IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'00), Istanbul, Turkey, 5–9 June 2000

PRINCIPLES OF FRACTIONAL DELAY FILTERS

V. Välimäki[1] *and T. I. Laakso*[2]

[1]Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing
P. O. Box 3000, FIN-02015 HUT, Espoo, Finland
vesa.valimaki@hut.fi, http://www.acoustics.hut.fi

[2]Helsinki University of Technology, Laboratory of Signal Processing and Computer Technology
P.O. Box 3000, FIN-02015 HUT, Espoo, Finland
timo.laakso@hut.fi, http://wooster.hut.fi

ABSTRACT

In numerous applications, such as communications, audio and music technology, speech coding and synthesis, antenna and transducer arrays, and time delay estimation, not only the sampling frequency but the actual sampling instants are of crucial importance. Digital fractional delay (FD) filters provide a useful building block that can be used for fine-tuning the sampling instants, i.e., implement the required bandlimited interpolation. In this paper an overview of design techniques and applications is given.

1. INTRODUCTION

The sampling rate must satisfy the Nyquist criterion in order for a sample set to represent adequately the original continuous signal. This problem has been addressed in the sampling theory literature, e.g., [1], [2]. However, the appropriate sampling rate alone is not sufficient for many applications—also the *sampling instants* must be properly selected. For example, in digital communications, the decisions of the received bit or symbol value are made based on samples of the received continuous-time pulse sequence which should be taken exactly at the middle of each pulse to minimize probability of erroneous decision. This requires that both the sampling frequency and the sampling instants must be synchronized to the incoming signal.

Another class of problems is modeling of musical instruments which involves discretization of differential equations describing a physical system producing acoustical vibrations. Propagation delays caused by a finite speed of vibration in strings, tubes, and other musical resonators must be simulated accurately—otherwise the instrument will sound out of tune. Such delays are not generally multiples of the sampling interval used. The actual playing of the instrument requires changing of system parameters on-line and thus special care must be taken to avoid transient phenomena due to these changes.

Both examples are typical applications of fractional delay (FD) filters, i.e., situations where uniform sampling is used and interpolation between samples is required [3], [4], [5][†]. Fractional delay means, assuming uniform sampling, a delay that is a non-integer multiple of the sample interval. Employing fractional delay filters facilitates the use of traditional well-known methods developed for uniformly sampled signals and yet the observation of signal values at arbitrary locations between the samples.

In this paper, we review the principles of fractional delay filters. First we introduce the digital fractional delay problem and discuss and compare the known techniques for designing nonrecursive (FIR) and recursive (IIR, especially allpass) filters approximating a given FD value. A special case of approximating a very small delay ($D < 1$) is also considered. The implementation of time-varying FD filters and transient problems in time-varying recursive FD filters are briefly discussed. Some applications are reviewed.

2. IDEAL FRACTIONAL DELAY AND ITS APPROXIMATIONS

The ideal fractional delay element is a digital version of a continuous-time delay line. The delay system must be rendered bandlimited using an ideal lowpass filter while the delay merely shifts the impulse response in the time domain. Thus, the impulse response of an ideal fractional delay filter is a shifted and sam-

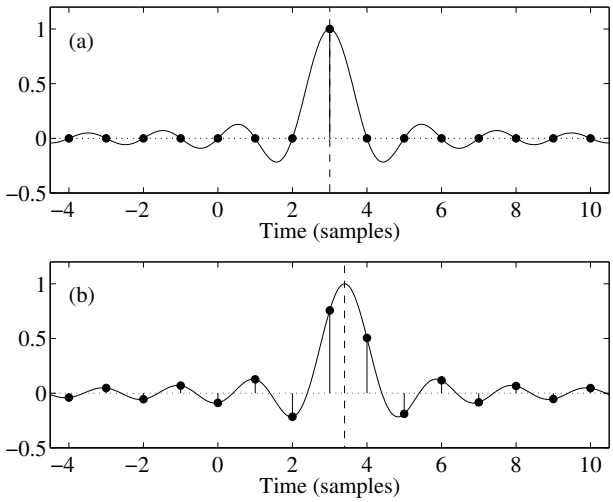

Fig. 1. Continuous-time (solid line) and sampled (dots) impulse response of the ideal fractional delay filter, when the delay is (a) $D = 3.0$ samples and (b) $D = 3.4$ samples. The vertical dashed line indicates the midpoint of the continuous-time impulse response in each case.

[†] Due to space limitations, we cannot refer to all relevant publications here. Many more references can be found in [5].

pled sinc function, that is, $h(n) = \mathrm{sinc}(n - D)$, where n is the (integer) sample index and D is the delay with an integral part floor(D) and a fractional part $d = D - \mathrm{floor}(D)$. The floor function returns the greatest integer less than or equal to D. Figure 1 shows the ideal impulse response when $d = 0.0$ and $d = 0.4$ samples. In the latter case, the impulse response is infinite-length. For this reason, the impulse response corresponds to a *noncausal* filter which cannot be made causal by a finite shift in time. In addition, the filter is not stable since the impulse response is not absolutely summable. The ideal FD filter is thus *nonrealizable*. To produce a realizable fractional delay filter, some finite-length, causal approximation for the sinc function has to be used.

2.1 FIR Approximation of Fractional Delay

We consider five different approaches to design causal FD FIR filters [4], [5]:

1. Windowed sinc function (using an asymmetric window function with a fractional offset) [6], [7], [8];
2. Lowpass FD approximation with a smooth transition band obtained using a low-order spline function;
3. Maximally-flat FIR approximation (Lagrange interpolation);
4. Weighted least-squares (WLS) approximation;
5. Oetken's method (a quasi-equiripple FD approximation).

None of the above design methods requires an iterative algorithm. However, Oetken's method utilizes an odd-length equiripple linear-phase FIR filter, which can be designed using the Remez algorithm; it then obtains the almost equiripple FD approximation with a matrix operation. A limitation of Oetken's method is that it is only suitable for odd-order FIR FD filters, which have a transmission zero at the Nyquist limit when $d = 0.5$. All the other methods are available for both odd and even orders.

Lagrange interpolation is obtained as a maximally-flat approximation of FD. It may be derived in many different ways [9], [10]. The coefficients of the Lagrange interpolator are given by the following equation:

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

where D is the desired total delay and N is the order of the filter. Kootsookos and Williams [11] showed that Eq. (1) (for even N) can also be obtained using the *windowing method*: the window function must be a scaled binomial window. This result has been generalized for odd-order Lagrange interpolators in [12]. The following equation defines the windowing method for the design of a Lagrange interpolator for both odd and even N:

$$h(n) = C_{\mathrm{bin}}(D, N)\, w_{\mathrm{bin}}(n)\, \mathrm{sinc}(n - D) \qquad (2a)$$

where $n = 0, 1, 2, ..., N$, and the scaling coefficient $C_{\mathrm{bin}}(D, N)$ and the binomial window $w_{\mathrm{bin}}(n)$ are defined as

$$C_{\mathrm{bin}}(D, N) = (-1)^N \frac{\pi(N+1)}{\sin(\pi D)} \binom{D}{N+1} \quad \text{and} \quad w_{\mathrm{bin}}(n) = \binom{N}{n} \qquad (2b,c)$$

2.2 IIR Approximation of Fractional Delay

Allpass filters are particularly well suited to FD approximation since their magnitude response is exactly unity at all frequencies. The literature on IIR FD filters mainly concentrates on allpass filters, with notable exceptions, e.g., [13]. The following five allpass filter design methods were discussed in our review articles [4], [5]:

1. Least squares (LS) phase approximation;
2. LS phase delay approximation;
3. Maximally-flat group delay approx. (Thiran allpass filter);
4. Iterative WLS phase error design (enables almost equiripple phase approximation);
5. Iterative WLS phase delay error design (enables almost equiripple phase-delay approximation).

The design of allpass FD filters typically requires an iterative design algorithm or solving a set of linear equations. The simplest allpass FD filter design is the Thiran allpass filter that has closed-form coefficient formulas ($k = 0, 1, 2, ..., N$) [14], [15]:

$$a_k = (-1)^k \binom{N}{k} \prod_{n=0}^{N} \frac{N-n-D}{N-k-n-D} \qquad (3)$$

2.3 Comparison of FIR and Allpass Designs

We compare the above FD filters in terms of their frequency response error (FRE) magnitude. We have chosen to present results for a wideband specification where the passband of approximation is 80% of the Nyquist limit (i.e., normalized frequency 0.4). The length of the FIR approximations is 10 (i.e., $N = 9$), and the order of the allpass filters is 5; an equal number of multiplications is needed to implement these filters.

The Chebyshev window function with a 35-dB ripple level was selected for sinc windowing. A second-order spline function was used in lowpass FD approximation. In the windowed sinc and lowpass FIR FD filter design, the coefficients must be scaled to obtain the best approximation (e.g., so that their frequency response at $\omega = 0$ equals unity). The allpass filters are automatically scaled by the design algorithms.

The useful range of delay D is different for FIR and allpass filters. In the case of FIR filters, the best approximation is obtained when interpolating between the middle taps in the case of even-length filters and within half a sample from the middle tap in the case of odd-length ones. The squared approximation error functions (as a function of D) are symmetric about the midpoint of the FIR filter. Here we consider the worst case, which occurs with $d = 0.5$ when the filter length is even.

In the case of allpass filters, the error curves are asymmetric [12]. Furthermore, the stability of the allpass filters must be taken into account. For example, the Thiran interpolator is stable only for $D > N - 1$, which implies that there is a lower limit for the delay to be approximated [15]. In general, the optimal range of D (width of one sampling interval) for allpass filters is about $N - 0.5$ to $N + 0.5$ [12]. In our example $N = 5$ and the desired delay is $D = 4.5$.

Figure 2 presents the FRE magnitude for the five FD FIR

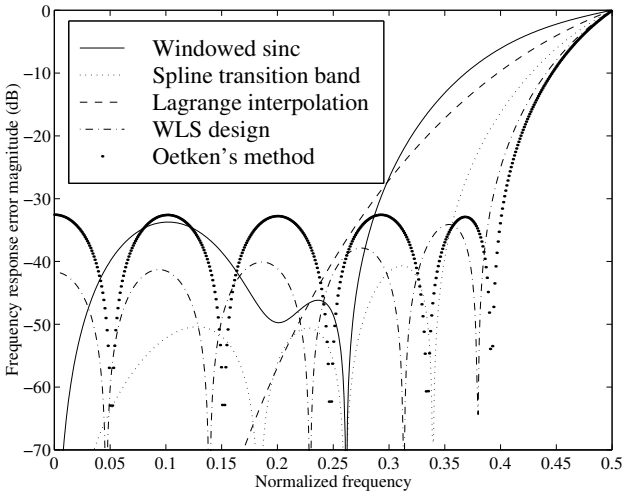

Fig. 2. Frequency response error magnitude of five 10-tap FIR fractional delay filters ($D = 4.5$, $\omega_p = 0.8\pi$).

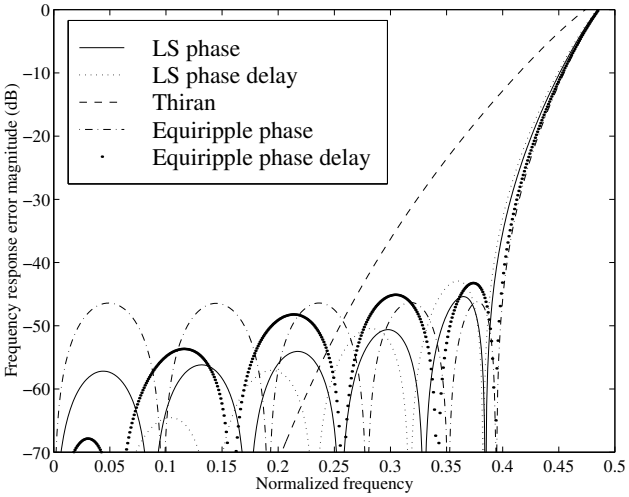

Fig. 3. Frequency response error magnitude of five 5th-order allpass fractional delay filters ($D = 4.5$, $\omega_p = 0.8\pi$).

approximations. In terms of peak error, the best result (–32.5 dB) is achieved using Oetken's method, which gives a quasi-equiripple error on the approximation band. The WLS method is not much worse but its FRE increases towards the upper edge of the approximation band reaching the value of –25.1 dB. The WLS approximation is optimal in the sense that the error power is smallest possible in the approximation band—sometimes this may be more important than the peak error. Lagrange interpolation is not competent in this kind of comparison, since it merely minimizes the error in the vicinity of $\omega = 0$. Nevertheless, for narrow-band cases it is a useful technique.

Figure 3 displays the FRE magnitude of the five allpass filters. In general, they appear to be better than the FIR approximations of Fig. 2. The best performance (the smallest peak FRE in the approximation band) is obtained with the equiripple phase and phase delay allpass filter approximations (–45.8 dB and –42.0 dB, respectively). The peak errors of the LS phase and phase delay approximations are not much worse (–35.3 dB and –32.9 dB, respectively). The Thiran allpass filter gives a clearly poorer performance in terms of peak FRE. It is a maximally flat group delay approximation at $\omega = 0$ and is thus comparable to Lagrange interpolation. The Thiran allpass filter is easy to design and is attractive for a narrow-band approximation.

MATLAB functions that can be used for designing filters in this comparison are currently available at http://www.acoustics.hut.fi/software/fdtools/.

3. VERY SMALL DELAYS

FD filters yield the best approximation when the total delay D is close to $N/2$ for FIR filters and close to N for allpass filters. It is, however, of interest to consider the case of a very small D with respect to N. Figure 4 shows how the approximation error behaves as the order of a Lagrange interpolation filter is increased from $N = 1$ (i.e., linear interpolation) to $N = 9$ but the required delay is kept constant ($D = 0.5$ samples), i.e., the integer part of the delay is not allowed to increase with filter order. It is seen that the error decreases only at very low frequencies whereas at high frequencies the error gets larger.

All the five allpass filter designs discussed above yield an unstable filter when the desired delay is smaller than $N - 1$ (e.g., $D < 4$ when $N = 5$). Thus, delays between 0 and 1 samples can be best realized with a first-order allpass filter (see Fig. 4), which yields a coarse approximation but which is anyway stable.

It is, nevertheless, possible to design recursive FD filters which have an all-pole response. One possibility is the Thiran all-pole filter [14], which has been used to devise the maximally-flat allpass FD filter. With this filter we notice that the error does not decrease much even when the order is increased. Examples of the first and 10th-order Thiran allpass filters are presented in Fig. 4. The Thiran all-pole design fails because the approximation concentrates on the group delay, and the magnitude response

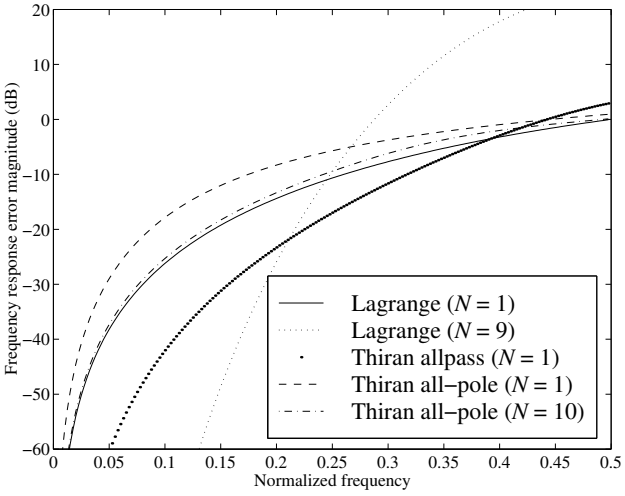

Fig. 4. Frequency response error magnitude of linear interpolation, 9th-order Lagrange interpolation, first-order Thiran allpass filter, and 1st and 10th-order Thiran all-pole filters for approximating a small delay ($D = 0.5$).

remains narrow lowpass even when increasing the filter order.

As a conclusion, the design of high-quality FD filters is difficult if a very small delay is required. One possibility is to use a higher sampling rate if accurate and small fractional delays are needed. This helps the approximation task in two ways: the unit delays are scaled down by the new sampling interval, and the normalized approximation bandwidth is reduced.

4. TIME-VARYING FD FILTERS

Many applications of fractional delay filters require the delay parameter to vary over time (see Section 5 below). Farrow [16] suggested that every filter coefficient of an FIR FD filter could be expressed as an Nth-order polynomial in the variable delay parameter D. This results in a structure containing $N + 1$ FIR filters $C_k(z)$ with constant coefficients. An implementation of this structure using Horner's rule is given in Fig. 5. During the last decade, the Farrow structure has become the most popular method for implementing time-varying FIR FD filters (e.g., [17], [12], [18]). Vesma and Saramäki have proposed a modified Farrow structure which is a polynomial of $2D - 1$ [18]. The advantage of their structure is that the fixed subfilters are linear-phase FIR filters. In the case of time-varying recursive filters, transients may become a problem. They are caused by the feedback from the state variables of the filter, which have been computed using previous coefficient values. A method for suppressing the transients is discussed in [19].

5. APPLICATIONS

Fractional delay filters are useful in numerous signal processing tasks. A basic example is sampling-rate conversion for incommensurate ratios, such as between 44.1 kHz and 48 kHz, a problem that occurs often in digital audio. The FD filter can be used to compute output signal samples with a different delay value each time. Obviously, the FD filter must be time-varying in this case. In practice, the situation is often made even harder by the fact that the sampling-rate ratios are not only incommensurate but also time-varying, which is caused, e.g., by variations in clock frequencies due to temperature, aging, or external disturbances. Various other examples include synchronization of digital modems, digital simulation of the Doppler effect in virtual reality systems, or elimination of wow in old gramophone disc recordings. For more information and references on these and other applications of FD filters, see [5].

6. ACKNOWLEDGMENTS

The work of the first author has been supported by the Academy of Finland. The authors are grateful to Dr. Tony Verma for his helpful comments.

7. REFERENCES

[1] A. J. Jerri, "The Shannon sampling theorem—its various extensions and applications: a tutorial review," *Proc. IEEE*, vol. 65, no. 11, pp. 1565–1596, Nov. 1977.

[2] R. J. Marks II, *Introduction to Shannon Sampling and Interpolation Theory*. New York: Springer-Verlag, 1991.

[3] G. D. Cain, N. P. Murphy, and A. Tarczynski, "Evaluation of several FIR fractional-sample delay filters," in *Proc. IEEE ICASSP '94*, Adelaide, South Australia, Apr. 1994, vol. 3, pp. 621–624.

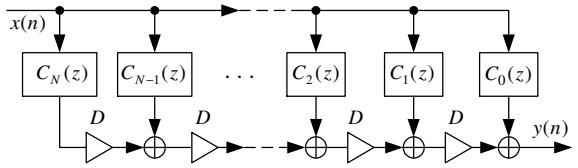

Fig. 5. Farrow structure for time-varying FD filtering.

[4] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay—tools for fractional delay filter design," *IEEE Signal Process. Mag.*, vol. 13, no. 1, pp. 30–60, Jan. 1996.

[5] V. Välimäki and T. I. Laakso, "Fractional delay filters—design and applications," in *Theory and Applications of Nonuniform Sampling*, F. Marvasti, ed. New York: Plenum/Kluwer, to be published in 2000.

[6] G. D. Cain, A. Yardim, and P. Henry, "Offset windowing for FIR fractional-sample delay," in *Proc. IEEE ICASSP'95*, Detroit, MI, May 1995, vol. 2, pp. 1276–1279.

[7] T. I. Laakso, T. Saramäki, and G. D. Cain, "Asymmetric Dolph–Chebyshev, Saramäki, and transitional windows for fractional delay FIR filter design," in *Proc. 38th Midwest Symp. Circ. Syst.*, Rio de Janeiro, Brazil, Aug. 1995, vol. 1, pp. 580–583.

[8] A. Yardim, G. D. Cain, and P. Henry, "Optimal two-term offset windowing for fractional delay," *Electron. Lett.*, vol. 32, no. 6, pp. 526–527, Mar. 1996.

[9] E. Hermanowicz, "Explicit formulas for weighting coefficients of maximally flat tunable FIR delayers," *Electron. Lett.*, vol. 28, no. 20, pp. 1936–1937, Sept. 1992.

[10] S. Minocha, S. C. Dutta Roy, and B. Kumar, "A note on the FIR approximation of a fractional sample delay," *Int. J. Circ. Theor. Appl.*, vol. 21, no. 3, pp. 265–274, May–June 1993.

[11] P. J. Kootsookos and R. C. Williamson, "FIR approximation of fractional sample delay systems," *IEEE Trans. Circ. Syst.—Part II*, vol. 43, no. 3, pp. 269–271, Mar. 1996.

[12] V. Välimäki, *Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters*. Doctoral dissertation. Report no. 37, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland, Dec. 1995. Available at http://www.acoustics.hut.fi/~vpv/publications/vesa_phd.html.

[13] A. Tarczynski and G. D. Cain, "Design of IIR fractional-sample delay filters," in *Proc. 2nd Int. Symp. DSP for Commun. Syst.*, Adelaide, South Australia, Apr. 1994.

[14] J.–P. Thiran, "Recursive digital filters with maximally flat group delay," *IEEE Trans. Circ. Theory*, vol. 18, no. 6, pp. 659–664, Nov. 1971.

[15] A. Fettweis, "A simple design of maximally flat delay digital filters," *IEEE Trans. Audio and Electroacoust.*, vol. 20, no. 2, pp. 112–114, June 1972.

[16] C. W. Farrow, "A continuously variable digital delay element," in *Proc. IEEE Int. Symp. Circ. Syst.*, Espoo, Finland, June 1988, vol. 3, pp. 2641–2645.

[17] L. Erup, F. M. Gardner, and R. A. Harris, "Interpolation in digital modems—part II: implementation and performance," *IEEE Trans. Commun.*, vol. 41, no. 6, pp. 998–1008, June 1993.

[18] J. Vesma and T. Saramäki, "Interpolation filters with arbitrary frequency response for all-digital receivers," in *Proc. Int. Symp. Circ. Syst.*, Atlanta, GA, May 1996, vol. 2, pp. 568–571.

[19] V. Välimäki and T. I. Laakso, "Suppression of transients in variable recursive digital filters with a novel and efficient cancellation method," *IEEE Trans. Signal Process.*, vol. 46, no. 12, pp. 3408–3414, Dec. 1998.