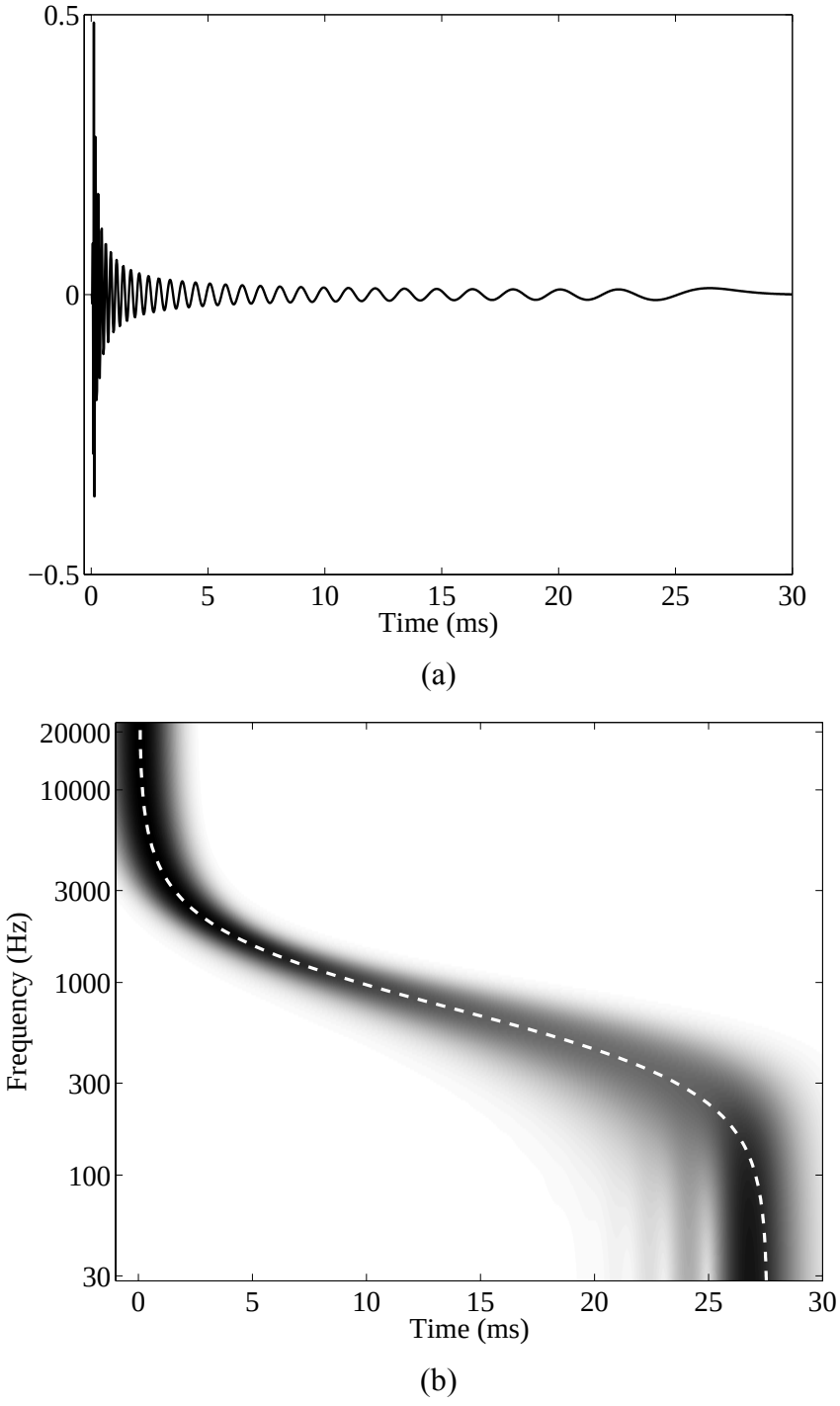
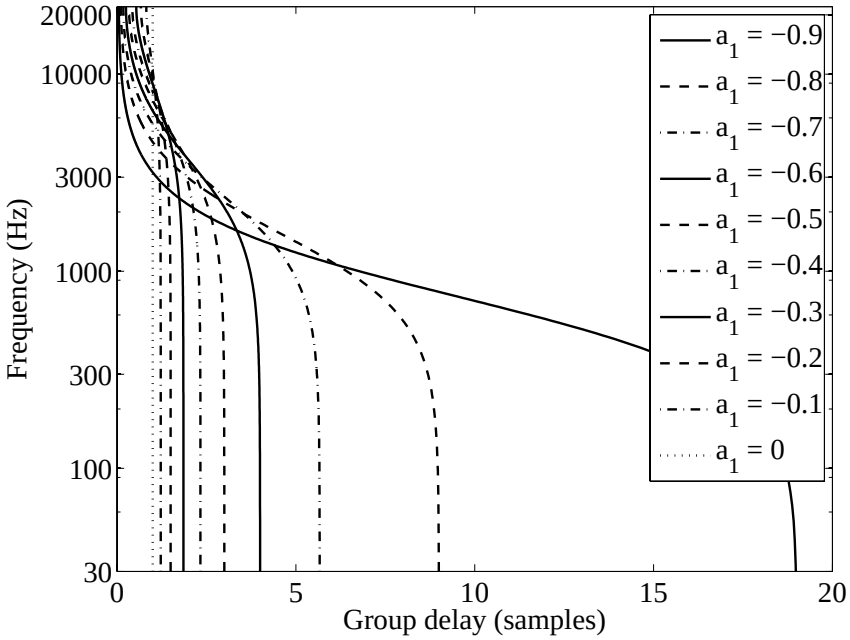
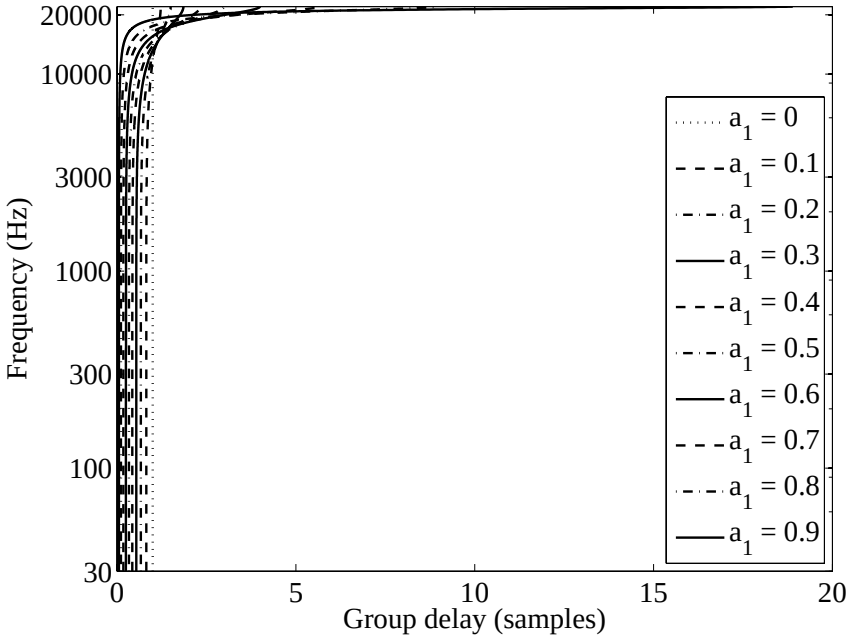


Figure 4. (a) Impulse response of the cascade of 64 first-order allpass filters (a_1 = –0.9) and (b) its spectrogram (linear magnitude of the discrete-time Fourier transform evaluated at 256 logarithmically located frequencies with the Blackman window and a hop size 1 sample), which shows a downwards chirp. The dashed line is the group delay of one allpass filter (a_1 = –0.9) multiplied by 64.



(a)



(b)

Figure 5. The group delay curves of Figure 3 displayed with interchanged time and frequency axes. These curves show the shape of the chirps produced by these allpass filters.

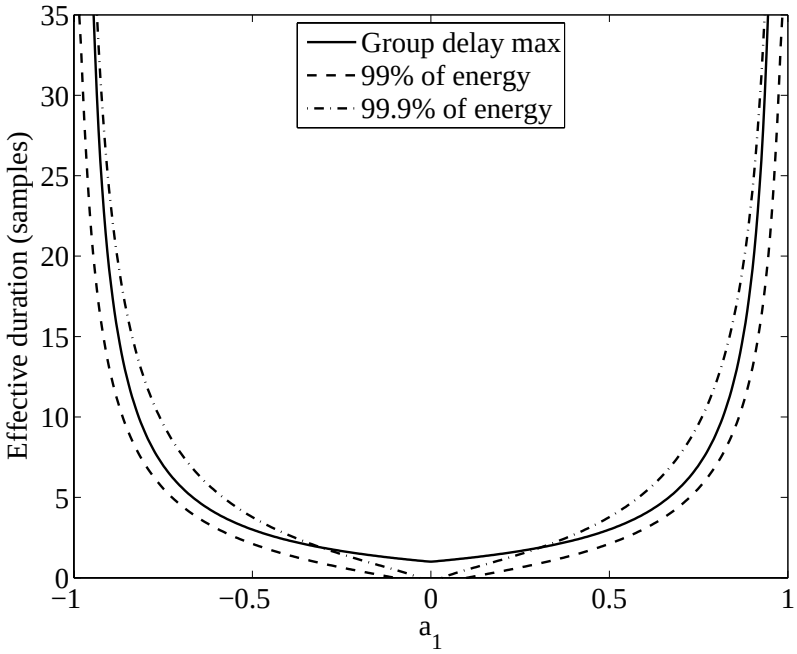


Figure 6. Three different estimates for the effective duration of a first-order allpass filter impulse response as a function of coefficient a_1: the group delay maximum and the energy-based effective duration for 99.0% and 99.9% of the total energy.

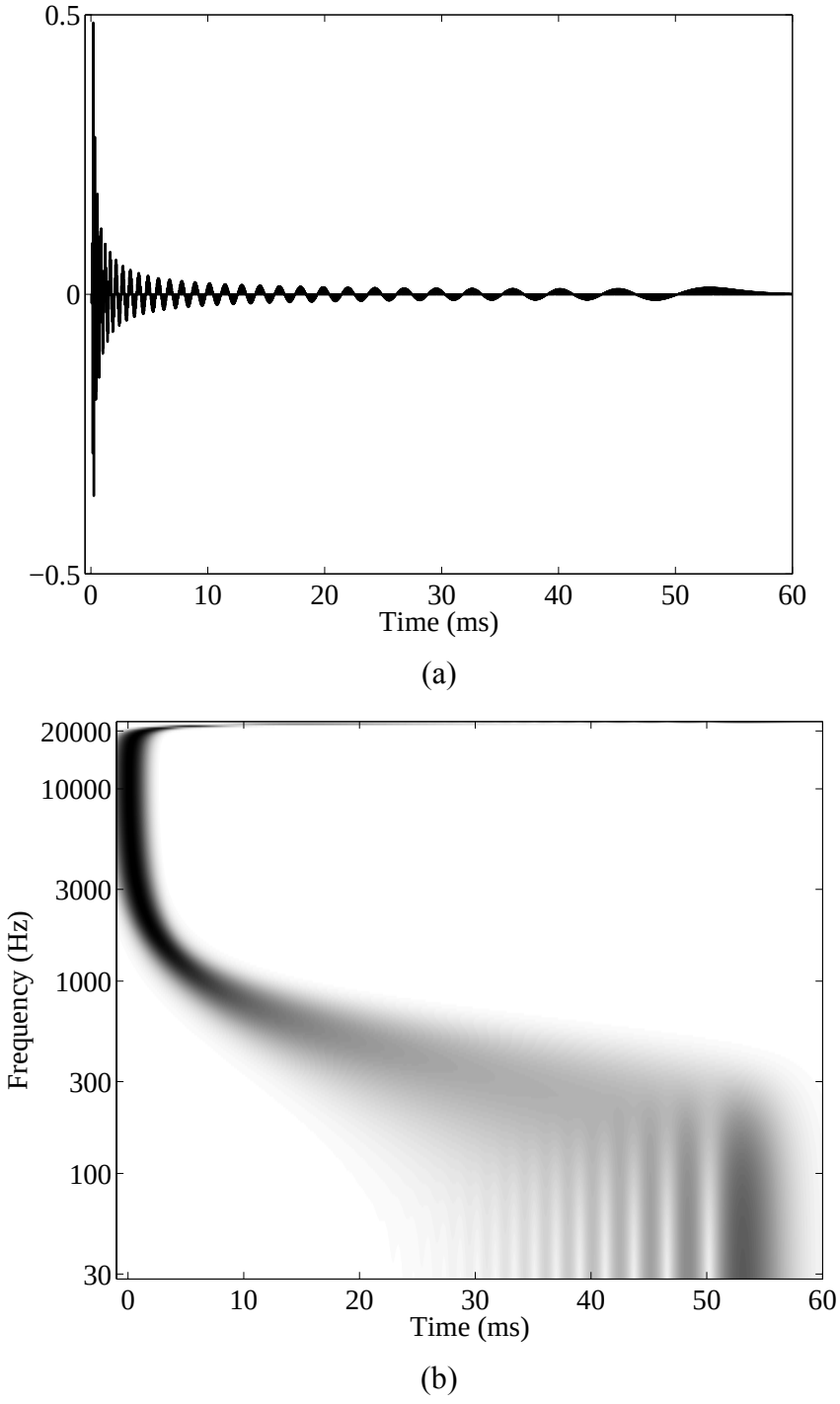


Figure 7. (a) Impulse response of a cascade of 64 stretched allpass filters ($a_1 = –0.9$) when each section contains two unit delays ($K = 2$) and (b) its spectrogram computed with the same parameters as in Figure 4. Now two chirps appear going in opposite directions in frequency.

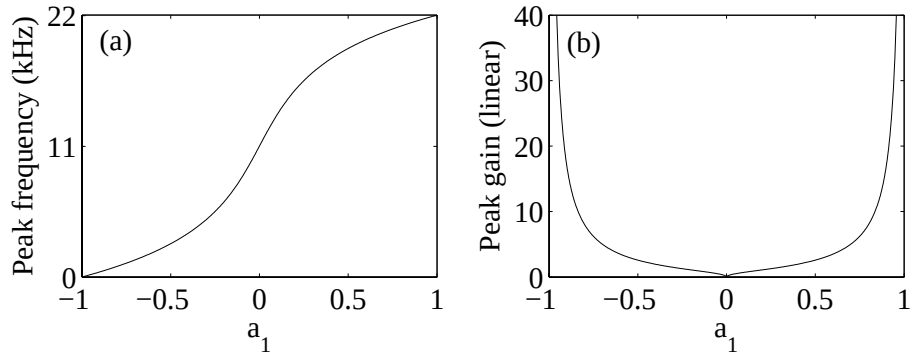


Figure 8. The peak frequency [see Eq. (16)] and peak gain of the equalizer for a spectral delay filter as a function of allpass filter coefficient a_1. The peak frequency has been multiplied by $f_s/(2\pi)$, where is f_s = 44.1 kHz.

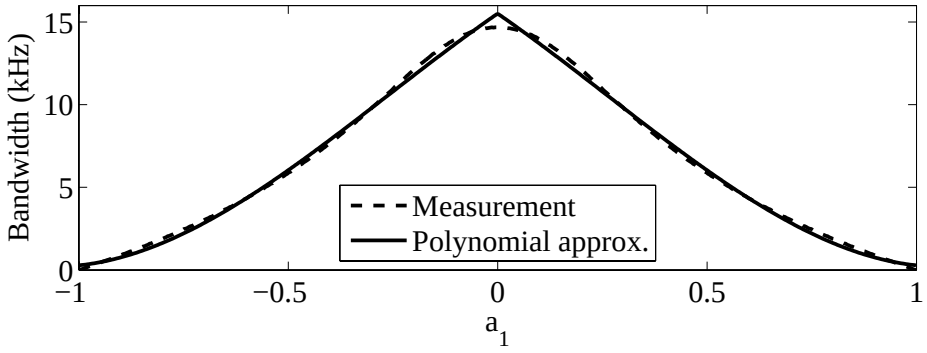


Figure 9. Measured bandwidth of the equalizer and its polynomial approximation [Eq. (18) multiplied by $f_s/(2\pi)$].

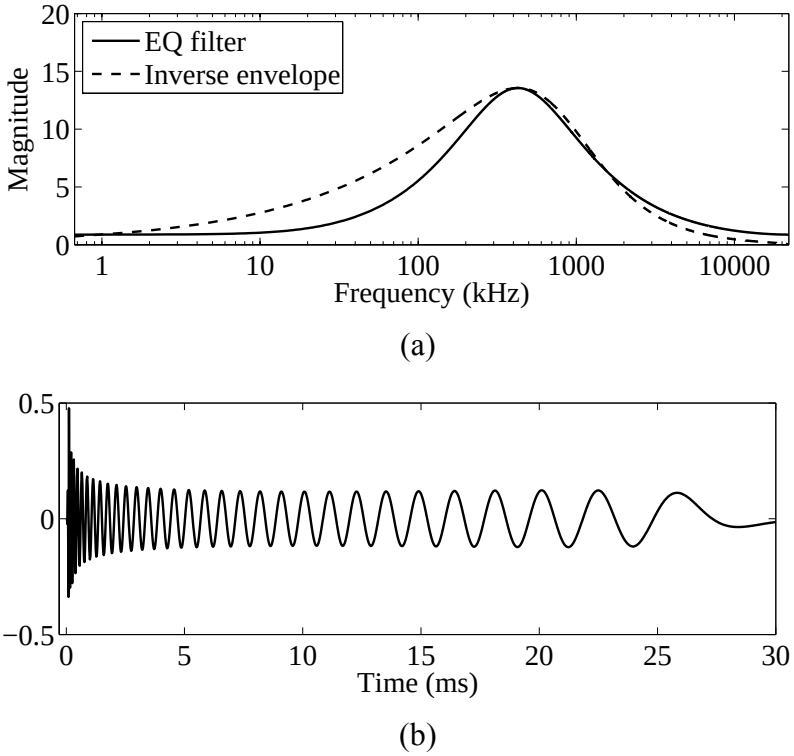


Figure 10. (a) The inverse envelope of an allpass chirp and an equalization filter's magnitude response. (b) The equalized impulse response [cf. Figure 4(a)].

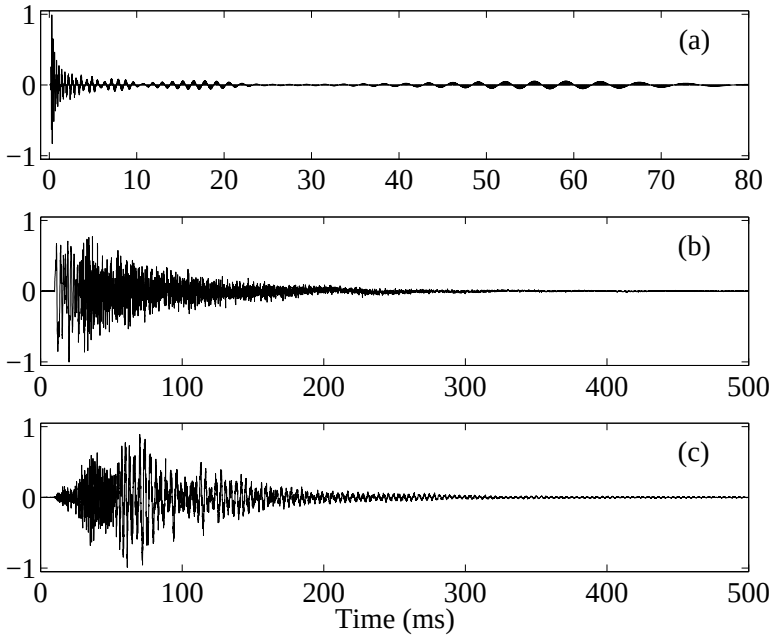


Figure 11. (a) Impulse response of a spectral delay filter (a_1 = –0.9, M = 100, K = 2, non-equalized), (b) waveform of a snare drum sample, and (c) the snare drum sample after filtering with (a).

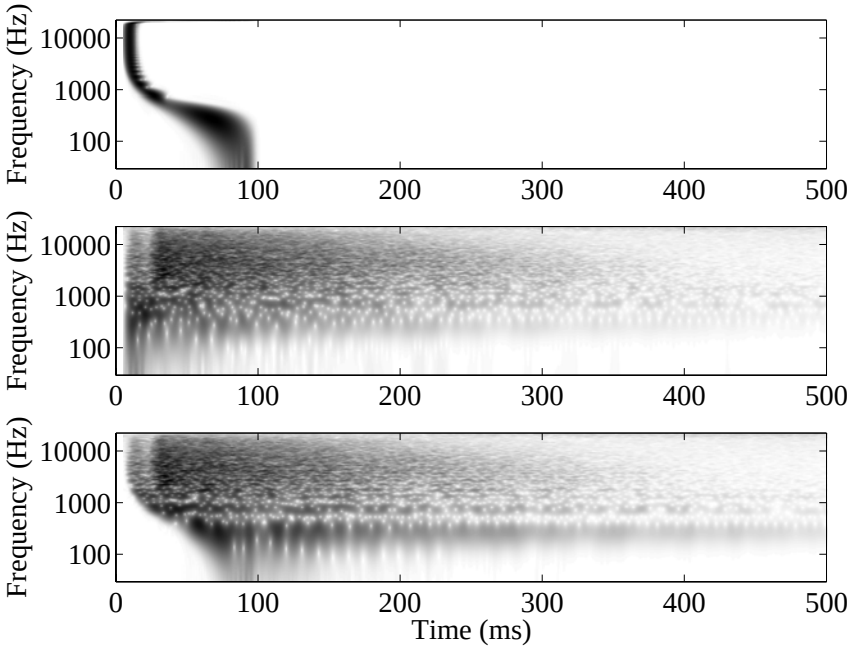


Figure 12. Spectrograms of signals in Figure 11: (a) spectral delay filter impulse response (shifted by 10 ms for visualization), (b) snare drum sample, and (c) snare drum sample filtered with the spectral delay filter.

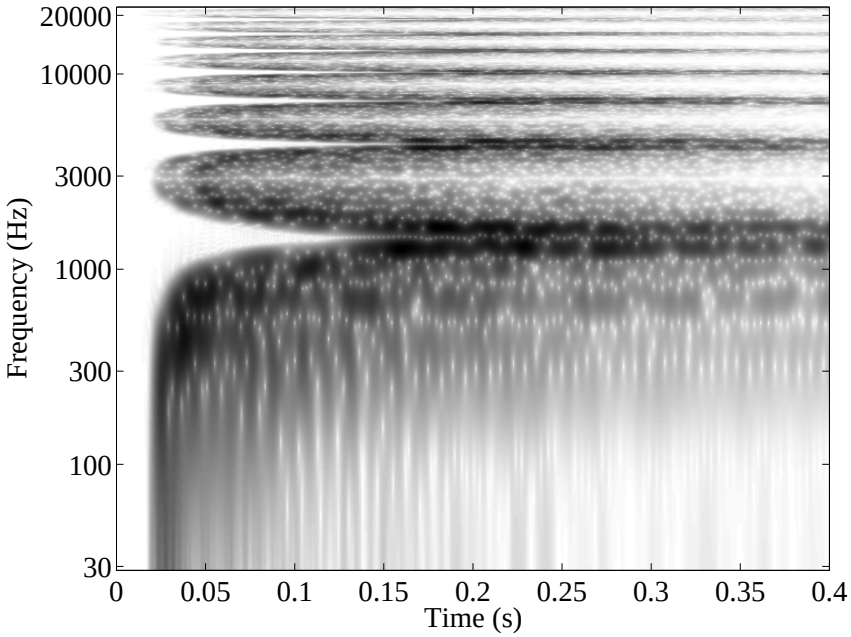


Figure 13. Spectrogram of a special effect ("alien hiccup") that is produced by filtering a snare drum sample with a stretched spectral delay filter ($a_1 = 0.6$, $M = 150$, $K = 15$).

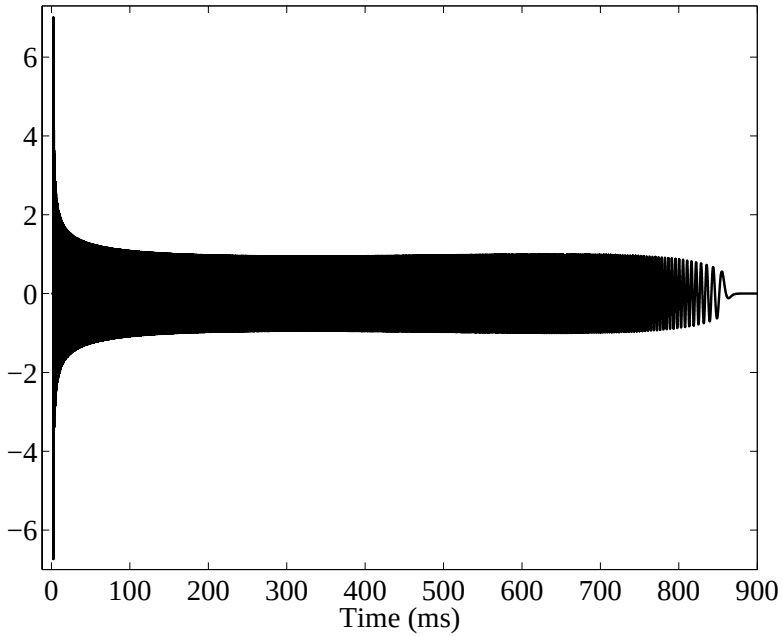


Figure 14. Impulse response of the spectral delay filter used for producing the shooting star effects displayed in Figure 1(b). The spectral delay filter consists of 2000 first-order allpass filters ($a_1 = –0.9$) and a second-order equalizer.

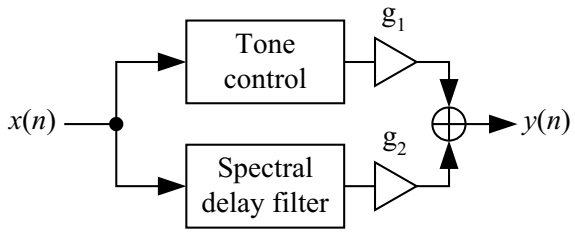


Figure 15. Mixing the spectral delay effect to an arbitrary audio signal $x(n)$. The tone control may be used to give room for the effect for example at middle frequencies. Coefficients g_1 and g_2 are adjusted to mix the dry and wet signal appropriately.