5. On the Exclusion of Filterbank Deadbands

5.1 Overview

As observed in section 4.6, the subband hierarchy of a practical QMF filterbank has deadbands which must be excluded from the MAS paradigm in order to guarantee high-quality synthesis. *Logical exclusion* means that deadbands are excluded preemptively by a modification to the MAS allocation policy outlined in section 3.2.3. An alternative is to employ a filterbank in which deadbands are eliminated so that the subband hierarchy is that of the idealised case illustrated in Fig. 3.7. Such a technique may be described as *physical exclusion* and can be achieved by exploiting the properties of complex signal representation in a novel form of oversampled synthesis filterbank, termed the Physical Exclusion Filterbank (PEF) for convenience; the development and evaluation of which the latter part of this chapter is devoted to.

Note that the conclusion of Chapter 3, in that a hierarchical interpretation of a binary-tree subband decomposition is the optimum paradigm for MAS (in terms of both efficiency and flexibility), is accepted *per se* and forms the context for PEF design. There are two principal motives for the PEF concept. The first is to maximise the range of filterbank solutions available to MAS in terms of cost and performance: the PEF forms a useful counterpoint to the QMF approaches. Secondly, as noted in section 4.5, phase normalisation is difficult to achieve for the FIR, and unfeasible for IIR, QMF's. In order to address the desired ideal of complete phase transparency and prove that it is possible in MAS, the PEF is demonstrated to have this unique property. However, the first topic is an analysis of the effect of logical exclusion upon a conventional QMF filterbank.

5.2 Logical Exclusion in QMF Filterbanks

5.2.1 A Formal Allocation Algorithm

Given an oscillator x parameterised by $\{f_{min}(x), f_{max}(x)$, eqn. (5.1) is true for subband $s_{k,l}$ in a QMF filterbank paramaterised by Δ_f if $s_{k,l}$ completely bounds x. Using the subband hierarchy notation of section 3.2.2, the first clause pertains to odd(l) which represents

the $H0$ subbands of QMF stages in which the deadband reduces $f_{max}(s_{k,l})$. Conversely, the second clause pertains to even(l) which represents the $H1$ subbands where the deadband has the opposite effect of increasing $f_{min}(s_{k,l})$. Note that $a(s_{0,1},x)$ is true by default because it is classical AS. To take into account deadband inheritance, the optimal allocation level k in the subband hierarchy is given by the formal condition that for all $i{:}\{0 \le i \le k\}$ there exists a value of $l{:}\{1 \le l \le 2^i\}$for which $a(s_{i,l},x)$ is true and that, conversely, for all $i{:}\{k < i \le K\}$ there exists *no* value of $l{:}\{1 \le l \le 2^i\}$ for which $a(s_{i,l},x)$ is true. The optimal subband, given k, is therefore designated by the case of $a(s_{k,l},x)$ that is true for a single value of $l{:}\{1 \le l \le 2^k\}$.

$$a(s_{k,l},x) = \left\{ \begin{array}{l} (k>0) \wedge \text{odd}(l) \\ \wedge (f_{min}(x) > 2^{-k-1} f_s(l-1)) \wedge \\ (f_{max}(x) < 2^{-k-1} f_s(l-2\Delta_f)) \end{array} \right\} \vee \left\{ \begin{array}{l} (k>0) \wedge \text{even}(l) \wedge \\ (f_{min}(x) > 2^{-k-1} f_s(l-1+2\Delta_f)) \\ \wedge (f_{max}(x) < 2^{-k-1} f_s l) \end{array} \right\} \vee (k=0)$$

(5.1)

5.2.2 Allocation Maps for Logical Exclusion

Logical exclusion may be interpreted as an "allocation map" which is a function of K and Δ_f with Cartesian co-ordinates ($f_{min}(x)$, $f_{max}(x)/f_{min}(x)$). It is assumed that K=3. A point on the map denotes the optimal subband for x according to the algorithm of section 5.2.1. The first map, illustrated in Fig. 5.1, is that for an ideal subband hierarchy with Δ_f=0.0 for comparison with the second map which has the realisable figure of Δ_f=0.05. Observe that the allocation patterns illustrated in section 3.2.3 are horizontal "slices" through the Δ_f=0.0 map. The maps also help visualise the gradual transition in the proposed MAS paradigm from supporting fixed-pitch notes in the terminal integer series to those which are low-Q and require large pitch variation as discussed in section 3.2.3.

Note that, in Fig. 5.1, the severest deterioration in the subband hierarchy, as Δ_f increases from Δ_f =0.0→0.05, is in the terminal integer spaced subbands series $s_{k,l}{:}\{k=3\}$ which are prone to the inheritance of ancestor deadbands, causing a marked increase in the cost of allocating a fixed-pitch note. Indeed, a complete terminal integer spaced subband series

of bandwidth less than $\Delta_f/2$ is not possible as those either side of $f_s/4$ fall completely within the deadband from level k=0, which is another limitation on K. In contrast, note, that the relative extra cost of allocating a note with large pitch variation into the fully-overlapping octave spaced subband series $s_{k,l}$:{l=1}is slight as each inherits only the deadband of its immediate parent which causes a negligible reduction in passband width of $\times(1\text{-}2\Delta_f)$.

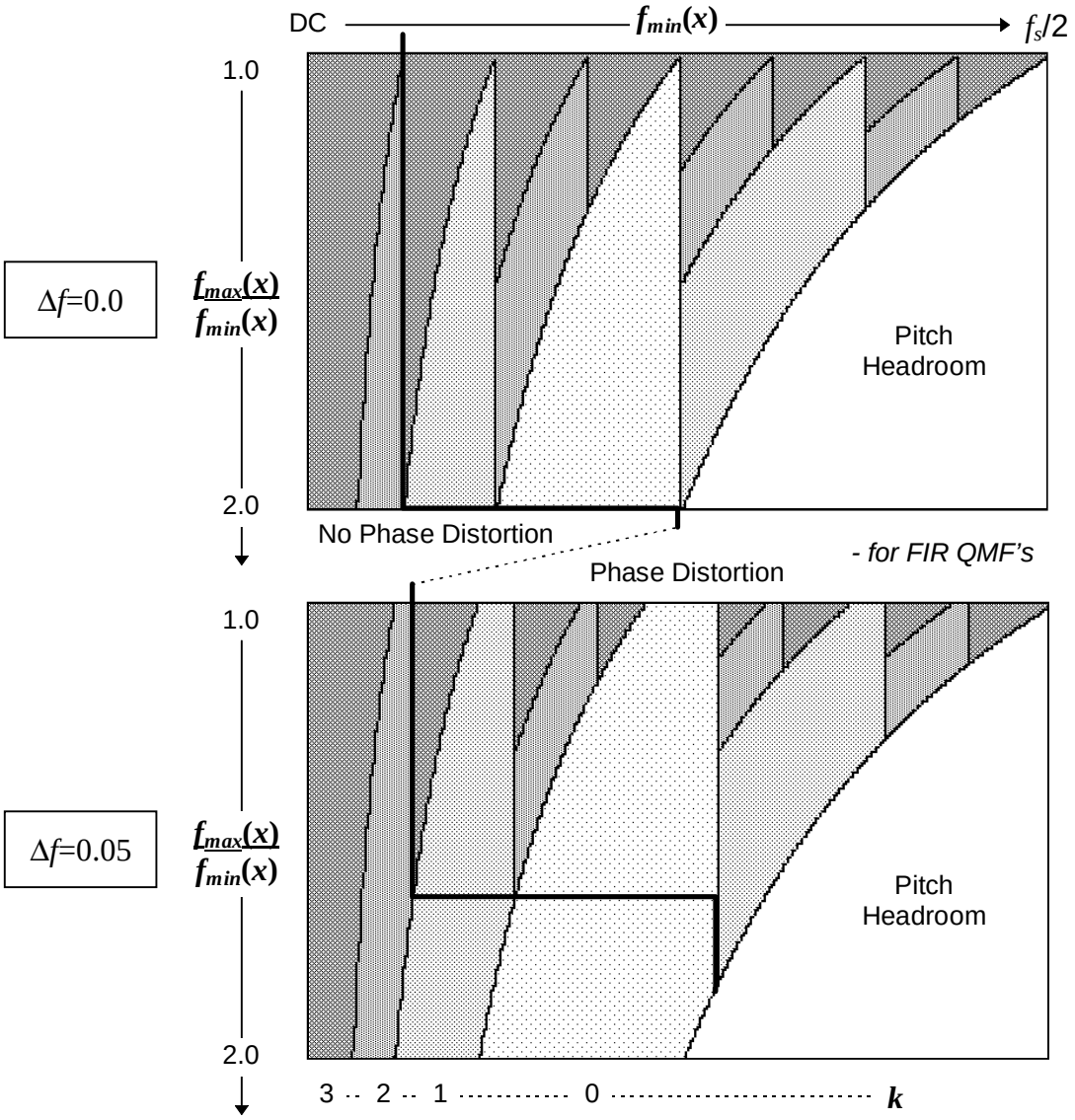

Figure 5.1 Logical Exclusion Allocation Maps for Δ_f=0.0, Δ_f=0.05 (K=3)

5.2.3 The Implications of Allocation Maps for PM-FIR and PA-IIR Filterbanks

Fig. 5.1. also differentiates between subbands $s_{k,l}$:{l=1} which involve a pure low-pass filterbank path of H0's which has the property (as introduced in section 4.3.3) that baseband phase is equivalent to output phase. For all other subbands $s_{k,l}$:{$l>1$} the filterbank path involves phase-distorting H1's. The frequency threshold between these two subband sets is illustrated to be a function Δ_f and $f_{max}(x)/f_{min}(x)$. This threshold is also improved, interestingly, by wider Δ_f. The point of section 4.3.3. - in that partial phase relationships are most important, and should be preserved, at lower frequencies and that this requirement has a convenient correspondence with PM-FIR QMF filterbank functionality - is reinforced by Fig. 5.1. The allocation map for the PA-IIR filterbank is close to the Δ_f=0.0 map, but with nil phase transparency in all subbands.

5.3 Physical Exclusion Filterbanks

QMF filterbanks are critically sampled. This is a desirable property because it minimises computation. However, the requirement for logical exclusion of deadbands causes a cost increase for oscillator computation because many oscillators are promoted to higher subbands than would be the case with an ideal hierarchy. The fundamental idea of the Physical Exclusion Filterbank (PEF) proposed in this section is that oversampling permits deadbands to be placed outside the signal baseband. In order to achieve this, extra flexibility is required in how signals may be manipulated in frequency domain and this functionality is provided by the properties of complex number representation. As low-pass filters (akin to H0 in the QMF) are used, signal basebands rather than sidebands are interpolated leading to the property of complete phase transparency (see section 4.3.3). Unlike the QMF, the PEF has no complementary analysis filterbank as it is an application-specific interpolation structure for MAS. A complex-output multirate oscillator bank is required but there are several efficient algorithms, such as CORDIC, for computing complex sinusoids which are documented in Chapter 7.

5.3.1 Discrete-Time Complex Signals

Fig. 5.2 illustrates the spectrum of a discrete time-complex signal sampled at f_s for which a frequency $e^{j\omega}$ may be conceptualised as a rotating phasor that is anticlockwise for

positive ω and clockwise for negative ω. In conventional "single-phase" sampling, where the quadrature imaginary component is absent and zero by default, a spectrum of bandwidth β (e.g. the sum of oscillators in a subband) resolves into a complex-conjugate pair of sidebands symmetric about integer multiples of f_s and therefore to avoid aliasing, must be sampled at f_s=2β: the definition of Nyquist's theorem. This relationship is re-interpreted for complex signals where positive and negative frequency are distinct. The minimum sample rate is f_s=β, but a *couplet* of samples - (*real*, *imaginary*) - is required for each sample interval and so the data bandwidth is conserved. The baseband span for such a signal is not from DC to β/2, as in the single-phase case, but from -β/2 to +β/2 with DC exactly mid-baseband. These requirements are restatements of Double Sideband (DSB) and Single Sideband (SSB) modulation (Oppenheim and Schafer, 1989).

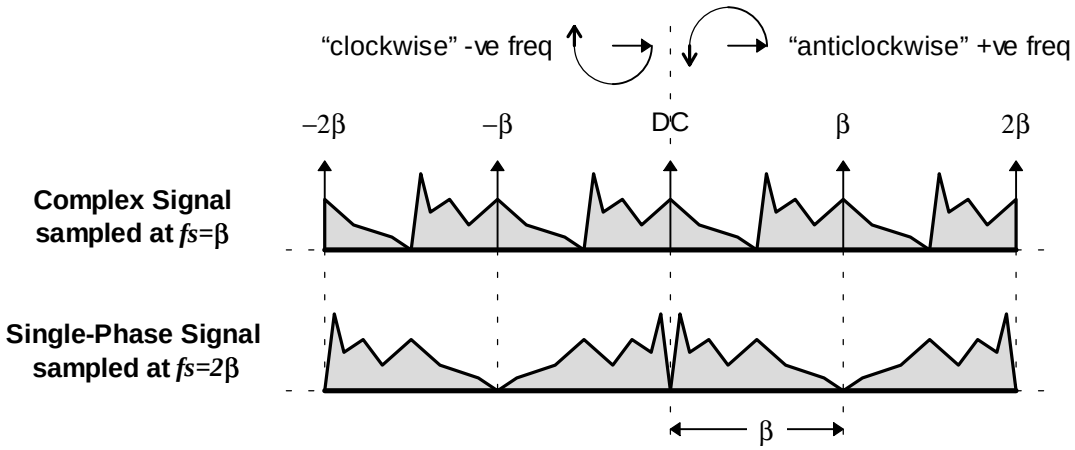

Figure 5.2 Spectra of Discrete-Time Complex and Single-Phase Signals

5.3.2 PEF Filterbank Design

The frequency domain operation of a PEF stage illustrated in Fig. 5.3 is analogous to the QMF synthesis model of section 3.2.1 in that two baseband signals $X0(z^2)$ and $X1(z^2)$, both of bandwidth β, are interpolated by ↑2 (to generate $X0'(z^2)$ and $X1'(z^2)$) and then concatenated in frequency domain to create an output $Y(z)$ of bandwidth 2β. A fundamental characteristic of the PEF is that both $X0(z)$ and $X1(z)$ are oversampled by a factor λ:{λ>1}. After ↑2 interpolation (real and imaginary signal components are filtered in separation by identical filters) a frequency shift is performed of $-\omega_n$ for $X0'(z^2)$, such that the upper frequency bound of the $X0'(z^2)$ baseband coincides with DC. Similarly, a

complementary frequency shift of $+\omega_h$ for $X1'(z^2)$ causes the lower frequency bound of $X1'(z^2)$ to coincide with DC. $\pm\omega_h$, the heterodyne frequencies, are related to λ by $|\omega_h| = \pi/2\lambda$.

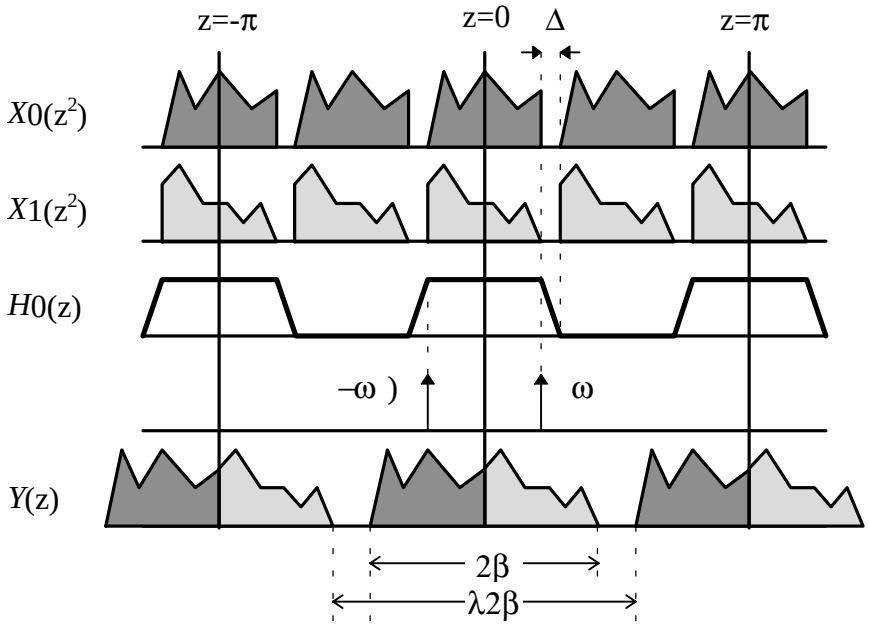

Figure 5.3 Frequency Domain Operation of PEF Stage

Superposition of the interpolated and shifted input spectra creates a perfect concatenation of basebands at DC and therefore eliminates the deadband. The transition width of $H0$ is physically excluded by mapping it on to the guard-bands between input sidebands created by oversampling. Like $X0(z^2)$ and $X1(z^2)$, output $Y(z)$ is a complex signal oversampled by λ and may therefore be input to the $X0(z)$ or $X1(z)$ input of another PEF stage. One of the hallmarks of the PEF approach is that sample rates are $\times\lambda/2$ that of the corresponding QMF stage. If $\lambda<2$ then oscillator bank PWL envelope uncompression is reduced compared to the QMF case, which is significant in that minimisation of this bottleneck is a traditional goal of AS optimisation. The transfer eqn. of the PEF stage is expressed by eqn. (5.2) where '*' denotes frequency-domain convolution.

$$Y(z) = ((X0(z^2)H0(z)) * -\omega_h + ((X1(z^2)H0(z)) * \omega_h \qquad (5.2)$$

Fig. 5.4. illustrates the structure of the PEF stage. ↑2 interpolation of the four input signal components is carried out by a FIR filter that is functionally equivalent to the optimised *H*0 of the PM-FIR QMF stage described in sections 4.2.1 and 4.2.2. By a careful choice of λ, ω_h is a rational fraction of π and may be generated by a short LUT. Frequency shifts are carried out by a complex multiplication using the identity $(a+bj)(c+dj) = (ac-bd)+(ad+cb)j$ which constitutes four multiplies per $y[m]$. Therefore, in contrast to a QMF stage which requires i multiplies per $y[m]$, a PEF stage requires $\lambda(2i+4)$ taking into account the PEF/QMF sample rate ratio of λ/2 and is thus more expensive to compute.

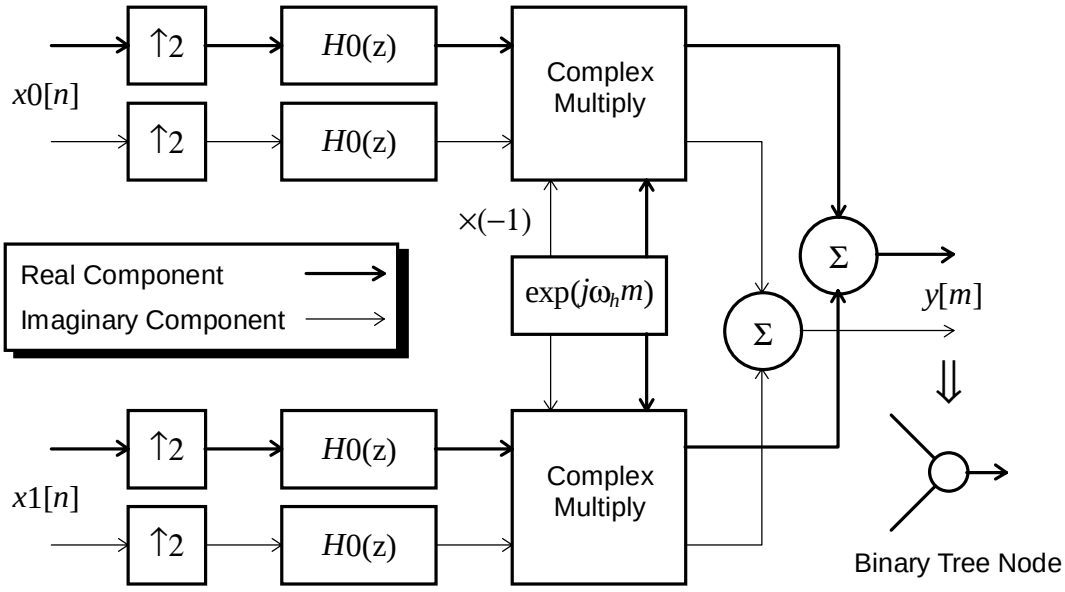

Figure 5.4 PEF Stage Dataflow

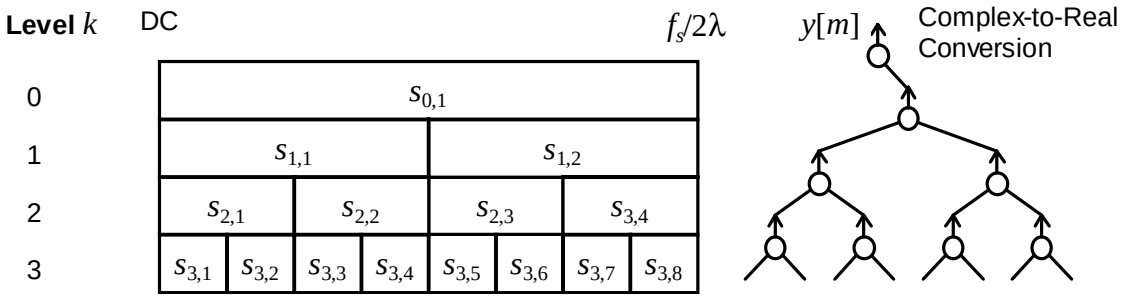

Figure 5.5 PEF Filterbank Topology for Depth K=3

In the filterbank tree of Fig. 5.5 of depth K=3, a final complex-to-real stage is necessary in order to generate digital audio. This is achieved by a PEF stage with a single $x1[n]$

input: the complex signal is interpolated and shifted to be positive frequency as before. However, the imaginary component of $y[m]$ is discarded causing a conjugate image of $x1[n]$ to appear in place of the negative-frequency shifted baseband of $x0[n]$: thus $y[m]$ has the desired characteristics of a single-phase signal. There is a direct mapping of ideal subband hierarchy to the corresponding filterbank tree because no frequency-inverting $H1$ filters are used. Also, it can be appreciated how the insertion of a complex-to-real stage "halves" sample rates in comparison to the corresponding subbands in the QMF case of Fig. 3.7, though, in practice, oversampling creates a practical figure of $\times\lambda/2$.

5.3.3 Determination of Oversampling Factor λ

As the PEF is designed to be phase-transparent, linear-phase FIRs must be used in order to avoid the complexity of phase-equalisation in phase non-linear PA-IIRs. λ directly determines the FIR design for $H0$ via substitution of relationship $\Delta_f=(1-1/\lambda)/2$ in eqns. (4.4) to (4.6) resulting in the inverse proportionality of M versus λ illustrated in Fig. 5.6. Clearly, a value of $\lambda\cong1$ results in an FIR with unacceptable latency and computational cost, though oscillator bank redundancy is minimised as it approaches critical sampling. At the other extreme, $\lambda>>1$ results in an FIR of diminishing latency and cost, but with high redundancy in an oversampled oscillator bank.

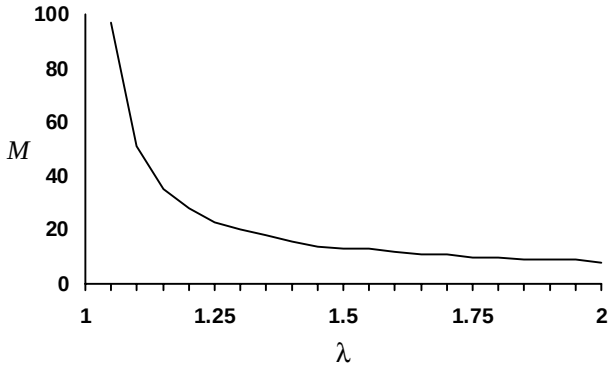

Figure 5.6 FIR M versus λ

Such characteristics are a parallel to generalised MAS economics as summarised in eqn. (3.1). A compromise value is necessary. Experience from simulations (see section 5.3.8) suggest that $\lambda=1.25$ is a good choice for a PEF filterbank with depth $K=3$, requiring an

FIR of M=23 (Δ_f=0.1). f_s= 2λ×20kHz= 50kHz in order to provide a full audio bandwidth of DC-20kHz, though a slightly lower industry-standard rate such as DAT @ f_s=48kHz would be utilised in practice giving an audio spectrum ceiling of 19.2kHz. Another reason for a rational value of λ=1.25 is that ω_h=0.4π and may thus be generated by a trivial five couplet LUT (e^{0j}=$e^{2\pi j}$, 5×0.4π=2π, *QED*).

5.3.4 Latency Normalisation

The scheme of section 4.3.1. applies equally well to the PEF filterbank, except that two delay-lines in parallel are required for real and imaginary signal components. A PEF filterbank has a latency $2T_{fb}(K)/\lambda$ (×1.6 for λ=1.25) of the corresponding PM-FIR QMF with stages of equal M due to the lower sample rates of the PEF filterbank.

5.3.5 Frequency Normalisation

There are two differences in the frequency normalisation scheme for a PEF filterbank compared to that of section 4.3.2. for a QMF filterbank. The first is that there are no frequency inversions caused by high-pass filtration via $H1$. Secondly, the subband normalised digital frequency envelope $\Omega_x[m]$ of an oscillator x allocated to subband $s_{k,l}$ satisfies $-\omega_h \leq \Omega_x[m] \leq \omega_h$ and thus may be positive or negative frequency as expressed in eqn. (5.3). After processing through the PEF filterbank, x emerges with the expected single-phase frequency $F_x[m]$.

$$\Omega_x[m] = (1-2a)\omega_h \text{ where } a = \left(\frac{F_x[m] - f_{min}(s_{k,l})}{f_{max}(s_{k,l}) - f_{min}(s_{k,l})} \right) \tag{5.3}$$

5.3.6 Phase Normalisation

Fig. 5.7 illustrates the path of a signal $x[n]$ in a PEF filterbank from an arbitrary subband $s_{k,l}$ to the output $y[m]$. The vector $\mathbf{v}$:$\{\mathbf{v}_i \in \{-1,+1\}\}$, where $0 \leq i \leq k$, defines the top-down path to $s_{k,l}$ through the binary tree of the filterbank by specifying the heterodyne frequency direction at each stage: $\mathbf{v}_0$=+1 denotes complex-to-real conversion. Latency-normalisation is shown (null for k=K) and PM-FIR $H0$ filters for ↑2 interpolation are conceptualised as delay-lines of length M. As there are no $H1$ filters, the phase-distortion

discussed in section 4.3.3. is eliminated. However, a mechanism is required to cancel the time-variance introduced by the heterodyne oscillators. An initial insight to a solution lies in the functional equivalence of the path **v** to single-stage interpolation by $I=\uparrow 2^{k+1}$ - with a delay of T samples obtained from eqn. (5.4) - followed by multiplication by a single oscillator ω_{eq} equal to the sum of the constituent heterodyne frequencies in **v** as expressed in eqn. (5.5).

$$T = (2^{K+1} - 1)M \tag{5.4}$$

$$\omega_{eq} = \omega_h \sum_{i=0}^{k} 2^{-i}\,\mathbf{v}_i \tag{5.5}$$

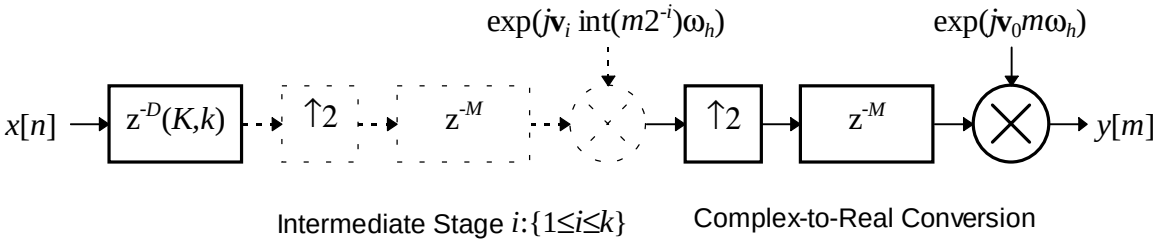

Figure 5.7 Path Model from Level k in a PEF Filterbank

To enable an oscillator x to emerge at $y[m]$ at a specific time $m=t$ with correct phase (given **v** and k), the phase offset introduced by heterodyning (denoted $\theta[t]$) is predicted and its negative used to initialise $\Phi_x\,[t\text{-}T]$ and thereafter, the PEF filterbank is phase-transparent upon condition that the partials for a note are accurately phase-accumulated under decimation using eqn. (4.10). $\theta[t]$ is given by eqn.(5.7) with ω_{eq} from eqn. (5.5) and $\theta[T]$ from eqn. (5.6) which originates in the assumption that the instantaneous phase of all heterodyne oscillators at $m=0$ is zero, and by tracing the signal path to the output $y[m]$ at $m=T$ and summing the phase of each heterodyne frequency as it is encountered. A constraint on t is that $t=T+Dc$ where $D=2^{k+1}$ and c is an integer.

$$\theta[T] = \sum_{i=0}^{k} \mathbf{v}_i (2^{K-i+1} - 1) M\omega_h \tag{5.6}$$

$$\theta[t] = (t - T)\omega_{eq} + \theta[T] \tag{5.7}$$

5.3.7 Functionally Equivalent Parallel-Form PEF Filterbanks

A complete binary-tree PEF is only one of a number of possible topologies and, as hinted in section 5.3.6, it is feasible to construct a parallel-form PEF with an individual interpolator and heterodyne oscillator for each subband. Their outputs are summed and undergo complex-to-real conversion. Consideration of the parallel-form is significant as single step interpolation for a high integer factor $\uparrow I$ has two efficient implementations as discussed in section 3.1.2. Table 5-1 records the performance, determined by a spreadsheet model, of the functionally equivalent multi-stage and single-stage parallel forms of a complete binary tree PEF of depth K=3, for input λ=1.25 and PM-FIRs of δ_p=δ_s=0.0001 (-80dB) derived using the design eqns. (4.4) to (4.6).

	Single-Stage	**Multi-Stage**	**Binary Tree**
Max Latency (ms)	7.86	4.86	6.90
Multiplier B/W (MHz)	6.3	4.5	3.1

Table 5-1 Equivalent Filterbank Performance for K=3

In latency terms, the performance ranking from best to worst is (1) multi-stage, (2) binary-tree and (3) single-stage. The binary-tree requires the lowest multiplier bandwidth because adding an extra level $K \rightarrow K+1$ doubles the number of PEF stages, but since they operate at half the sample rate of the ancestor level the computational cost is proportional to K. In contrast, all interpolators in a parallel form have the same final sample rate. The regularity of a binary-free based upon a single prototype stage implies a simpler control structure than that required by the variety of filters in a parallel form. In conclusion, a binary-tree topology appears the optimal choice for a PEF implementation.

5.3.8 Simulation

In order to validate the theoretical PEF design and its normalisation schema, a simulation was written in 'C' under UNIX. The program included a binary-tree PEF of depth K=3, with PM-FIR stages of latency M=23 and an allied complex multirate oscillator bank which could be initialised to generate classical AS waveforms. A sawtooth function with fundamental frequency ω_0, as expressed in eqn. (5.8) where t is the time index, is a satisfactory test signal as it includes every harmonic partial, thereby maximising the number of oscillators in the composite waveform, and also making phase distortion phenomena manifest itself readily in the waveform envelope. The imaginary component is taken from the complex-to-real stage of the PEF because eqn. (5.8) is a sum-of-sinusoids. In the simulation, all harmonics up to 20kHz are included.

$$f_{sawtooth}(\omega_0, t) = \sum_{i=1}^{\infty} \frac{\sin(it/\omega_0)}{i} \tag{5.8}$$

Fig. 5.8 illustrates the output - $y[m]$ *versus* sample index m - for a 110Hz note (A2) assuming f_s=50kHz with frequency normalisation, but excluding that for phase. The characteristic envelope of a sawtooth is corrupted by incorrect phase relationships. Fig. 5.9 includes the phase normalisation scheme of section 5.3.6 and validates the functionality of the proposed technique by generating a high purity sawtooth. Additionally, frequency normalisation as proposed in section 5.3.5 is verified. The latency of 345 samples or 6.9ms @ f_s=50kHz, as obtained via eqn. (5.4), is evident as is the transient response of combined filterbank paths to a step amplitude envelope. As the note is stationary, partials are allocated into the integer-series subbands at level k=3: all oscillators are operating at an effective sample rate of 50kHz/2^{k+1} =3.125kHz in comparison to the uniform f_s>40kHz of a TOB. Latency normalising delay lines are redundant in this example.

To validate phase accumulation in MAS according to the scheme of section 4.3.2, a 'chirped' sawtooth with linearly increasing frequency is an appropriate signal: the Fourier series is truncated to avoid partials aliasing at the maximum fundamental frequency. Fig. 5.10 illustrates an example of a chirp from 27.5Hz to 440Hz; four octaves from A0 to

A4. The output sawtooth maintains phase integrity under chirping. In contrast, partials in this example are allocated into the fully-overlapping octave spaced subbands series $s_{k,1}$ disposed across layers k=0..3. Latency normalising delay lines are therefore employed and demonstrated to operate correctly. The net conclusion of these and many other runs is that the binary-tree PEF permits a MAS paradigm with high functional transparency as desired in the performance criteria of section 1.2.1.

5.4 Review

A 'logical exclusion' algorithm is proposed to overcome the problem of deadbands in a QMF subband hierarchy. A side-effect is that oscillator computation is rendered more expensive in proportion to the dominance that deadbands play in the resulting subband hierarchy: high for the PM-IIR and low for the PA-IIR. In consideration of the phase normalisation problem of the QMF, a novel alternative filterbank design - the PEF- is proposed. Simulation validates that it achieves high functional transparency. However, the PEF has the highest computational overhead of filterbanks envisaged for MAS. The perceptual relevance of phase is the key determinant in which approach is adopted; a subjective issue and of an importance that cannot be dismissed lightly when high-fidelity synthesis is required. It has not been the purpose of this chapter to recommend a prescriptive solution. However, a PM-FIR QMF with limited phase transparency, as discussed in section 4.5, may prove an acceptable compromise between the extremes of PA-IIR QMF and PEF designs.

5.5 PEF Simulation Results

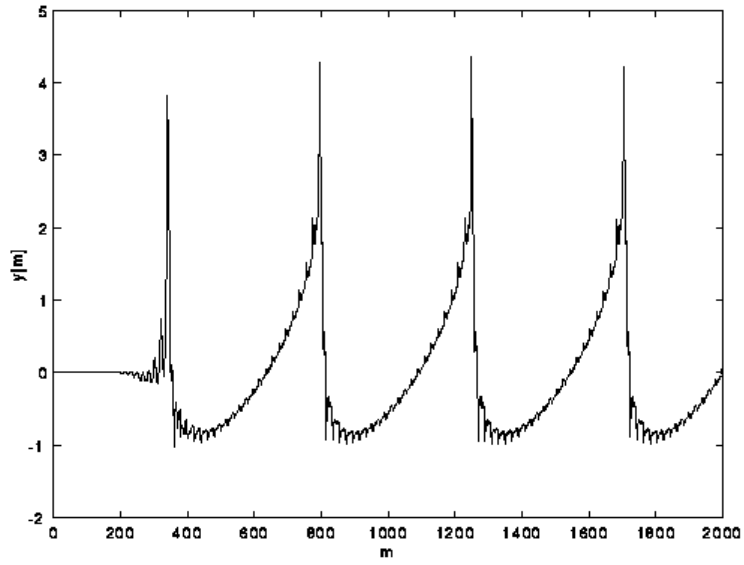

Figure 5.8 110Hz (A1) Sawtooth without Phase Normalisation: y[m] versus m

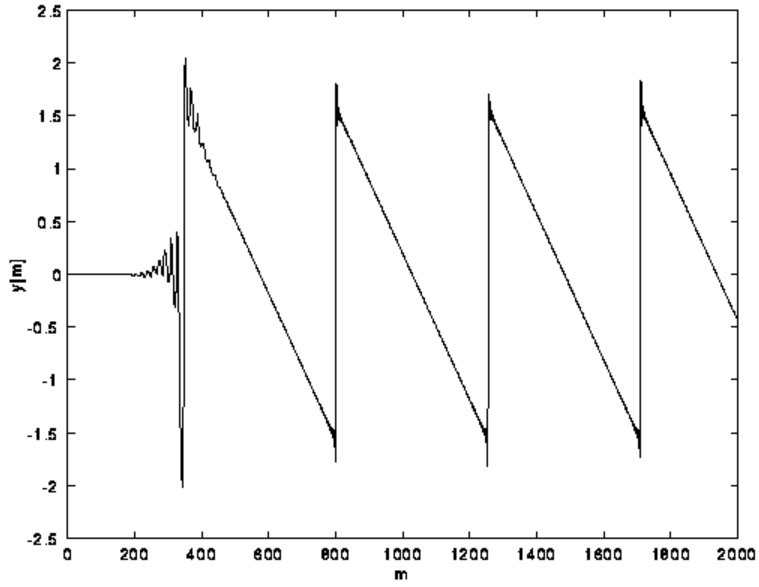

Figure 5.9 110Hz (A1) Sawtooth with Phase Normalisation: y[m] versus m

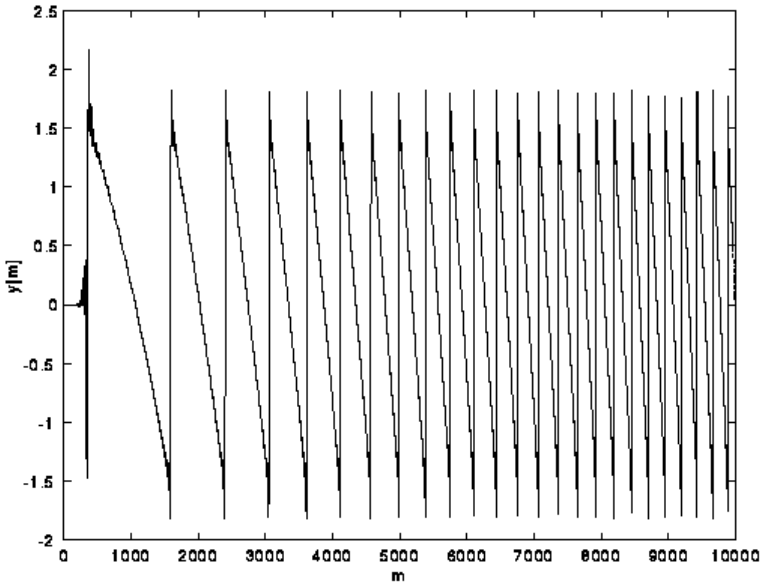

Figure 5.10 27.5Hz to 440Hz (A0 to A4) Chirped Sawtooth: y[m] versus m