Lecture 8
Introduction to Multirate

Topics for Today

- Upsampling and Downsampling
- Multirate Identities
- Polyphase
- Decimation and Interpolation
- Fractional Delay
- Sampling Rate Conversion
- Multirate Analysis of STFT Filterbank

Main References (please see website for full citations).

- Vaidyanathan, Ch.4,11
- Vetterli Ch. 3
- Laakso, et al. "Splitting the Unit Delay" (see 421 citations)
- `http://www-ccrma.stanford.edu/~jos/resample/`

Z transform

For this lecture, we do frequency-domain analysis more generally using the *z-transform*:

$$X(z) \triangleq \sum_{n=-\infty}^{\infty} x[n]z^{-n}; z \in \mathcal{ROC}(\mathbf{x})$$

To obtain the DTFT, evaluate at $z = e^{j\omega}$.

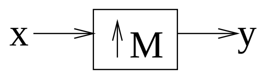
Signal Notation

We refer interchangeably to "$x[n]$" as the signal x evaluated at time n, *or* as the entire signal vector. Please excuse this lapse in rigor, as otherwise simple operations such as modulation (e.g. $w[n] = x[n]y[n]$) become cumbersome.

Upsampling and Downsampling

Upsampling

- Notation:



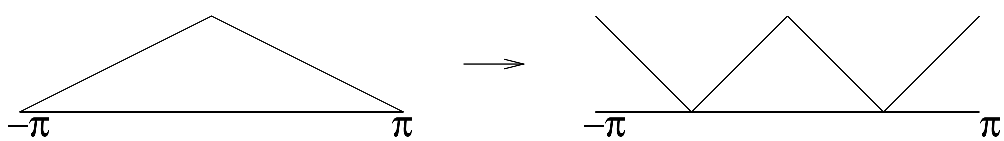
- Basic Idea: To upsample by integer M, stuff $M-1$ zeros in between the samples of $x[n]$.
- Time Domain

$$y[n] = \begin{cases} x[n/M], & M \text{ divides } N \\ 0 & \text{otherwise.} \end{cases}$$

- Frequency Domain

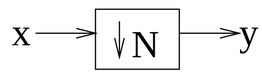
$$Y(z) = X(z^M)$$

- Plugging in $z = e^{j\omega}$, we see that the spectrum in $[-\pi, \pi)$ contracts by factor M and M *images* placed around the unit circle:



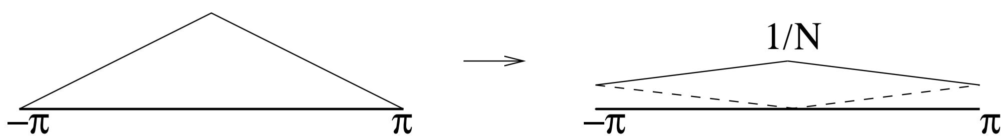
Downsampling

- Notation:



- Basic Idea: Take every N^{th} sample.
- Time Domain: $y[n] = x[Nn]$
- Frequency Domain:

$$Y(z) = \frac{1}{N} \sum_{k=0}^{N-1} X(e^{j2\pi k/N} z^{1/N})$$



Proof: Upsampling establishes a one-to-one correspondence between

- $\mathcal{S}_1$, the space of all discrete-time signals, and
- $\mathcal{S}_N$, space of signals nonzero only at integer multiples kN.

Consequently, if $w[n] \in \mathcal{S}_N$, and $x[n]$ is $w[n]$ downsampled by M, then $X(z) = W(z^{1/N})$. To complete the proof, we need only obtain $W(z)$ where $w[n] = \delta^N[n]x[n]$, $\delta^N[n]$ being the discrete unit impulse train with spacing N. By the modulation theorem:

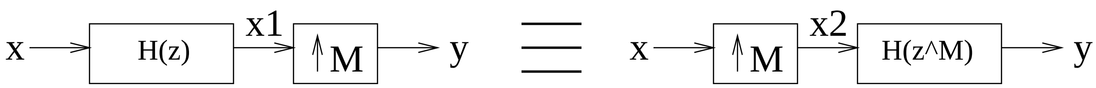
$$W(z) = \frac{1}{N} \sum_{k=0}^{N-1} X(e^{j2\pi k/N} z)$$

so replace $z \to z^{1/N}$ and we are done.

Remark: Consider downsampling as a two stage process: first, we discard information which leads to aliasing in the frequency domain; next we contract the signal in time, stretching out the spectrum.

Multirate Identities

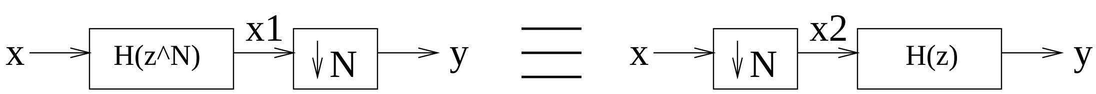
Interchange of Upsampling and Filtering



Proof:

$Y(z) = X_1(z^M) = H(z^M)X(z^M)$. Since $X_2(z) = X(z^M)$, filter with $H(z^M)$ to get $Y(z)$.

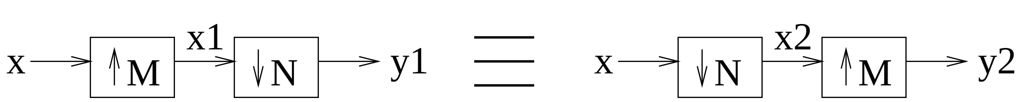
Interchange of Filtering and Downsampling



Proof:

$$
\begin{aligned}
X_1(z) &= H(z^N)X(z) \\
Y(z) &= \frac{1}{N}\sum_{k=0}^{N-1} H((e^{j2\pi k/N}z^{1/N})^N)X(e^{j2\pi k/N}z^{1/N}) \\
&= \frac{1}{N}\sum_{k=0}^{N-1} H(z)X(e^{j2\pi k/N}z^{1/N}) \\
&= H(z)\frac{1}{N}\sum_{k=0}^{N-1} X(e^{j2\pi k/N}z^{1/N}) \\
&= H(z)X_2(z).
\end{aligned}
$$

Interchange of Upsampling and Downsampling



iff M, N are relatively prime.

Proof: Here we give a time-domain proof. Good proofs of the standard frequency-domain version can be found in Vetterli 2.8 or Vaidyanathan 4.2 (pay particular attention to Vaidyanathan Ex. 4.12, "jumping painters").

Proceeding, we have:

$$\begin{aligned} x_1[n] &= x[n/M]\, 1_{\{M \text{ div. } n\}} \\ y_1[n] &= x[Nn/M]\, 1_{\{M \text{ div. } n\}} \\ x_2[n] &= x[Nn] \\ y_2[n] &= x[Nn/M]\, 1_{\{M \text{ div. } Nn\}} \end{aligned}$$

Hence, it is sufficient to show [(M div. n iff M div. Nn) iff M, N relatively prime]. Always if M div. n, M div. Nn, so it's left to check conditions for the converse. If M, N are relatively prime then M div. $Nn \rightarrow M$ div. n. Conversely, let $J/K = M/N$ in lowest terms and define $L = M/J$. Let $P/Q = K/L$ in lowest terms. Now let $n = JP$. Observe that M div. $N \cdot n = M \cdot K \cdot P$ but $n/M = J \cdot P/M = P/L$ which is in lowest terms. The only way P/L can be in lowest terms and an integer is for $L = 1$, which implies M, N are relatively prime, so we are done.

Other properties: such as linearity, commuting upsamplers and downsamplers through multipliers and summing junctions, etc. should be obvious (see beginning of Vaidyanathan 4.2).

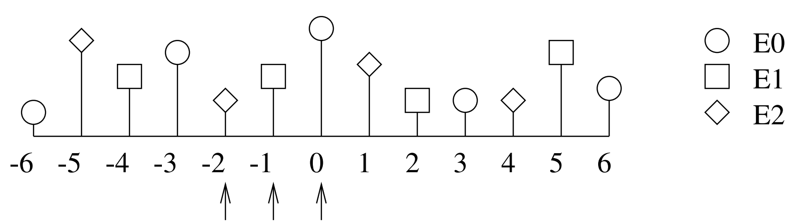
Polyphase

- Remember that if $x[n] \in \mathcal{S}_N$, it is trivial to obtain the spectrum of $x[Nn]$ as $X(z^{1/N})$. Also if $x[n]$ is a filter, it is straightforward to apply the appropriate interchange identities. Can we decompose any signal as a linear combination of a finite number of shifted versions of signals in $\mathcal{S}_N$? This is the subject of *polyphase decomposition*.

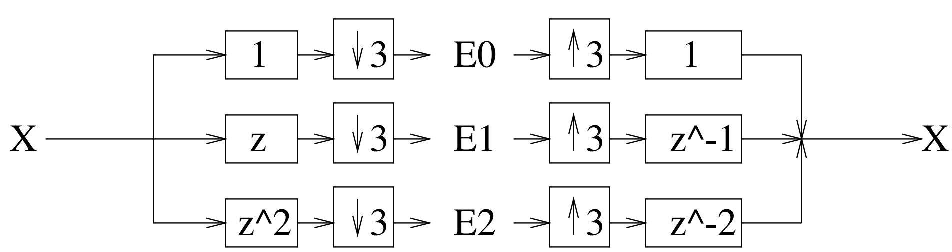
- Frequency Domain:

$$X(z) = \sum_{k=0}^{M-1} z^{-k} E_k(z^M)$$

- Time Domain: The basic idea is to decompose $x[n]$ into its periodically interleaved subsequences.



- System Diagram (M=3)



- **Remark:** Our first PR filterbank! Aliasing in the subband signals (here: the polyphase components) cancels upon reconstruction. Unfortunately, we don't get any frequency localization here. We'll see much better filterbanks at the end of this lecture.

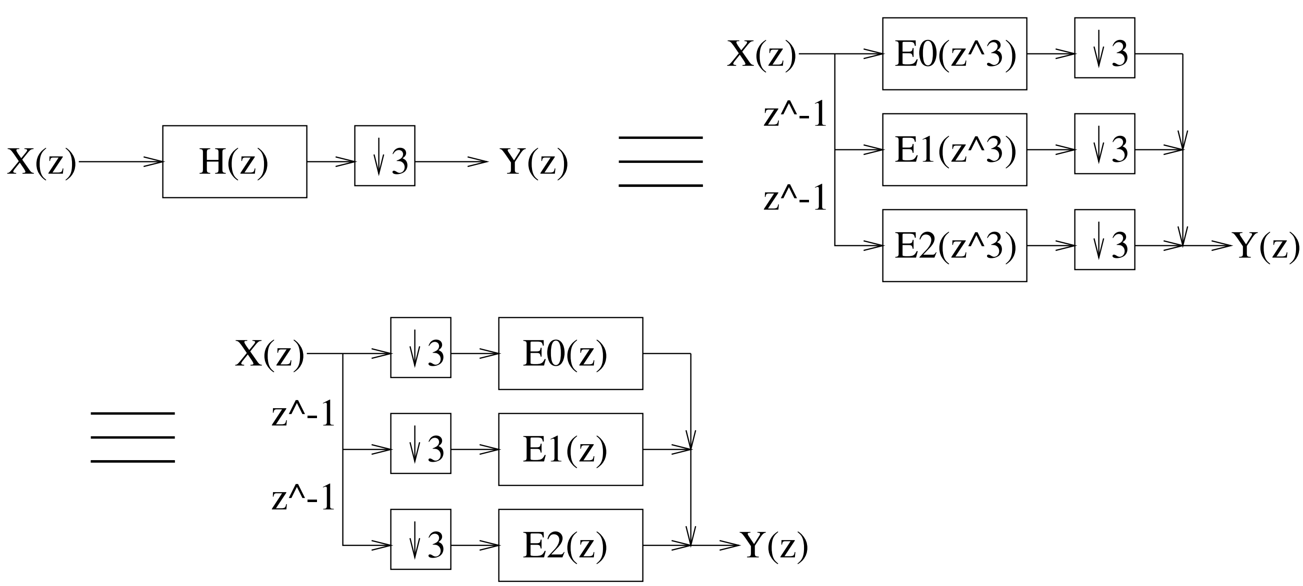
- The reverse polyphase decomposition is also defined:

$$X(z) = \sum_{k=0}^{M-1} z^k R_k(z^M).$$

Decimation and Interpolation

Decimation Filtering

- Consider filtering followed by downsampling by N. We throw away $N-1$ out of N samples: intuition tells us there should be some way to reduce computations by a factor of N.
- **Thought Exercise: Blocking and Averaging** Consider filtering by $1/N$ times length-N rectangular window, then downsampling by N. Convince yourself this is the same as taking length-N blocks and outputting the average of each block.
- With polyphase decomposition on $H(z)$ and some identities, we extend the above idea to a general filter $H(z)$:



- Recognizing that the input structure just does a *reverse* polyphase decomposition on the input, the polyphase decimation filter is written:

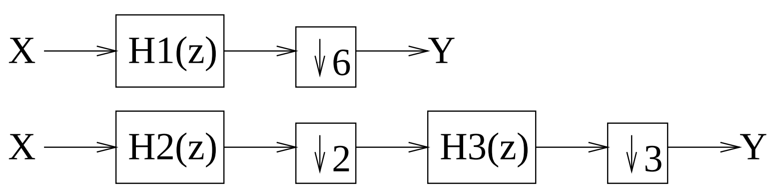
$$Y(z) = \sum_{k=0}^{N-1} R_{(X,k)}(z) E_{(H,k)}(z)$$

- Complexity Analysis: We compute as many filter taps, but per N samples. Hence, complexity is reduced by factor N.

- **Multistage Decimation**

  - If decimation by a large factor is required, it's often better to break it up into stages. This is because we can allow aliasing in the intermediate stages, insofar as it is rejected by the antialias filter for the final stage.

  - For example:



  $H_1(z)$ must have cutoff $\pi/6$, with minimal transition width. However $H_2(z)$ is allowed transition width from $\pi/2$ to $5\pi/6$, because the lowest foldover frequency after downsampling by 2 is $\pi/3$, which exceeds the cutoff for $H_3(z)$.

  - Tradeoff: In the passband, we have a cascade of filters, so there may be significant propagation of ripple (Please review HW#2, Problem 9).

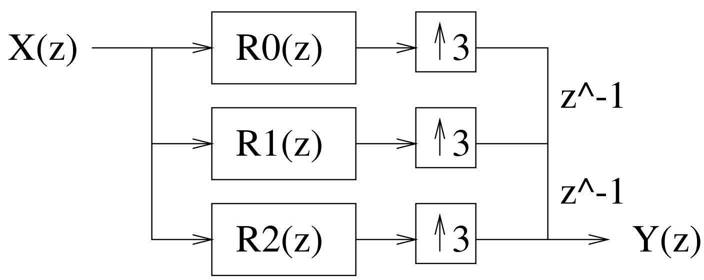
  - See (Vaidyanathan 4.4.2) for exact computational analysis.

Interpolation Filtering

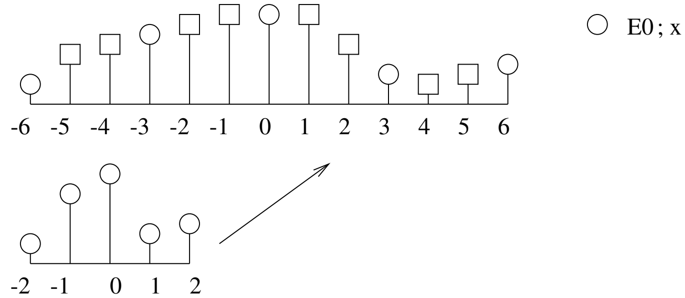
- Consider upsampling by M and filtering with $H(z)$. Intuition tells us we can reduce computations by factor of M, because $M-1$ out of M inputs to $H(z)$ are zero.
- The corresponding polyphase decomposition is:

$$E_{(Y,k)}(z) = R_{(H,k)}(z)X(z); \quad k = 0 \ldots M-1.$$

- Picture:



- **Exercise:** Show this directly, imitating the polyphase decomposition for the decimation filter. Later, we discuss (Hermitian) *transpose structures*: afterwards, you'll see no work at all is required in the derivation.
- Consider $H(z)$ satisfying the *Nyquist(M)* property: e.g. the zeroth polyphase component is an impulse. In this case,

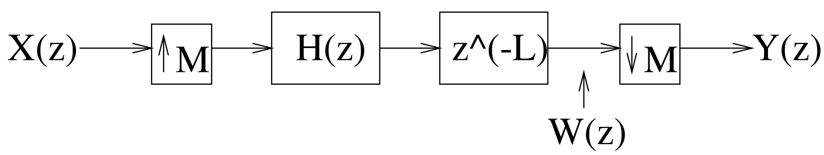


polyphase decomposition obtains $y[kM] = x[k]$. Evidently, the interpolated values "go through" the original data points.

Examples include linear, sinc and Lagrange interpolation: we'll return to these in 421, as well as discuss IIR structures. You can now read the Laakso paper "Splitting the Unit Delay" (see 421 citations).

Fractional Delay

- We wish to delay $x[n]$ by an amount D which is not an integer number of samples. Formally, let's design a filter for which magnitude response is constant (allpass), but phase response is -$\omega D/2$. An important application is delay line interpolation; this to allow, for instance, continuous variation in the pitch of an instrument model.
- Consider first the case of D rational. Then there exists L, M such that $D = L/M$, and we apply the following scheme:

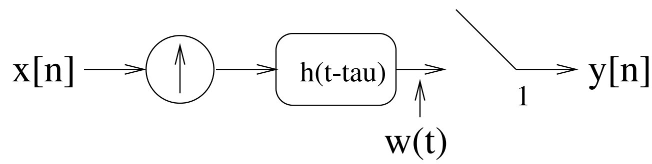


- Ideally, $H(z)$ should be zero-phase, bandlimited to π/M, with flat response in the passband. Only the sinc function satisfies all three criteria. In practice, we must be content with shorter filters.
- The fractional delay system is LTI, and equivalent to $R_{(H,L)}(z)$.
- $\mathbf{Proof:}$ Let $G(z) = z^{-L}H(z)$. Downsampling by M extracts the zeroth polyphase component of $W(z)$, which by polyphase decomposition for interpolation filtering equals $R_{(G,0)}(z)X(z)$, and $R_{(G,0)}(z) = R_{(H,L)}(z)$.
- $\mathbf{Remark:}$ Since $H(z)$ is arbitrary, $W(z)$ is usually *not* bandlimited. As $W(z)$ contains the filtered images $X(z^M)$, one might expect alias components of $X(z)$ to remain in the output. As the proof shows, there are in fact no alias components. Here

we have another example of *aliasing cancellation*, which figures prominently in our subsequent study of filterbanks. (There is, however, aliasing of $z^{-}LH(z)$).

"Canceled" alias components reappear when the delay amount becomes time-variant. Intuitively, think of this as Doppler shift: the effective "downsampling rate" is no longer exactly M. A *first order* solution may be found in fixed sampling-rate conversion techniques. In any event, this alone is an argument for $H(z)$ to resemble an ideal lowpass (sinc) filter.

- **Exercise:** Give an alternative proof using only the frequency-domain properties of upsampling and downsampling.
- **Exercise:** Suppose that we downsample by N instead of M. Determine necessary and sufficient conditions on the pair $\{M, N\}$ such that the output is free of alias components of $X(z)$ uniformly for all L and $H(z)$.
- **Exercise:** Why is it a good idea for $H(z)$ to be Nyquist(M)?
- Fix as constant the number of taps in each polyphase filter $R_{(H,L)}(z)$. Since computational load is now invariant to M, we can get any desired resolution for the delay amount $\tau = L/M$ provided M gets arbitrarily large. We must store $\mathcal{O}(M)$ coefficients for the polyphase filters, but this cost is obviated assuming a "closed-form" expression exists for this family of filters, indexed by the *continuous* variable τ.
- To handle the continuous case, we define the τ-*polyphase decomposition* of $h(t)$ as $e_{h,\tau}[n] = h(n+\tau)$; $r_{h,\tau}[n] = h(n-\tau)$. To corroborate prior results, consider the diagram:



(The symbol $(\uparrow)$ denotes "conversion to impulse train"; that is, $x[n] \to (\uparrow)$ obtains $\Sigma_{k=-\infty}^{\infty} x[k]\delta(t-k)$).
Proceeding:

$$\begin{aligned} w(t) &= \left[\sum_{k=-\infty}^{\infty} x[k]\delta(t-k)\right] *_t h(t-\tau) \\ &= \sum_{k=-\infty}^{\infty} x[k]h(t-(\tau+k)) \\ y[n] &= w[n] \\ &= \sum_{k=-\infty}^{\infty} x[k]h(n-(\tau+k)) \\ &= x[n] *_n r_{(h,\tau)}[n] \end{aligned}$$

- **Remark:** Even if it is possible to compute $r_{h,\tau}[n]$ in closed form, this may be too expensive. For example, the windowed sinc kernel requires at least one divide per tap. We store $h(t)$ in a lookup table, taking us back to the realm of discrete polyphase filters, BUT with the added opportunity to use a cheap interpolation scheme (eg. linear) to recover extra resolution on τ.

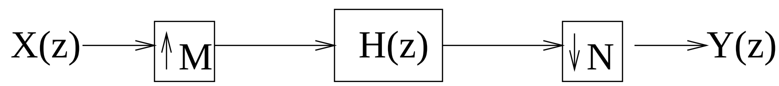
- **Exercise:** Work out the τ-polyphase decomposition for the linear interpolation kernel:

$$h(t) = (1-|\tau|)1_{\{\tau\in[-1,1]\}}$$

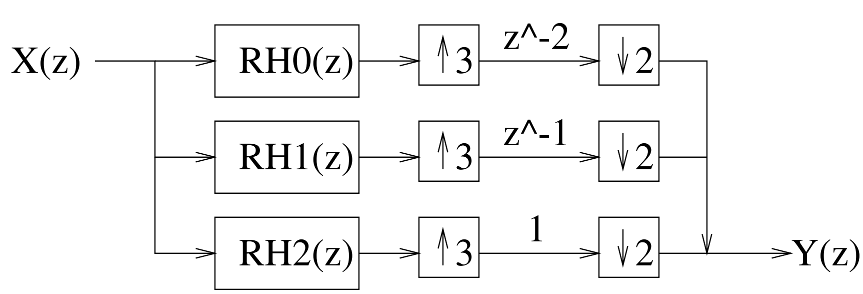
Does this make sense? When $\tau = 0$ or 1, the magnitude response of $r_{h,\tau}[n]$ is allpass as desired. In the worst case, ($\tau =$ 0.5) we obtain a halfband filter. Investigate the nonlinearity of the phase response for various values of τ.

Sampling Rate Conversion

- To begin, consider the fixed, rational case:

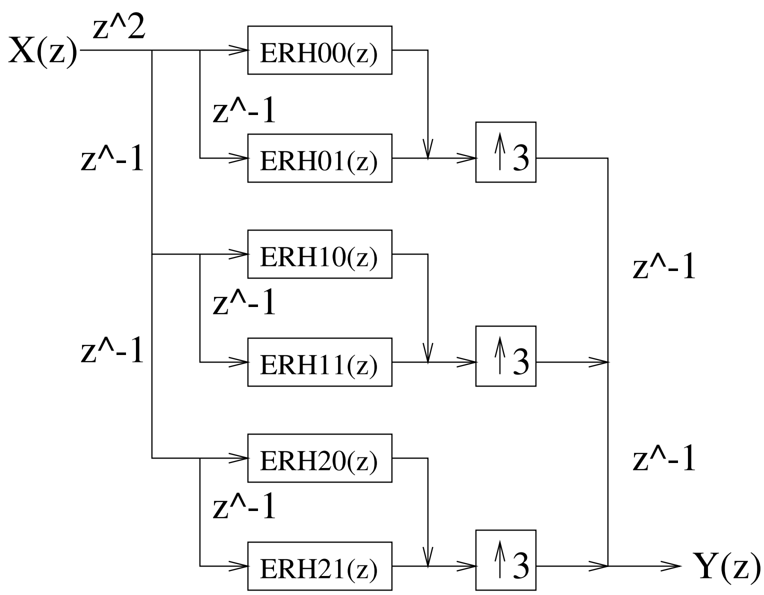


- No loss of generality is evidenced in the *requirement* that M, N be relatively prime. When either M or $N = 1$ we reduce to the familiar cases of interpolation/decimation filtering.
- $H(z)$ must be both a good reconstruction filter for upsampling by M and a good antialiasing filter for downsampling by N. Hence cutoff is set to $\min(\pi/M, \pi/N)$. Additionally, we desire zero-phase and flat magnitude response: all of which lead to the appropriate sinc filter.
- There are $M \cdot N$ times as many computations as necessary, for the input to $H(z)$ has $M-1$ out of M samples zero and we discard $N-1$ out of N samples in the downsampler. To recover this waste, we perform *two* polyphase decompositions.
- Upsampler decomposition:



To perform the second decomposition, we need some way of interchanging upsamplers and downsamplers. Intervening delays seem to prohibit this. However, as $M = 3$ and $N = 2$ are relatively prime, Euclid's theorem says we can always factorize $z^{-k} = z^{m_k M} z^{-n_k N}$ for some integers m_k, n_k. The "$z^{m_k M}$ parts" interchange with the upsamplers and the "$z^{-n_k N}$ parts" with the downsamplers. We may now perform:

- Downsampler decomposition:



Obtaining $\{m_k, n_k\}$ is left as an exercise: check your results against the above figure. Another good exercise is to verify that the number of computations per *output sample* are $\mathcal{O}(1)$ w.r.t. L.

Bandlimited Interpolation

- For synthesis and sound processing applications, fixed-rate conversion is not very useful, as topology changes drastically with M, N. However, the *idea* of two polyphase decompositions is important and should aid your conceptual understanding of what is "really used", e.g. extensions to the (continuous) domain of the τ-polyphase decomposition.

- You should now look at *Digital Resampling Home Page* by Julius Smith:

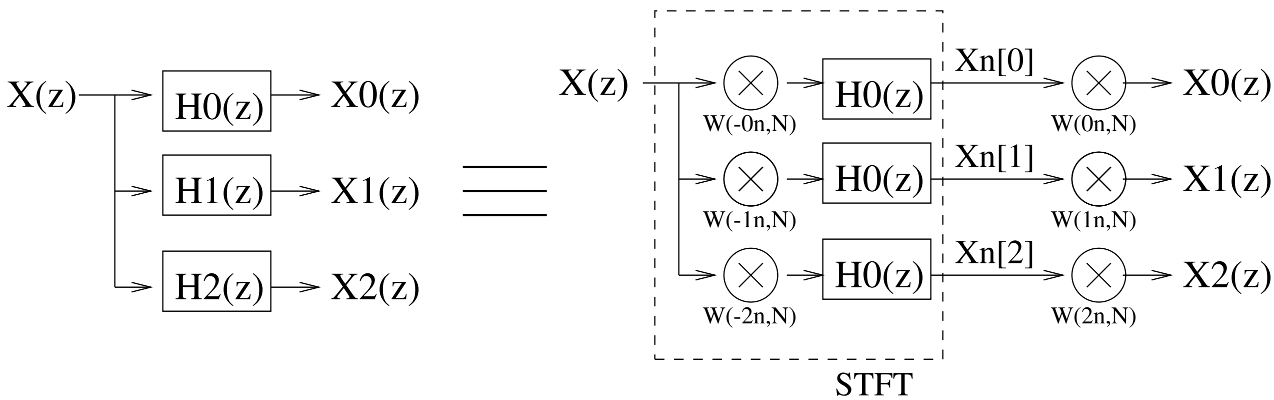
  `http://www-ccrma.stanford.edu/~jos/resample/`

  where continuous-time approach is used to introduce the topic. We return to bandlimited interpolation in Music 421.

Multirate Analysis of Filterbanks

Modulated STFT Filter Bank

- Recall the complex Portnoff analysis bank, where $H_k(z)$, $k = 1 \ldots N$ are Nth-band bandpass filters related a to lowpass prototype $H_0(z)$ by modulation (e.g. $H_k(z) = H_0(zW_{k,N}), W_{k,N} \triangleq e^{-j\frac{2\pi k}{N}}$):



- Convolution obtains $X_n[k] = \Sigma_{m=-\infty}^{\infty} x[m]h_0[n-m]W_{-km,N}$. Evidently we recover the sliding-window STFT implementation, where $h_0[n-m]$ is the sliding window "centered" at time n, and $X_n[k]$ is the kth DTFT bin at time n. After remodulating the DTFT channel outputs and summing, we obtain perfect reconstruction of $X(n)$ provided $H_0(z)$ is Nyquist(N) (e.g. $E_{(H,0)}(z) = 1$).
- For $H_0(z)$ to be a good lowpass filter, its length certainly must exceed the number of bins in the DTFT. (Otherwise, the best we

have is the rectangular window, which gives only -13 dB stopband rejection.)

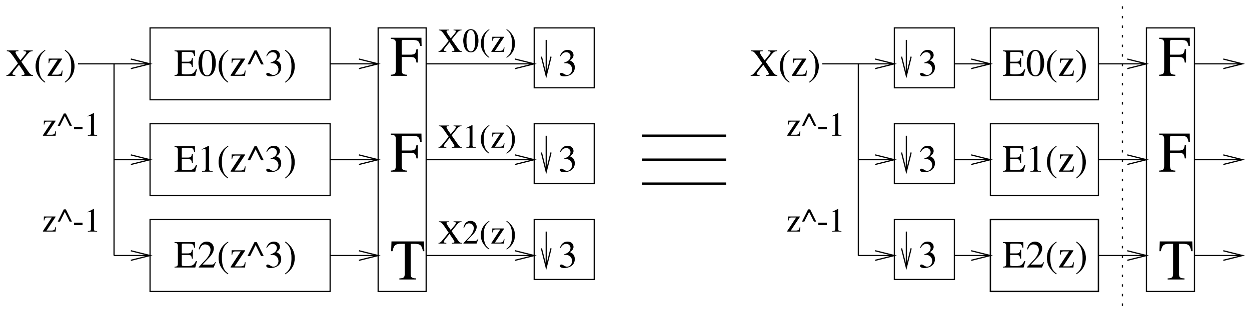
- You see the problem: Let the window length be $L > N$, then the FFT gives samples of the DTFT at $2\pi/L$-intervals; but we require these samples to be computed at $2\pi/N$-intervals. In Lecture 7, it is mentioned that we recover use of the FFT provided $h_0[n-m]$ is appropriately *time-aliased* about $[n, n+N-1]$.
- With polyphase analysis we obtain this result, along with an efficient FFT implementation.
- Proceeding, recall that $H_k(z) = H_0(zW_{k,N})$. Now by the polyphase decomposition:

$$\begin{aligned}
H_0(z) &= \sum_{l=0}^{N-1} z^{-l} E_{(H_0,l)}(z^N) \\
H_k(z) &= \sum_{l=0}^{N-1} (zW_{k,N})^{-l} E_{(H_0,l)}((zW_{k,N})^N) \\
&= \sum_{l=0}^{N-1} z^{-l} E_{(H_0,l)}(z^N) W_{-kl,N}
\end{aligned}$$

Consequently,

$$\begin{aligned}
H_k(z)X(z) &= \sum_{l=0}^{N-1} (z^{-l} E_{(H_0,l)}(z^N) X(z) W_{-kl,N} \\
\begin{bmatrix} H_0(z) \\ \dots \\ H_1(z) \end{bmatrix} &= \begin{bmatrix} \\ W_{-kl,N} \\ \\ \end{bmatrix} \begin{bmatrix} E_{(H,0)}(z^N) z^{-0} X(z) \\ \dots \\ E_{(H,N-1)}(z^N) z^{-(N-1)} X(z) \end{bmatrix}
\end{aligned}$$

- If $H_0(z)$ is a good Nth-band lowpass, the subband signals $x_k[n]$ are approximately bandlimited to a region of width $2\pi/N$. As a result, little information is lost when we downsample each of the subbands by N. Commuting the downsamplers to get an efficient implementation is straightforward:



We see that the polyphase filters compute the appropriate time-aliases of the *flipped* window $H_0(z^{-1})$, and this window is hopped by N samples.

- **Exercise:** Interpret when $E_{(H,k)}(z) = 1$ for all k.