

ENERGY BEHAVIOR IN TIME-VARYING FRACTIONAL DELAY FILTERS FOR PHYSICAL MODELING SYNTHESIS OF MUSICAL INSTRUMENTS

Jyri Pakarinen, Matti Karjalainen, Vesa Välimäki

Helsinki University of Technology
Lab. of Acoustics and Audio Signal Processing
P.O.Box 3000, FI-02015 HUT, Finland

Stefan Bilbao

Sonic Arts Research Centre,
Queens University Belfast
United Kingdom, BT7 1NN

ABSTRACT

Time-varying fractional delays are applied for example in physics-based modeling of musical instruments, particularly for string and wind instruments. While Lagrange interpolation and allpass filters are used routinely in such sound synthesis models, they are found somewhat problematic for example in plucked string simulation when the length of the string is varied due to glissando or vibrato. There can be problems with signal energy levels and aliasing. In this paper we study two variable delay filter designs that have physically realistic energetic behavior and keep undesirable side effects, such as aliasing, in control. The first one is sliding termination point simulation with energy correction and the second one is based on controllable wave digital filter delay lines.

1. INTRODUCTION

Modeling and sound synthesis of string and wind instruments is realized typically by digital waveguides (DWG) [1] or by their simplified version, the Karplus-Strong algorithm [2, 3, 4]. Figure 1(a) shows the basic idea of a DWG, based on a two-directional delay line and terminating wave reflection filters. For example a plucked string model is excited by feeding a pulse into the lines and the output may be the force at the bridge or a magnetic pickup signal.

Models based on the Karplus-Strong principle reduce the DWG into a single-delay loop [4], see Fig. 1(b). It consist of the total delay of the DWG and a loop filter implementing losses and dispersion in the string. The model is typically extended in various ways, such as using submodels for the two polarizations of vibration in a string. In both DWG and KS models, fractional delays are needed to adjust the pitch correctly. In many instruments the length of the delays must be continuously controllable, which introduces the specific problems discussed in this paper.

Fractional delays needed in the modeling can be realized and approximated in various ways [5, 6]. The most popular principles in physics-based modeling are the allpass filters and Lagrange interpolation. Figure 2(a) shows the case where an integer length part z^{-M} is cascaded with a fractional part made as a first-order allpass filter. It has the advantage of ideal magnitude response, but it requires special techniques to avoid transients when the delay parameter is changed in a stepwise manner [7].

Figure 2(b) depicts a typical implementation of fractional delay by Lagrange interpolation. It consists of an integer length part z^{-M} and an FIR structure with controllable tap coefficients, in this case a third order interpolator. Typically third or fifth order interpolators are used to make the non-ideal magnitude response flat enough [6]. The transient behavior is more easily in control by Lagrange interpolation than with the allpass structure [6].

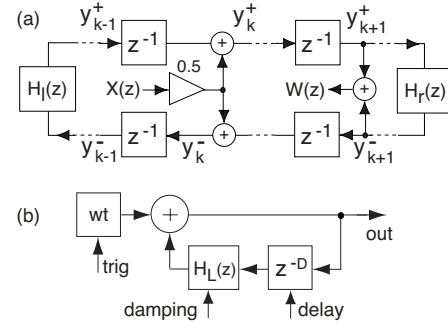


Fig. 1. (a) The principle of digital waveguide for simulating two-directional wave propagation and (b) a reduced version of it, a single loop Karplus-Strong type of model.

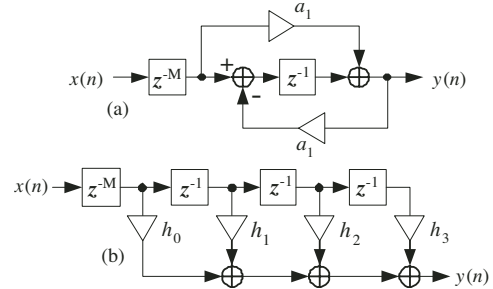


Fig. 2. Practical realizations of controllable fractional delays: (a) allpass structure and (b) third order Lagrange interpolator.

In this paper we start from an assumption that an ideal fractional delay filter with controllable delay is available. We can approach this for example by choosing the order of a Lagrange interpolator large enough. An ideal fractional delay D consisting of an integer part D_{int} and a fractional part d , $D = D_{\text{int}} + d$, is obtained as a bandlimited interpolation of signal in a delay line, which is easily show to correspond to infinite order sinc function interpolation

$$y(n) = x(n - D) = \sum_{k=-\infty}^{\infty} x(k) \text{sinc}(n - D - k) \quad (1)$$

By limiting the range of summation and using proper weight coefficients instead of sinc function values, the Lagrange interpolator results, which is a maximally flat FIR approximation of ideal fractional delay. The computation of tap coefficients for an N th order Lagrange interpolation is done according to formula:

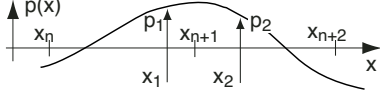


Fig. 3. Definitions of sliding point interpolation. The point moves from x_1 to x_2 in a sample period. Corresponding interpolated signal values are p_1 and p_2 . Points x_n are integer delay points.

$$h(n) = \prod_{k=0, k \neq n}^N \frac{D-k}{n-k} \quad \text{for } n = 0, 1, 2, \dots, N \quad (2)$$

Two methods will be studied in this paper to guarantee energetically correct behavior of varying delay lines in physics-based modeling. In Section 2, energy-preserving fractional delay techniques are developed and the aliasing properties are discussed. In Section 3, the use of energy-preserving Wave Digital Filters (WDF) is introduced.

2. SLIDING POINT INTERPOLATION

Let us start for simplicity reasons from the single delay loop KS type structure of Fig. 1(b), although it does not correspond to two-directional wave propagation for example in a string.

2.1. Energetic behavior

Let us first look at an analog transmission line with a wave propagating only in one direction in a medium, for example in a tube such as a cylindrical bore of a wind instrument (Fig. 3). The energy E of the wave within a section (x_1, x_2) of the delay-line is

$$\Delta E = E(x_1, x_2) = \int_{x_1}^{x_2} \frac{p^2(x)}{Z_0} dx \quad (3)$$

where p is a wave variable such as the sound pressure, x is the longitudinal coordinate, and Z_0 is wave impedance properly formulated from an energy point of view, assumed to be constant over the medium. For discrete-time computation we can interpret Eq. (3) so that position variable x directly represents the fractional delay, i.e., $x = D = D_{\text{int}} + d$.

Let us now assume that we probe the wave variable p , traveling in the positive direction of x -axis, and feed it back to the starting point $x = 0$ of the delay line. Although this is not a physically realistic two-directional model, it helps us to deal with the EKS model of Fig. 1(b) that is important in practical sound synthesis.

If the position of interpolation point makes a jump of Δx from x_1 to x_2 , the amount of energy gain/loss ΔE is obtained according to Eq. (3). To conserve the energy in the delay line loop, we have to compensate for ΔE when feeding back the interpolated value to the beginning of the delay line.

Accurate computation of the energy correction ΔE is a numerically expensive task because the interpolated waveform should be known everywhere between x_1 and x_2 . It is thus meaningful to find simple approximations for it. Let us first consider the case where the frequency of a wave traveling in the line is very low so that wavelength $\lambda \ll |\Delta x|$, where $\Delta x = x_2 - x_1$. Thus we can assume that the value of p is practically constant within the range (x_1, x_2) . This can be called the *zeroth order energy preserving interpolation*. In order to compensate the gain or loss of energy due to the step $x_1 \rightarrow x_2$, we have to use a corrected output value p_c probed from the delay line so that

$$\frac{p_c^2}{Z_0} = \frac{p^2}{Z_0} - \Delta E = \frac{p^2}{Z_0} - \Delta x \frac{p^2}{Z_0} \quad (4)$$

By solving for p_c we obtain

$$p_c = \sqrt{1 - \Delta x} p = g_c p \quad (5)$$

which for small steps $\Delta x \ll D_1$, where D_1 corresponds to a unit step in the delay line, can be approximated by

$$p_c \approx (1 - \frac{\Delta x}{2}) p = g_{ca} p \quad (6)$$

These compensation gain factors g_c and g_{ca} are useful for small steps per sample interval and low-frequency waves.

When the signal frequency is not low or the jump Δx is large, more accurate approximation of ΔE can be tried. Next we study the case where the signal output value from the delay line is known in two points x_1 and x_2 . This needs in general the fractional delay interpolation to be computed twice, first for the previous sample position x_1 and then for the new position x_2 . Let us assume that the signal values obtained are p_1 and p_2 . Then the energy change ΔE can be approximated by assuming linear variation of p between x_1 and x_2 , where x_1 is the starting point of interpolation and x_2 is the new interpolation point. Also let us denote $p = a\xi + b$ for linear interpolation of p between x_1 and x_2 so that $b = p_1$ and $a = p_2 - p_1$ and $\xi = x - x_1$. From Eq. (3) we obtain

$$\Delta E(x) = \int_0^{\Delta x} \frac{p^2(\xi)}{Z_0} d\xi = \frac{1}{Z_0} \int_0^{\Delta x} (a\xi + b)^2 d\xi \quad (7)$$

$$= \frac{1}{Z_0} \left\{ \frac{(p_2 - p_1)^2}{3} (\Delta x)^3 + (p_2 - p_1)p_1 (\Delta x)^2 + p_1^2 \Delta x \right\} \quad (8)$$

Similarly to Eq. (4) we obtain

$$p_c = p_2 \sqrt{1 - \frac{(p_2 - p_1)^2}{3p_2^2} (\Delta x)^3 - \frac{(p_2 - p_1)p_1}{p_2^2} (\Delta x)^2 - \frac{p_1^2}{p_2^2} \Delta x} \quad (9)$$

It is obvious from Eq. (9) that this *first order energy preserving interpolation* is computationally expensive because it requires two interpolations to get p_1 and p_2 and then computing the corrected term p_c from Eq. (9) for every sample period. Even higher order correction terms can be derived, using a higher number of spatially sampled points between x_1 and x_2 , but the complexity needed hardly motivates using them in practice for real-time implementations. In Section 2.3 we will evaluate the method developed above.

The study of sliding interpolation above was made for the KS type of single delay loop structure. A physically more relevant model is the dual delay line structure of Fig. 1(a), where moving the interpolation point has to be applied to both delay lines. In fact the interpolation must be done from one line through termination filter into the other one. The latter case can be done by deinterpolation techniques [6]. The study of the energetic behavior of such a case is left for future.

2.2. Aliasing behavior

Another important problem with delay loop models using time-varying delays is aliasing. When the length of the delay changes, the signal is resampled so that the frequencies are scaled accordingly. This is similar to the Doppler effect in acoustics where signal frequencies are changed if the source or the receiver moves.

In shortening the delay the frequency components are scaled up in frequency and there is a risk of aliasing, which happens in a very specific way. When the interpolation point of the delay line changes by Δx per sample period, frequency f is scaled to a new value f' by rule $f' = (1 - \Delta x)f$, where $\Delta x = 1$ means lengthening by a unit step. For shortening of the delay Δx is negative and the frequency is scaled up. For continuous shortening

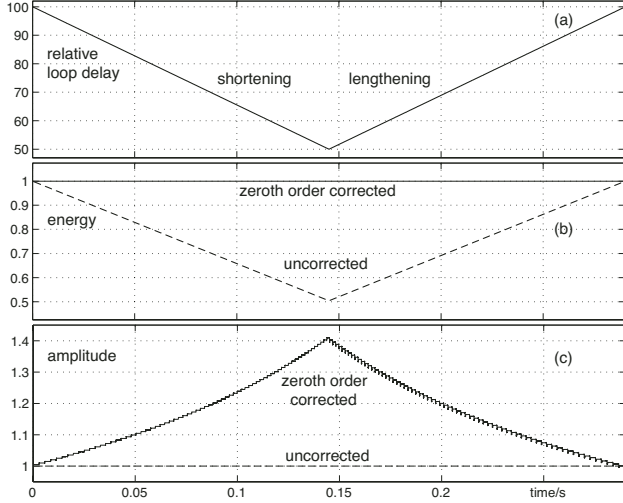


Fig. 4. Evaluation of fractional delay filter with (solid line) and without (dashed line) energy compensation: (a) sliding loop delay from 100 % (128 taps) to 50 % and back to 100 %, (b) energy level, and (c) signal amplitude level.

of delay, frequencies close to the Nyquist frequency can thus get aliased in a normal way. If the sliding of the delay is then stopped, something interesting happens. The waveform in the delay loop, see Fig. 1(b), continues looping around regularly, which means that the signal must again become harmonic, and the aliased signal components have contributed to the harmonic components of the new state. Therefore it is not possible to observe inharmonically aliased signal in the new steady state after the sliding period.

In the delay loop case the aliasing is more severe than in the case of just reading from a sliding point of a delay line. For signal components that have travelled around in the loop the final frequency scaling factor is proportional to the relative delay length so that $f' = (l_1/l_2)f$, where l_2 is the new length and l_1 is the original length. For example in a guitar string, if the length slides from open string to fret #12, i.e., the length becomes one half of the full length, and harmonic frequencies get doubled independently of the speed of sliding. Fortunately, the high frequencies after plucking a string die out quite rapidly so that the aliasing problem is not necessarily very severe in practice. It can be alleviated further by using oversampling, which also helps to improve the magnitude response of Lagrange type of FIR interpolation.

2.3. Evaluation and experiments

The energy-compensated and uncompensated fractional delay filters were compared using a lossless KS loop model. Interpolation was computed through FFT as a 256 point interpolation, being close to ideal. The delay was changed from 128 taps to 64 taps and back to 128 taps in about 300 ms for the sample rate of 44100 Hz with step size $|\Delta x| = 0.01$. This approximates a fast sliding of guitar string termination by an octave up from 345 Hz and back.

For illustration purposes, see Fig. 4, the delay line was initially filled with a DC signal value, but similar results of amplitude and energy envelopes are obtained with signal frequencies not too close to the Nyquist limit. In Fig. 4(a) the energy of the filter corrected according to Eq. (6) stays constant while the uncorrected one loses energy in relation to the delay length, but gains it back when the original length is resumed. Figure 4(b) shows how the

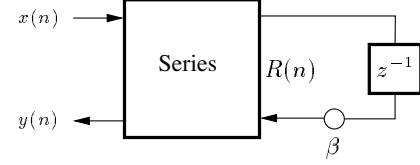


Fig. 5. Wave digital one-port, corresponding to a generalization of a first-order allpass section.

amplitude in the corrected case increases 3 dB when the delay is halved, and returns to the original value when the delay is resumed. The first-order energy correction does not improve the model behavior noticeably in practical cases.

The energy preserving KS loop model was successfully used also in a real-time electric guitar model calibrated after the Fender Stratocaster. A fifth-order Lagrange interpolator with zeroth-order energy correction was found perfect enough for good sustain.

3. TIME-VARYING WDF DELAY LINE

This section will briefly consider a varying-tension string model, implemented using time-varying wave digital one-ports.

A first-order allpass filter, relating an input sequence x_n to an output sequence y_n , is defined as

$$y(n+1) = \beta (ax(n+1) + x(n)) - ay(n) \quad (10)$$

with $\beta = \pm 1$, and is stable for $|a| < 1$. It also possesses the energy-conserving property

$$\sum_{n=-\infty}^{\infty} (y(n))^2 = \sum_{n=-\infty}^{\infty} (x(n))^2 \quad (11)$$

The matter of generalizing such a filter to the time-varying case is a delicate matter; though, for slowly varying coefficients a_n , the above recursion is a perfectly acceptable generalization. For faster variation, there may be some variation in the energy (squared norm) as per (11), and for extreme variation, the energy can become infinite even for finite energy input signals [9]. For physical modeling applications in particular, it is very useful to have a guarantee of signal energy conservation.

One means of generalizing such an allpass is through the use of wave digital filters, and in particular those of the power-normalized variety. A wave-digital one-port is shown in Fig. 5 where we refer to [10] for a description of basic WDF principles. The important property of such a design is that, regardless of time-variation of the terminating port resistance $R(n)$, the energy-conserving property (11) holds (this is a consequence of the orthogonality of the scattering matrix in the power-normalized case). The design shown in Fig. 5 can be rewritten as a recursion

$$y(n+1) = \beta (a(n+1)x(n+1) + \phi(n)x(n)) - \phi(n)a(n)y(n) \quad (12)$$

where

$$\phi(n) = \sqrt{\frac{1 - a(n+1)^2}{1 - a(n)^2}} \quad (13)$$

and reduces to (10) when $a(n) = a$, a constant.

We can now construct a string model by connecting these power-normalized one-ports in a loop, as illustrated in Fig. 6. Since the a parameter of Eq. (12) controls the phase delay of the one-ports, we can also vary the total loop delay and the resulting fundamental frequency of the string through a . It is important to note that while this WDF string is a single delay loop model such as the

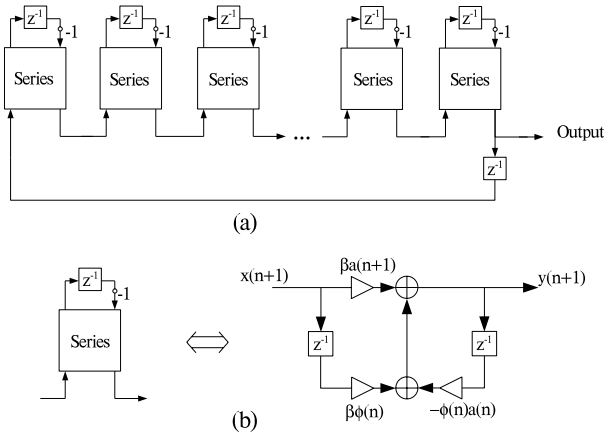


Fig. 6. (a) The WDF string consists of power-normalized wave digital one-ports, connected in a loop. If the wave digital formalism is used, the string can also be seen as a loop of inductors connected in series, or as a branch of capacitors, connected in parallel. (b) Each one-port is essentially a time-varying IIR filter with coefficients given as in Eq. (12).

KS structure, the same energetic properties hold also for the physically valid dual delay line model, although the former is considered here for simplicity. Also, since the spatial resolution of the string is preserved (the number of one-ports is fixed), changing the a parameter is analogous to varying the string tension, rather than the string length, as it was in Sec. 2.

3.1. Experiments

Experiments reveal that the time-varying WDF structure simulates well the varying tension string. For the simulation, we used a lossless WDF string model consisting of 126 one-ports. The fundamental frequency of the string was again rapidly slid from 345 Hz to 690 Hz and back, and the resulting output amplitude and normalized string energy behavior are shown in Fig. 7. The delay lines were again initialized with a DC signal for illustrational purposes, but similar results can be obtained even with impulsive excitations. The string energy was evaluated as in Eq. (3), except that this time, due to the change in tension, the wave impedance was varied, instead of the integration limits. The wave impedance was calculated as [8]

$$Z_0 = \sqrt{T\epsilon} = 2\epsilon L f_0, \quad (14)$$

where T is the tension, ϵ is the linear mass density, L is the length, and f_0 is the fundamental frequency of the string. As can be seen in Fig. 7(b), the energy of the WDF string remains constant.

It is interesting to note, that since the wave digital one-port is a time-varying generalization of the simple first-order allpass filter, the phase delay remains unaltered at high frequencies [5], thus leading to inharmonicity near the Nyquist limit. On the other hand, the same effect also prevents the aliasing problem, discussed in Sec. 2.2.

4. DISCUSSION AND CONCLUSIONS

In this paper we have studied two ways to realize energy-preserving fractional delay filters for physics-based modeling applications. In the first one the signal interpolated from a sliding termination point is compensated by a correction factor. In the second one, time-varying wave digital one-ports are used in a varying-tension string

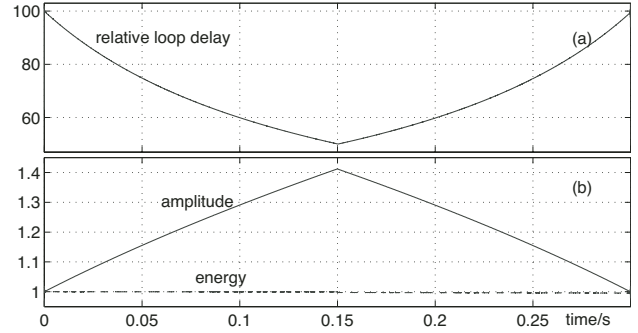


Fig. 7. Evaluation of the WDF string energy. The relative delay time (in %) of the WDF loop was varied as shown in (a). The resulting variation in the output amplitude and string energy are shown in (b) with solid and dashed lines, respectively.

loop to preserve energy. Both methods are found useful in string instrument modeling, such as in synthesizing the slide guitar.

5. ACKNOWLEDGMENTS

This work is supported by the ALMA project (IST-2001-33059), the Academy of Finland projects SA 104934 and SA 53537, and the Helsinki Graduate School of El. & Comm. Eng.

6. REFERENCES

- [1] J. O. Smith, *Principles of Digital Waveguide Models of Musical Instruments*, Chapter 10 in *Applications of Digital Signal Processing to Audio and Acoustics*, (ed. Kahrs and Brandenburg), Kluwer Academic Publishers, 1998.
- [2] K. Karplus and A. Strong, "Digital Synthesis of Plucked-String and Drum Timbres," *Computer Music J.*, vol. 7, no. 2, pp. 43–55, 1983.
- [3] Jaffe and J. O. Smith, 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.
- [4] M. Karjalainen, V. Välimäki, and T. Tolonen, "Plucked String Models: from Karplus-Strong Algorithm to Digital Waveguides and Beyond," *Computer Music J.*, vol. 22, no. 3, pp. 17–32, 1998.
- [5] T. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the Unit Delay," *IEEE Signal Processing Magazine*, vol. 13, no. 1, pp. 30–60, 1996.
- [6] V. Välimäki, *Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters*. Doctoral thesis. Report no. 37, Helsinki Univ. of Tech., Lab. of Acoustics and Audio Signal Processing, Espoo, Finland, 1995. Available online at: <http://www.acoustics.hut.fi/~vpv/publications/vesa.phd.html>
- [7] V. Välimäki, T. I. Laakso, and J. Mackenzie, "Elimination of Transients in Time-Varying Allpass Fractional Delay Filters with Application to Digital Waveguide Modeling," in *Proc. Int. Computer Music Conf.*, pp. 327–334, Banff, Alberta, Canada, Sept. 1995.
- [8] J. O. Smith, *Physical Audio Signal Processing*, August 2004 Draft <http://ccrma.stanford.edu/~jos/pasp/>.
- [9] S. Bilbao, "Time-varying Generalizations of Allpass Filters", unpublished manuscript, 2004.
- [10] A. Fettweis, "Wave Digital Filters: Theory and Practice," *Proc. of the IEEE*, vol. 74, no. 2, pp. 270–327, 1986.