Preprint 3030 (E-2)

Digital delay networks for designing artificial reverberators

Jean-Marc Jot, Digitec SA, Chatou France
Antoine Chaigne, Enst, département SIGNAL,
Paris, France

Presented at
the 90th Convention
1991 February 19-22
Paris

AES

AN AUDIO ENGINEERING SOCIETY PREPRINT

DIGITAL DELAY NETWORKS
FOR DESIGNING ARTIFICIAL REVERBERATORS

Jean-Marc JOT *+, Antoine CHAIGNE *

* *ENST, dept SIGNAL, 46 rue Barrault, 75634 PARIS, FRANCE*
+ *DIGITEC SA, 25 av. de l'Europe, 78400 CHATOU, FRANCE*

ABSTRACT

In the design of artificial reverberators, a critical problem is to avoid unpleasant resonances in the response to short transients. It is proposed to solve this problem by ensuring that all resonances in any narrow frequency band have the same decay time. This condition appears equivalent to providing each delay with a frequency dependent attenuation that is analogous to absorption in the air, irrespective of the reverberator structure. In addition, the design of a tone corrector associated in series with the system is described, to allow separate control of the decay characteristics and the frequency response. This results in a general method for designing a multiple feedback "reverberant filter" simulating the late mono reverberation. Questions concerning improvement of the early response and binaural aspects are briefly mentionned in conclusion.

0. INTRODUCTION

Artificial reverberators are used to add reverberation to studio recordings in the music and film industry, or to modify the acoustics of a listening room. Early analog devices - using springs or plates - have been replaced in the last decade by digital units. These units perform real-time computations to process the input signal (converted to digital form). A variety of algorithms based on recursive digital delay networks have been proposed in the literature since the pioneering work of Schroeder in the early 60's [1]-[5]. Even with a large number of delay taps or with time-varying delay lengths (as suggested in [3]), it is often difficult to eliminate unnatural resonances, which cause a characteristic sound often referred to as "metallic". Within the framework of Schroeder's parallel comb filter reverberator ([1], [2]), a technique is proposed for controlling the decay characteristics, concentrating on the modal decomposition of the system response. Then, general properties of multiple feedback systems are investigated, allowing extension of the technique to any recursive delay network, which provides low-cost solutions to the problem of increasing the echo density in the reverberator response.

1. REVERBERANT FILTER

Real-time artificial reverberation

The existence of multiple acoustical paths from a source to a listener placed in an enclosure results in a dense pattern of echoes which forms the reverberation process. Accurate simulation would imply taking many physical factors into account: geometry and acoustical properties of the walls and obstacles, positions and directivity patterns of the source and receiver... the computational requirements are still beyond the scope of real-time implementation with today's hardware. Assuming that all physical phenomena are linear, the reverberation process is fully described by a binaural impulse response. To simulate usual rooms, real-time convolution of the input signal by a measured or calculated impulse response also requires a very large processing power. However, although extremely accurate simulation is necessary for some applications (echo cancellation), such accuracy is not necessary to achieve a convincing artificial reverberation effect. Here, our purpose is to realize real-time artificial reverberation that is *perceptively* indistinguishable from real reverberation.

Perceptive requirements for the reverberant filter

Many psychoacoustical studies have been carried out in order to propose criteria for describing the acoustical quality of listening rooms, starting with the definition of Sabine's reverberation time. Since the work of Barron [6], it has been common usage to distinguish two parts in the reverberation process: *first reflections* (approximately the first 80 ms in the impulse reponse) and *late reverberation* (the remaining part of the decay). The early response is made of discrete echoes whose time and amplitude distributions depend critically on the shape of the room and the positions of the source and receiver. These echoes play a key role in the subjective *spatial impression*. In contrast, late reverberation lends itself to a more statistical description, and can be considered as a characteristic of the enclosure itself, *independent of the source and receiver positions*. These observations justify a design procedure focussing, first of all, on designing a *"reverberant filter"* supposed to simulate the late monaural reverberation. Then, binaural aspects are considered and solutions for improving the realism of the early response are investigated, which, in earlier work ([2], [3]), involved physical modelling of sound propagation in the room. Recent work has resulted in the proposition of a manageable number of independant psychoacoustical criteria of room acoustic quality [7], [8]. For late mono reverberation alone, it indicates that 6 independent command parameters should be sufficient to control the reverberant filter: the reverberation level and the reverberation time must be frequency dependent and controlled separately, in three frequency bands.

Perceptive comparison of the comb filter and the all-pass filter

The comb filter and the all-pass filter ([1], [2]) are single-delay IIR filters differing only by the addition of a direct path in the case of the all-pass filter (Fig.1). Schroeder [1] noticed that this simple modification was sufficient to give the all-pass filter a flat frequency response (for a particular amplitude of the initial pulse in the time response). In the case of a stationnary input signal, the all-pass filter eliminates the strong coloration caused by the comb filter.
However, the all-pass filter's *response to short transients* still reveals two major shortcomings:
a) The *"echo density"* in the time response is not high enough (causing a "fluttering" sound).
b) The comb filter's timbre is still present. This is clearly observed by *listening to the impulse responses themselves* for comparison. The effect depends largely on the magnitude of the feedback gain g. It should be noticed that when the magnitude of g approaches 1 (the stability limit), the all-pass filter's impulse response is reduced to the initial pulse (of amplitude -g, while the second pulse is of amplitude $1-g^2$, (Fig.1b)) In this case the all-pass filter has simply no effect on the input signal! For smaller magnitudes of g, the comb filter's timbre can be heard in the end of the impulse response of the all-pass filter, and for $|g| = 1/\sqrt{2}$, it becomes very difficult to hear a difference between these two filters by listening only to their impulse responses. This result is not surprising because the two time responses are identical except for the first pulse, as shown in fig.1. These observations illustrate the influence of the input signal on listening tests, and indicate that listening to the impulse response itself gives useful information for assessing the quality of a reverberator. A comparable test consists in running a Short-Time Fourier Transform (sonogram) along the impulse response, as suggested in [5].

Schroeder's parallel comb filter structure

To improve the "echo density" (number of echoes per second in the impulse response) while avoiding coloration, Schroeder proposed two basic combinations of the above unit filters: series association of all-pass filters, and parallel association of comb filters.
A *series all-pass filter* (Fig.2) yields a new all-pass filter (having a flat frequency response). It also produces a build-up of the echo density along the impulse response, in a way that is similar to what occurs in real rooms. Unfortunately, unnatural colorations are still present in the response to short transients, as reported by Moorer [2].
With the *parallel comb filter* (Fig.3), a flat frequency response cannot be achieved. However, if the frequency response exhibits a sufficient number of peaks per Hz, due to the addition of resonances from all comb filters, it becomes closer to that of a real room. More interestingly, the timbres of the individual comb filters will disappear *even in the response to impulsive sounds*, provided that the reverberation times of all unit comb filters are made equal. This will be studied below by decomposing the system response on its individual eigenmodes.

2. PARALLEL COMB FILTER: CONTROL OF DECAY CHARACTERISTICS

Pole study of the parallel comb filter

The transfer function of the comb filter shown in Fig.1a can be written:

$$C(z) = \frac{g}{z^m - g} = \frac{1}{m} \cdot \sum_{k=0}^{m-1} \frac{z_k}{z - z_k} \qquad (1)$$

The poles z_k ($0 \leq k \leq m-1$) are defined by $z_k = \gamma \cdot e^{j\omega_k}$, where $\gamma = g^{1/m}$ and $\omega_k = 2k\pi/m$.
By inverse z-transform, this yields the impulse response:

$$C(nT) = \frac{1}{m} \cdot \sum_{k=0}^{m-1} z_k^n \qquad (\text{for } n \geq 0) \qquad (2)$$

By grouping pairs of conjugate poles, this response can be expressed as a sum of exponentially decaying sinusoïds (if $g < 1$), forming an harmonic pattern of resonant frequencies. As shown in Fig.4a, the decay times of the eigenmodes - determined by the magnitudes of the associated poles - are all equal, and so are their excitations (or weigths).
When P such comb filters are associated in parallel, the system transfer function becomes:

$$C(z) = \sum_{p=0}^{P-1} \frac{g_p}{z^{m_p} - g_p} = \sum_{p=0}^{P-1} \sum_{k_p=0}^{m_p-1} \left[\frac{1}{m_p} \cdot \frac{z_{k_p}}{z - z_{k_p}} \right] \qquad (3)$$

The response results from the sum of the eigenmodes of all comb filters. If the delays are incommensurate, all the resonant frequencies are distinct (except at $\omega = 0$ or $\omega = \pi$) and the total number of resonances is equal to half the sum of the delay lengths expressed in samples.

Avoiding unnatural resonances: equal magnitude of the poles

If some resonances have a slower decay than the others, they will stand out in the end of the impulse response, revealing in the present case the timbre of a particular comb filter. To avoid any such disturbance, we must make sure that all modes have the same decay time i.e. that *all the system poles are of equal magnitude*. For the parallel comb filter, this yields the condition:

$$\gamma = g_p^{1/m_p} \qquad \text{for any } p \qquad (4)$$

which is a relation between the feedback gains of all comb filters, depending on the choice of delay lengths. If this condition is fulfilled, an echogram similar to the one shown in Fig.5 is obtained. The typical modal decomposition of the response is shown in Fig.4b.

Frequency density

The condition (4) garanties that the timbre of the impulse response remains unchanged along the decay. However, the comb filters with longer delays produce eigenmodes with weaker weigths, because of the normalizing coefficients $1/m_p$ in Eq.(3). Thus, for the same "modal density" (average number of resonances per Hz), the perceived coloration will be lower if the delay lengths are kept within a close range (which may explain why Schroeder proposes approximately 1:1.5 [1]). This suggests introducing a distinction between the (theoretical) *"modal density"* and a (perceived) *"frequency density"* (the latter being always lower than the modal density). There are at least two ways of detecting insufficient frequency density:
- the response to an impulsive input signal will produce "ringing" of particular modes, or beating of pairs of modes.
- the response to a quasi-stationnary input will produce excessive levels for some particular frequencies (as with some notes of a flute or a voice).

Nonetheless, Eq.(3) shows that if the outputs of the comb filters are weighted before summation (proportionally to their individual delay lengths), the frequency density will be kept equal to the modal density (that is, to the sum of the delay lengths expressed in seconds), since all modes will always have equal excitations (Fig.6).

Time density

In the case of a time response similar to that of Fig.5, the term "echo density" has an obvious meaning because the amplitudes of the echoes exactly follow an exponentially decreasing envelope. However, this is not the case in real rooms, where successive echoes can be of very different amplitudes, and soon overlap each other so closely that no distinct echo can be extracted from the response. This suggests using the term *"time density"* to refer to the perceived sensation, as above in the frequency domain. In the case of Fig.5, the "time density" will be considered equal to the echo density (that is, the sum of the reverses of the delay lengths).

For a parallel comb filter with P delays whose lengths τ_p are distributed over a close range around an average of τ seconds, we may write approximately:

Frequency density: $$D_f = \sum_{p=0}^{P-1} \tau_p \approx P.\tau \tag{5}$$

Time density: $$D_t = \sum_{p=0}^{P-1} \frac{1}{\tau_p} \approx P/\tau \tag{6}$$

This leads to: $$P \approx \sqrt{D_f \,.\, D_t} \quad \text{and} \quad \tau \approx \sqrt{D_f / D_t} \tag{7}$$

If, as suggested by Schroeder, we want $D_t = 1000$ echoes per second and $D_f = 0.15$ modes per Hertz, we need 12 comb filters in parallel, for an average delay length of about 12 ms. However, for a sampling frequency of 44.1 Hz, this time density still sounds insufficient (it has been noticed that the time density should be made larger when the bandwidth of the reverberation device is increased [5]). Actually, about 10000 echoes per second are considered necessary to obtain a smooth response to short clicks, which means that we need as many as 40 comb filters! With fewer filters, it is necessary to reduce the delay lengths so that the time density remains unchanged, but this causes the coloration to increase, as the frequency density becomes lower. Furthermore, Schroeder showed that the average separation of extrema in the frequency response curves of large rooms is inversely proportional to the reverberation time. This indicates that the frequency density should be made proportional to the reverberation time, raising even further the number of comb filters required for long reverberation times (a density of 0.15 modes / Hz is for a reverberation time of approximately 1 second [1]).

Reverberation time control

Eq.(4) can be more conveniently written if γ and g_p are expressed in decibels:

$$\Gamma = G_p / m_p \qquad \text{for any } p \tag{8}$$

where $\Gamma = 20 \log(\gamma)$ and $G_p = 20 \log(g_p)$. All eigenmodes have the same decay time, and Γ is their attenuation rate expressed in decibels per sample period. Thus, the reverberation time is defined by:

$$Tr = -60\, T / \Gamma = -60\, \tau_p / G_p \tag{9}$$

where T is the sampling period and $\tau_p = m_p.T$ is a delay length expressed in seconds.
As a consequence, there are two ways of modifying the reverberation time:
a) divide all the feedback gains G_p (expressed in decibels) by the same coefficient α.
b) multiply all the delay lengths τ_p by the same coefficient α.

Both of these operations will multiply the reverberation time by α. Method (b) will also modify the time density and frequency density, and is somewhat analogous to multiplying the room dimensions by α. The attenuations G_p are analogous to absorptions due to propagation of the sound in the air, since they are proportional to the delay lengths (and, consequently, to the distance covered by the sound). However, they also account for the average absorption of the walls, which should be considered separately when one wants to simulate a modification of the room size.

Obtaining a frequency dependent reverberation time

Schroeder [1] suggested making the feedback gains frequency dependent so that the reverberation time could become a function of frequency (it is usually longer at low frequencies in real rooms). Moorer [2] inserted a first order IIR filter in each feedback loop, and sought to optimize the filter coefficients in order to simulate the low-pass effect of air absorption.

Replacing each attenuation g_p by an "absorbent filter" with tranfer function $h_p(z)$, as shown in Fig.7, we will now study how these transfer functions should be chosen to avoid unnatural resonances and how they are related to the reverberation time as a function of frequency. The condition expressed by Eq.(4) is extended as follows: in any sufficiently narrow frequency band (where the reverberation time can be considered constant) all eigenmodes must have the same decay time. Equivalently, system poles corresponding to neighbouring eigenfrequencies must have the same magnitude. This condition can be called *"continuity" of the pole locus* in the z-plane (Fig.8). The magnitude of the poles becomes a function of the normalized frequency ω, and we must select the transfer functions $h_p(z)$ so that if a pole is located at frequency ω, then its magnitude expressed in decibels as in Eq.(9) is:

$$\Gamma(\omega) = -60\,T / Tr(\omega) \qquad 0 \leq \omega \leq \pi \tag{10}$$

Moorer [2] studied how the poles of a comb filter are modified by inserting a filter in the loop. This problem was also investigated in [9] and [10], where an algorithm based on a modified comb filter similar to the one shown in Fig.7 is used to synthesize plucked-string timbres.

If the absorbent filter realizes a pure frequency-dependant gain $g_p(\omega) \doteq |h_p(e^{j\omega})|$, the resonant frequencies of the comb filter are unchanged and a pole located at normalized frequency ω is of magnitude:

$$\gamma(\omega) = |h_p(e^{j\omega})|^{1/m_p} \tag{11}$$

This condition replaces Eq.(4) and must be respected for any p to avoid unnatural resonances. In this case, multiplying both sides in Eq.(10) by m_p:

$$\log_{10}|h_p(e^{j\omega})| = -3\,\tau p / Tr(\omega) \tag{12}$$

More generally, we must consider the phase delay introduced by the tranfer function $h_p(z)$, which makes the effective loop length frequency dependent, thus slightly modifying the eigenfrequencies of the comb filter. For our purpose, it is not necessary that the eigenfrequencies remain exactly unchanged when the decay characteristics are modified. However, to fulfill exactly the condition of "continuity" of the pole locus, the delay length must be replaced by the effective loop length in the above equations, and Eq.(12) becomes:

$$\log_{10}|h_p(e^{j\omega})| = \frac{-3}{Tr(\omega)} \cdot \left(\tau_p - \frac{\angle h_p(e^{j\omega})}{\omega} \right) \tag{13}$$

Where $\angle$ denotes the argument of a complex number.

It is more convenient to use Eq.(12) for designing the absorbent filters with a desired $Tr(\omega)$ curve. Again, it appears that each absorbent filter should realize an attenuation analogous to absorption in the air, in agreement with Moorer's method, although these filters must also account for the frequency dependent wall absorption. If more accuracy is needed, rather than using Eq.(13), we may design linear phase absorbent filters producing a frequency dependent gain with a frequency independent phase delay of one sample (the delay lengths m_p being reduced by 1). This will keep the effective loop length constant, and the resonant frequencies unchanged. In our implementation with first order IIR absorbent filters (described below), a satisfactory result was obtained by using simply Eq.(12).

Independant control of frequency response and decay characteristics

As mentionned in part 1 of this paper, the reverberant filter must allow independant control of the reverberation time and the reverberation level (both functions of frequency). Unfortunately, making the feedback gain frequency dependant in the comb filter modifies the envelope of the frequency response, as shown in Fig.7. This is due to the fact that modifying the magnitude of a pair of conjugate poles in Eqs.(1) and (2) modifies the decay time, but also the energy of the associated eigenmode. The energy of the mode, expressed in the time domain, is the energy of the corresponding exponentially decaying sinusoïd. Thus, if the excitation of the mode remains unchanged, its energy is simply proportional to its decay time.

Choosing a reference ω_0 on the frequency scale, we may write:

$$\frac{l(\omega)}{l(\omega_0)} = \frac{Tr(\omega)}{Tr(\omega_0)} \tag{14}$$

Where $l(\omega)$ is the energy of an eigenmode located at frequency ω. If the condition of "continuity" of the pole locus is fulfilled for all comb filters, the envelope of the global frequency response can easily be kept independant of the decay characteristics by associating a tone corrector $t(z)$ in series with the parallel comb filter, as shown in Fig.9, provided that:

$$\frac{|t(e^{j\omega T})|}{|t(e^{j\omega_0 T})|} = \frac{Tr(\omega_0)}{Tr(\omega)} \tag{15}$$

Implementation using first order filters

As a simple application of the above method, a technique for designing first order IIR absorbent filters and an adequate tone corrector will now be described. The tranfer fonction $h_p(z)$ is decomposed as shown in Fig.10, using the zero frequency as a reference:

$$h_p(z) = k_p \,.\, \delta k_p(z) \qquad \text{where} \qquad \delta k_p(z) = \frac{1 - b_p}{1 - b_p.z^{-1}}$$

The first order filter $\delta k_p(z)$ is similar to the one used by Moorer [2] and realizes a low-pass function if $0 \leq b_p < 1$, with unity gain at $\omega = 0$. Here, instead of simulating explicitly the air absorption, we will focus on defining the corresponding $Tr(\omega)$ curve, so that we can fulfill the condition of "continuity" of the pole locus expressed by Eq.(12) and compensate the frequency response using Eq.(15). A remarkable point is that, if the coefficients b_p are small compared to 1, the frequency envelope itself is well approximated by another first order IIR response (see Fig.7), making it easy to correct the frequency response with a first order FIR filter (Fig.9).

The gains k_p are determined as before by the reverberation time at zero frequency:

$$K_p = 20 \log_{10}(k_p) = -60\, \tau_p / Tr(0)$$

The calculations (detailed in annexe) show that each coefficient b_p must be proportional to the corresponding delay length for Eq.(12) to be fulfilled (approximately) for all comb filters. If we use $\alpha = Tr(\pi) / Tr(0)$ as a command parameter to shape the reverberation time versus frequency, we obtain:

$$b_p \approx K_p \,.\, \frac{\ln(10)}{80} \,.\left[1 - \frac{1}{\alpha^2}\right]$$

$$t(z) = \frac{1 - b.z^{-1}}{1 - b} \qquad \text{with} \qquad b \approx \frac{1 - \alpha}{1 + \alpha}$$

These formulas are valid for small values of the coefficients b_p, i.e. for not too small values of α and not too long delays (on Fig.9, the frequency envelope deviates sligthly from a perfectly flat curve). Unfortunately, first order filters provide only two parameters to control the decay characteristics, whereas psychoacoustical studies indicate that the reverberation time should be controllable in at least three frequency bands to cover the full range of perception [8]. However, the general method described here can be used with more complex absorbent filters.

3. GENERALIZATION TO MULTIPLE FEEDBACK DELAY NETWORKS

Improving the "time density"

The above study of the parallel comb filter showed that its strongest limitation lies in the difficulty of obtaining a sufficient *time density* with a reasonable number of unit filters, given that the total delay length determines the maximum *frequency density* one can obtain. As far as the time density is concerned, the series all-pass filter (Fig.2) has a more suitable behaviour: the number of echoes per second is obviously larger and increases along the impulse response (compare Fig.11a and Fig.5), although it becomes more difficult to propose a measure of the time density. This is due to the series association of the unit filters, or, more precisely, to the fact that the output of some delay units is fed into several of them, producing a build-up of the echo density in the time response. A more general approach to multiple feedback delay networks was proposed by Stautner and Puckette [3]. It suggests studying the properties of a general delay network, with the purpose of generalizing the method of absorbent filters for controlling the decay characteristics, and simultaneously obtaining an improved time density with a small number of delay units.

General delay network

The most general single-input, single-output network one can build including N delay units with lengths $\tau_i = m_i.T$ is one where the output signal y(t) is a linear combination of the input signal x(t) and the outputs $q_i(t)$ of all delay units, and so is the input of each delay unit:

$$y(t) = \sum_{i=1}^{N} c_i.q_i(t) + d.x(t) \tag{16}$$

$$q_j(t+m_j) = \sum_{i=1}^{N} a_{ij}.q_i(t) + b_j.x(t) \qquad (\text{for } 1 \leq j \leq N) \tag{17}$$

The resulting network is shown on Fig.12, where, as proposed in [3], the *"feedback matrix"* **A** contains all the feedback gains a_{ij} (from delay unit i to delay unit j). Using z-transform and matrix notation, equations (16) and (17) become, respectively:

$$y(z) = \mathbf{c}^T.\mathbf{q}(z) + d.x(z) \tag{18}$$

$$\mathbf{q}(z) = \mathbf{D}(z).[\,\mathbf{A}.\mathbf{q}(z) + \mathbf{b}.x(z)\,] \tag{19}$$

where:

$$\mathbf{q}(z) = \begin{bmatrix} q_1(z) \\ \vdots \\ q_N(z) \end{bmatrix} \qquad \mathbf{D}(z) = \begin{bmatrix} z^{-m_1} & & 0 \\ & \ddots & \\ 0 & & z^{-m_N} \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_N \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix}$$

The vectors **b** and **c** become matrices for multiple-input and multiple-output systems ($\mathbf{c}^T$ is the transpose of **c**). **D**(z) will be called the *"delay matrix"*. Eliminating **q**(z) in Eqs.(18) and (19) yields the system transfer function:

$$\frac{y(z)}{x(z)} = \mathbf{c}^T.[\,\mathbf{I} - \mathbf{D}(z).\mathbf{A}\,]^{-1}.\mathbf{D}(z).\mathbf{b} + d = H(z) \tag{20}$$

where **I** is the unit matrix. Since the delay matrix **D**(z) is diagonal, we have $[\mathbf{D}(z)]^{-1} = \mathbf{D}(z^{-1})$, and the transfer function can be written:

$$\boxed{H(z) = \mathbf{c}^T.[\,\mathbf{D}(z^{-1}) - \mathbf{A}\,]^{-1}.\mathbf{b} + d} \tag{21}$$

Pole study of the general delay network

The system poles are the solutions of the characteristic equation:

$$\boxed{\det[\,\mathbf{A} - \mathbf{D}(z^{-1})\,] = 0} \tag{22}$$

where $\mathbf{D}(z^{-1})$ is diagonal (the inverse of the delay matrix), and $\mathbf{A}$ is the feedback matrix. It is not simple to solve this equation analitically in the general case, not to mention evaluate the excitations of the eigenmodes. We will focus on two particular classes of feedback matrices: unitary matrices (verifying $\mathbf{A}^T = \mathbf{A}^{-1}$) and triangular matrices (verifying $a_{ij} = 0$ if $i > j$).

In [3], it is shown that the stability of the system can be ensured by choosing for the feedback matrix the product of a unitary matrix and a diagonal matrix whose elements are of magnitude less than 1. A particular four delay network is also studied in detail, with the following feedback matrix, where g is used to control the reverberation time:

$$\mathbf{A} = \frac{g}{\sqrt{2}} \cdot \begin{bmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \qquad (g < 1)$$

It can be shown that choosing a unitary feedback matrix actually forces all the system poles to be on the unit circle. This is observed by writing a state variable formulation of Eq.(19), with a state vector $\tilde{\mathbf{q}}(t)$ composed of all the samples stored in the network's delay units at time t.

If the state variables are defined as follows:

$$\tilde{q}_i^k(t) = q_i(t+k) \qquad (\text{for } 1 \le i \le N \text{ and } 0 \le k \le m_i\text{-}1)$$

Eq.(17) is replaced by:

$$\tilde{q}_i^k(t+1) = \tilde{q}_i^{k+1}(t) \qquad (\text{for } 0 \le k \le m_i\text{-}2)$$

$$\tilde{q}_j^{m_j-1}(t+1) = \sum_{i=1}^{N} a_{ij}.\tilde{q}_i^0(t) + b_j.x(t) \qquad (\text{for } 1 \le j \le N)$$

and Eq.(19) becomes:

$$\tilde{\mathbf{q}}(z) = z^{-1} . [\, \tilde{\mathbf{A}}.\tilde{\mathbf{q}}(z) + \tilde{\mathbf{b}}.x(z)\,] \tag{23}$$

where the state transition matrix $\tilde{\mathbf{A}}$ is a square matrix whose dimension is the sum of the delay lengths and whose eigenvalues are the system poles. The form of this matrix reveals that it is unitary if (and only if) the feedback matrix $\mathbf{A}$ is unitary. Thus if $\mathbf{A}$ is unitary, the system poles are of magnitude 1. This means that the system has only non decaying eigenmodes.

A well-known case is when the feedback matrix $\mathbf{A}$ is diagonal, which brings us back to Schroeder's parallel comb filter. Since the determinant of Eq.(22) is the product of the diagonal terms, the characteristic equation becomes:

$$\prod_{i=1}^{N} (\, a_{ii} - z^{m_i}\,) = 0 \tag{24}$$

More interestingly, this is still valid when $\mathbf{A}$ is triangular (because $\mathbf{A} - \mathbf{D}(z^{-1})$ is also triangular). Thus, if $\mathbf{A}$ is triangular, the poles are the same as those of the parallel comb filter with the feedback gains a_{ii}. The eigenfrequencies are determined only by the selection of delay lengths and each delay unit is associated to a pattern of harmonic resonances. The decay rates of the modes are determined only by the diagonal coefficients in $\mathbf{A}$. The non diagonal coefficients in $\mathbf{A}$, as well as the coefficients b_i and c_i, only determine the excitations of the eigenmodes.

Schroeder's *series all-pass filter* (Fig.2) is itself a network with a triangular feedback matrix (whose diagonal elements are the feedback gains g_i). This reveals that a series all-pass filter has the same eigenmodes (resonant frequencies and decay rates) as the parallel comb filter with the same delay units and feedback gains (compare Fig.11b and Fig.6).

Control of decay characteristics: generalization

For all resonances to have equal decay times, the system poles must have the same magnitude, derived from Eq.(9):

$$\gamma = 10^{-3.T/Tr} \tag{25}$$

When the feedback matrix is triangular, this condition can be easily fulfilled: Eq.(24) shows that we only need to choose the diagonal elements a_{ii} so that:

$$a_{ii} = \gamma^{m_i} = 10^{-3.\tau_i/Tr} \tag{26}$$

This rule applies in particular to the series all-pass filter (Fig.11) and will avoid unnatural resonances in the response to short transients, just as well as with the parallel comb filter, yielding the additionnal benefits of providing an increased time density and a flat frequency response (in Fig.11b, the initial pulse in the time response is cancelled). The main problem with the series all-pass filter, as mentionned in the beginning of this paper, is that its time response tends to become a single pulse when the magnitudes of the feedback gains approach 1 (i.e. when the reverberation time is made long compared to the delay lengths).

However, the method described in part 2 above suggests to control the decay characteristics by modifying the tranfer function of each delay unit instead of modifying the coefficients of the feedback matrix. To generalize this method, we insert a gain k_i at the output of each delay in the general network of Fig.12, respecting the following condition, derived from Eq.(4):

$$k_i = \alpha^{m_i} \tag{27}$$

This is equivalent to replacing $\mathbf{D}(z)$ by $\mathbf{D}(z/\alpha)$ in Eq.(19), or replacing $\mathbf{D}(z^{-1})$ by $\mathbf{D}(\alpha.z^{-1})$ in Eq.(22). The effect is clearly to multiply all the system poles by α. Thus, starting from any *"reference filter"* having poles of equal magnitude, we can modify the reverberation time without violating the principle of equal magnitude of the poles. If the gains k_i are made frequency dependent, the "continuity" of the pole locus is obtained equally well by restricting the condition (27) to a narrow frequency band and making α frequency dependent.

This result allows us to generalize the method of "absorbent filters" to any delay network, as shown in Fig.13, provided that we know a "reference filter" having poles of equal magnitude. It is not a restriction to require that this magnitude should be 1 (which can always be obtained by choosing for α the reverse of the pole magnitude). Any network having a unitary feedback matrix is a reference filter, but this is not a necessary condition since one can also select a triangular feedback matrix with diagonal elements equal to ± 1 (a triangular matrix cannot be unitary unless it is diagonal). It is more convenient to select a reference filter with poles on the unit circle, because then $\alpha(\omega)$ coincides with the magnitude of the poles. In this case the absorbent filters $h_i(z)$ and the tone corrector $t(z)$ in Fig.13 can be designed directly as described in part 2 of this paper, using Eqs.(12) and (15).

The reverberation time can also be modified by multiplying all the delay lengths m_i by the same coefficient α, without modifying the feedback matrix. This is equivalent to replacing $\mathbf{D}(z)$ by $\mathbf{D}(z^{\alpha})$ in Eq.(19), which implies that all eigenfrequencies are devided by α. The magnitude of the poles expressed in decibels is also devided by α, which multiplies the reverberation time by α, as shown by Eq.(25) or Eq.(9).

It is tempting to imagine networks allowing reverberation time control by using a single gain g affecting all the feedback paths. However, we may notice that associating the gain k_i with each delay in the network of Fig.12 is equivalent to multiplying by k_i the corresponding column of the feedback matrix (and the coefficient b_i). Thus, if the feedback matrix is multiplied as a whole by a unique gain g, the conditions expressed by Eq.(27) or (26) cannot be fulfilled. These equations indicate that one needs to associate a separate gain (or absorbent filter) to each delay unit in the network. Otherwise, we strongly suspect that the condition of continuity of the pole locus cannot be fulfilled (certainly not if the feedback matrix is triangular).

Selecting the reference filter

The reference filter completely defines the structure of the reverberator as described on Fig.12 (that is: the number of delay units, the delay lengths, the input, output and feedback paths). Since the poles of a reference filter are of unit magnitude, its impulse response is made of only non decaying eigenmodes. A typical reference filter is one where the feedback matrix is the unit matrix, which corresponds to the parallel comb filter. Although we cannot propose a direct method for selecting the reference filter, a significant improvement can be gained from the parallel comb filter by using a unitary feedback matrix with no null coefficients. This produces a maximum time density with a small number of delay taps. A large variety of matrices can be tried.

Moorer reported that a very natural sounding artificial reverberation effect is obtained by convolving the input signal with an exponentially decaying pseudorandom Gaussian sequence. To test the validity of a reverberant filter structure, the reference filter can be compared to a pseudorandom noise generator: listening to its impulse response allows one to test the selected structure for repetitive patterns in the time response, insufficient time density and insufficient frequency density. If the frequency density is too low, the resulting noise sounds colored ("ringing" tones are heard). In this respect, a crucial factor is the distribution of excitations of the resonances, which should be as even as possible in order to produce a maximum frequency density with shorter delays. The number of delay units determines how fast the time density will build-up before stabilizing. Fig.14 shows the reference filter response in the case of a 12-delay network with a unitary feedback matrix and a total delay time of 1 second, featuring a fast build-up followed by a smooth colorless noise. The choice of delay lengths is not critical if they are kept incommensurate to avoid superpositions of peaks in the time and frequency responses [1]. It is also profitable to select particular matrices allowing simpler calculations (less than N^2 multiply-adds to compute the feedback paths).

4. OBSERVATIONS AND POSSIBLE IMPROVEMENTS

The reverberant filter designed by the above method is only supposed to simulate the final part of the reverberation decay, and produces a somewhat "idealized" reverberation. Of course, it is not intended here to simulate exactly the modal content of a particular room response. The condition of "continuity" of the pole locus should only be considered as a means to obtain a natural sounding reverberation effect with a cost effective delay network (in terms of total delay length and number of delay units), but it may not refer to a feature of real rooms. We suspect that one of the limitations for the general delay network studied here is that it exhibits a uniform *modal density* over the whole frequency range (it certainly does with a triangular feedback matrix). In real rooms, the modal density increases as the square of frequency, although the unevenness of modal excitations may result in a more uniform *frequency density.*

One of the necessary extensions of the model is to allow reproduction of spatial impression by providing several uncorrelated outputs to different channels. This can be done by using directly the outputs of the delay units (as proposed in [3]) or, more generally, by feeding a different combination of these output signals to each channel. A further study is necessary to determine how these combinations should be selected to control the perceived spatial impression, depending on the listening situation (number of channels, earphones or loudspeakers). This entails a more accurate control of early reflections, especially if one wants to simulate multiple or moving sources. Moorer [2] proposed to add a FIR filter in series with the recursive network to simulate the first 80 ms in the monaural impulse response (and increase the echo density). In [3], several early reflections can be assigned to each output channel (by providing multiple input paths to each delay unit). Both of these realizations use a geometric model of sound propagation in the room to evaluate the delays, amplitudes, and spectral contents of the first reflections. Moorer's approach can be extended for stereo headphone reproduction [11] by implementing the binaural head-related transfer function [12] to simulate the direction of incidence of the direct sound and each individual reflection.

Although accurate binaural processing of the direct sound is necessary to simulate the localization of the source [12], it is not clear how much accuracy is required for the delays, amplitudes, directions and spectral contents of early reflections. The main problem with geometric modelling of sound propagation is that it neglects the important contribution of sound *diffusion and diffraction* by walls and obstacles [2], [5]. Psychoacoustical criteria may be helpful to save computer time in controlling and computing the early response.

5. CONCLUSION

In this paper a general method is proposed for designing a multiple feedback delay network simulating the final part of a monaural room response. The "reverberant filter" is designed in two separate steps: 1.Design of a "reference filter" having non decaying eigenmodes and defining the reverberator structure. 2.Design of an "absorbent filter" allowing frequency dependent reverberation time control, with the associated tone corrector to compensate for the modification of the frequency response. The absorbent filter can be selected irrespective of the reference filter, on the basis of perceptive criteria of room acoustic quality. The design of the reference filter requires defining additional perceptive criteria for evaluating the naturalness of artificial reverberation: the "time density" and the "frequency density", which are not usually considered in room acoustics, and still need a more complete investigation.

6. ACKNOWLEDGEMENTS

This work is supported by a CIFRE grant from ANRT. We are grateful to Jean-Pascal JULLIEN for insightful discussions, and to Jean-Michel MERCIER for the development of the graphic software tools used in this paper.

REFERENCES

[1] M.R. SCHROEDER,
"Natural sounding artificial reverberation", J. Audio Eng. Soc., 10(3), 1962.
[2] J.A. MOORER,
"About this reverberation business", Computer Music Journal, 3(2), 1979.
[3] J. STAUTNER, M. PUCKETTE,
"Designing multi-channel reverberators", Computer Music Journal, 6(1), 1982
[4] J. SIKORAV,
"Implementation of reverberators on digital signal processors", Proc. 80th A.E.S. Convention, 1986.
[5] D. GRIESINGER,
"Practical processors and programs for digital reverberation", Proc. 7th A.E.S. Internat. Conf., 1989.
[6] M. BARRON,
"The subjective effects of first reflections in concert halls : the need for lateral reflections",
J. of Sound and Vibration, 15: 475-494, 1971.
[7] J.P. JULLIEN,
"Correlations among objective criteria of room acoustic quality", I.C.A. Toronto 1986, Vol.2, E 4-9.
[8] C. LAVANDIER,
"Validation perceptive d'un modele objectif de caractérisation de la qualité acoustique des salles",
Thèse de doctorat, Université du Maine, Le Mans, Juin 1989.
[9] K. KARPLUS, A. STRONG,
"Digital synthesis of plucked-string and drum timbres", Computer Music Journal 7(2), 1983.
[10] D.A. JAFFE, J.O. SMITH,
"Extension of the Karplus-Strong plucked-string algorithm", Computer Music Journal 7(2), 1983.
[11] F. RICHTER, A. PERSTERER,
"Design and applications of a creative audio processor", Proc. 86th AES Convention (Hamburg), 1989.
[12] J. BLAUERT,
Spatial Hearing, M.I.T. Press, 1983.
[13] M.P. SMYTH,
Linear Engineering Systems, Pergamon press, 1972.

ANNEXE: Reverberation time control using first order absorbent filters.

The tranfer fonction $h_p(z)$ can be decomposed using the zero frequency as a reference, as shown in Fig.10:

$$h_p(z) = k_p . \delta k_p(z) \quad \text{where} \quad \delta k_p(z) = \frac{1 - b_p}{1 - b_p . z^{-1}}$$

The first order filter $\delta k_p(z)$ realizes the frequency dependant gain defined by:

$$\left| \delta k_p(e^{j\omega}) \right|^2 = \frac{(1 - b_p)^2}{1 + b_p^2 - 2.b_p.\cos(\omega)} \tag{a1}$$

which is unitary at $\omega = 0$, and low-pass if $0 \le b_p < 1$. With the gains expressed in decibels, Eq.(12) yields:

$$K_p = 20 \log_{10}(k_p) = -60\, \tau_p / Tr(0) \tag{a2}$$

$$K_p + \Delta K_p(\omega) = -60\, \tau_p / Tr(\omega) = K_p . Tr(0) / Tr(\omega) \tag{a3}$$

Where $\Delta K_p(\omega) = 10 \log_{10} \left| \delta k_p(e^{j\omega}) \right|^2$. It is convenient to write:

$$Tr(\omega) = Tr(0) . \delta tr(\omega)$$

So that $\delta tr(\omega)$ will describe the variation of the frequency envelope in Eq.(14).
Then Eq.(a3) yields:

$$\frac{1}{\delta tr(\omega)} = 1 + \frac{\Delta K_p(\omega)}{K_p} = 1 - \frac{10}{K_p} . \log_{10}\left(\frac{1}{\left| \delta k_p(e^{j\omega}) \right|^2} \right) \tag{a4}$$

Which is nothing but the desired frequency response of the tone corrector t(z) in Eq.(15).
If b_p is small compared to 1, we may use the approximation:

$$\frac{1}{\left| \delta k_p(e^{j\omega}) \right|^2} = 1 + \frac{2.b_p}{(1 - b_p)^2} . (1 - \cos(\omega)) \tag{a5}$$

$$\approx 1 + 2.b_p.(1 + 2.b_p).(1 - \cos(\omega))$$

Replacing in (a4), we obtain:

$$\frac{1}{\left| \delta tr(\omega) \right|^2} \approx 1 - \frac{40}{\ln(10)} . \frac{b_p}{K_p} . (1 - \cos(\omega)) \tag{a6}$$

Eq.(a6) reveals that, if $b_p \ll 1$, there is an approximation of the function $Tr(\omega)$ that is the same for all comb filters, provided that each coefficient b_p is proportional to the associated delay length. This will fulfill (approximately) the condition of "continuity" of the pole locus.
Comparing Eqs.(a5) and (a6), we can see that the curve $Tr(\omega)$ is simply approximated by the frequency response of a first order IIR filter, with the same type of tranfer function as the absorbent filters. This holds for the frequency envelope in Eq.(14), thus we only need a transfer function t(z) of the reverse type (FIR filter) to correct the frequency response:

$$t(z) = \frac{1 - b.z^{-1}}{1 - b} \tag{a7}$$

provided that:

$$\frac{2.b}{(1 - b)^2} \approx \frac{-40}{\ln(10)} . \frac{b_p}{K_p} \approx \frac{1}{2} . \left[\frac{1}{\delta tr(\pi)^2} - 1 \right]$$

which yields formulas for the coefficients b_p and b, using $\alpha = \delta tr(\pi)$ as a command parameter shaping the reverberation time versus frequency:

$$b_p \approx K_p . \frac{\ln(10)}{80} . \left[1 - \frac{1}{\alpha^2} \right] \quad \text{and} \quad b \approx \frac{1 - \alpha}{1 + \alpha} \quad \text{where} \quad \alpha = \frac{Tr(\pi)}{Tr(0)} \tag{a8}$$

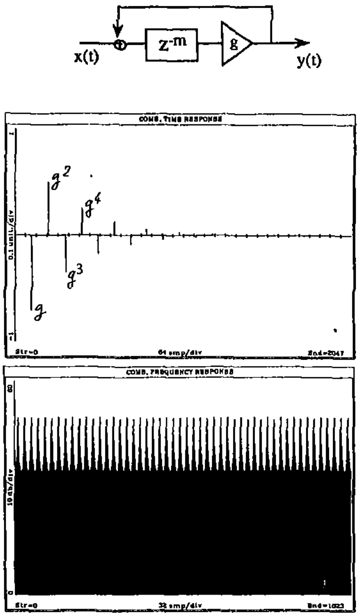

Fig.1a: Comb filter ($g = \frac{-\sqrt{2}}{2}$)

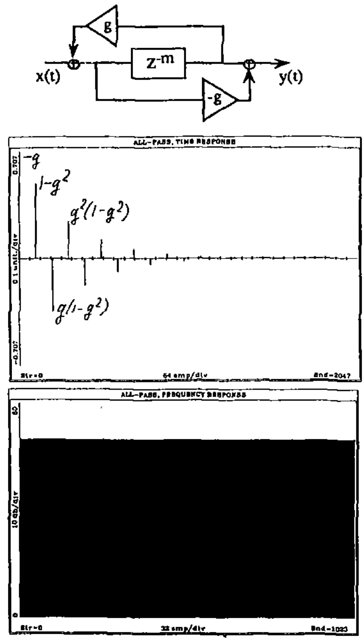

Fig.1b: All-pass filter ($g = \frac{-\sqrt{2}}{2}$)

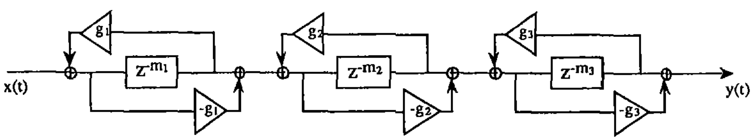

Fig.2: Series all-pass filter

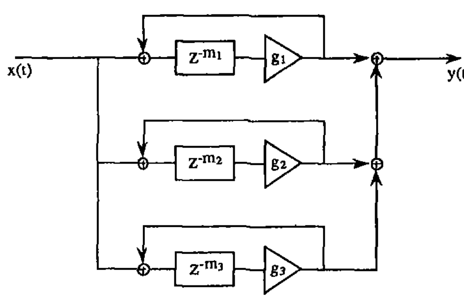

Fig.3: Parallel comb filter

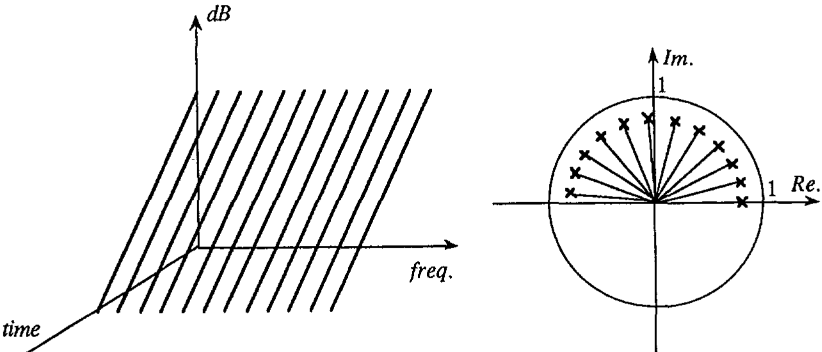

Fig.4a: Pole study of the comb filter

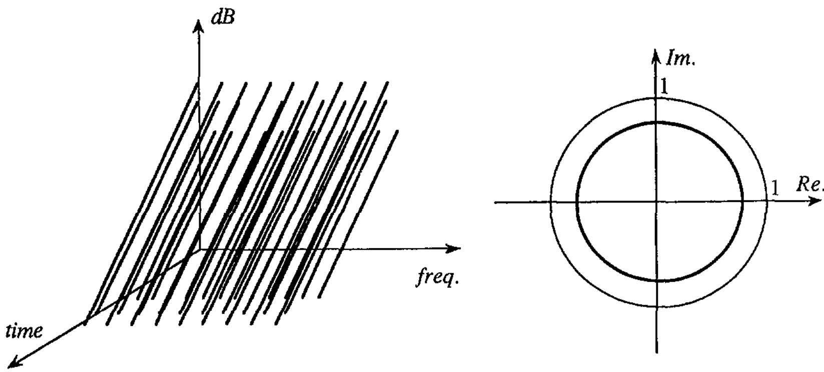

Fig.4b: Pole study of the parallel comb filter. Equal magnitude of the poles.

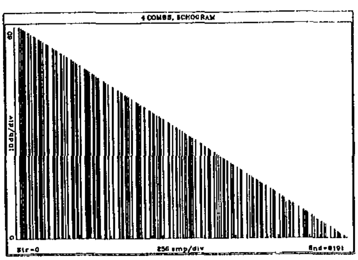

Fig.5: Echogram of parallel comb filter with equal magnitude of the poles.

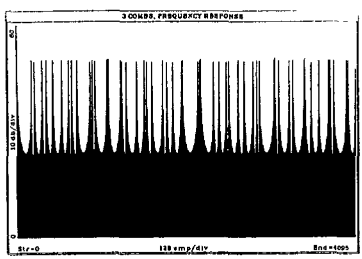

Fig.6: Frequency response of parallel comb filter with equal magnitude of the poles and equal excitations of the eigenmodes.

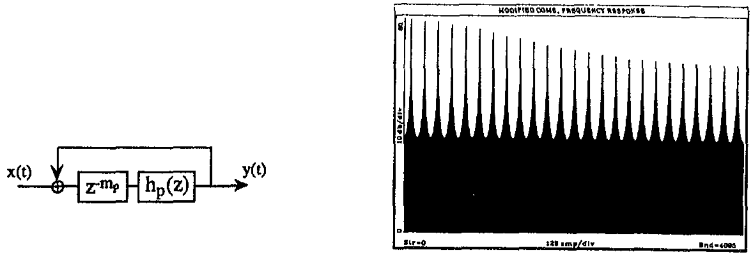

Fig.7: Modified comb filter. Frequency response for $Tr(0)/Tr(\pi) = 4$ ($h_p(z)$ is a first order IIR low pass function implemented as in Fig.10).

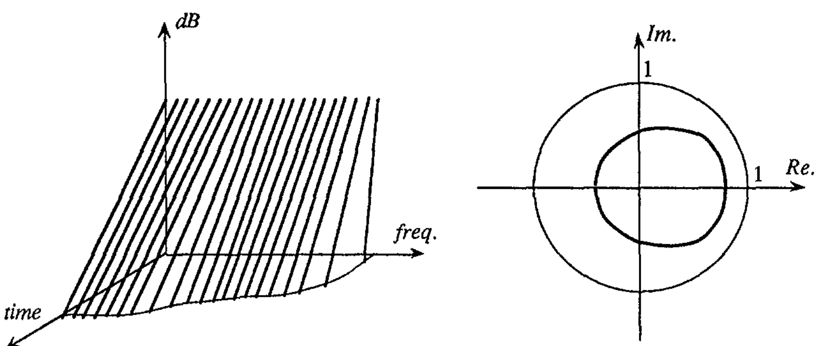

Fig.8: Pole study of comb filter with frequency dependent reverberation time. "Continuity" of the pole locus.

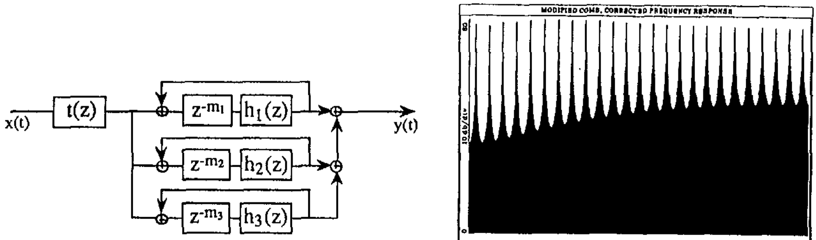

Fig.9: Modified parallel comb filter including absorbent filters $h_p(z)$ and tone corrector $t(z)$. Corrected frequency response for the comb filter of Fig.7.

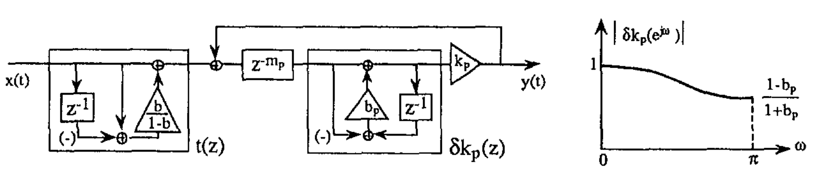

Fig.10: Implementation with first order filters.

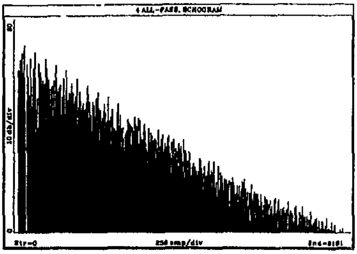

Fig.11a: Echogram of series all-pass filter with equal magnitude of the poles (delay lengths and feedback gains as in Fig.5)

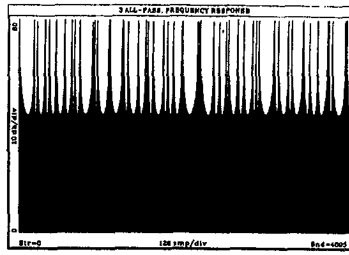

Fig.11b: Frequency response of series all-pass filter (without initial pulse, delay lengths and feedback gains as in Fig.6)

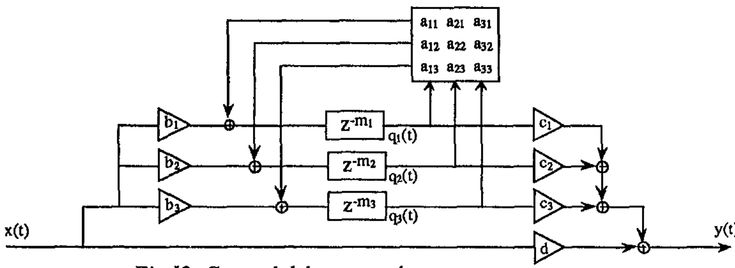

Fig.12: General delay network.

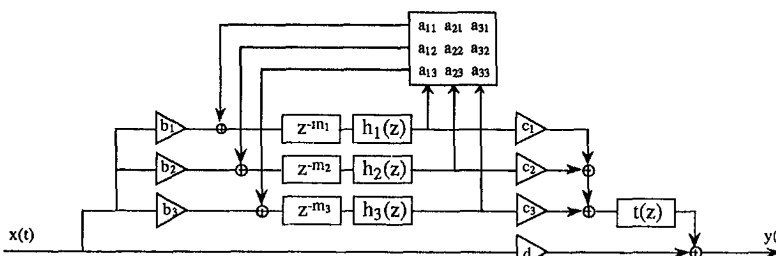

Fig.13: Modified general delay network including absorbent filters $h_i(z)$ and tone corrector $t(z)$.

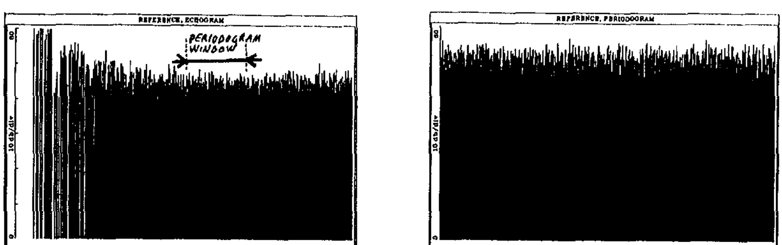

Fig.14: Echogram and periodogram of a reference filter response (12 delay units, unitary feedback matrix, total delay length = 1 s).