Multiscale System Theory

Albert Benveniste, *Fellow, IEEE,* Ramine Nikoukhah, *Member, IEEE,* and Alan S. Willsky, *Fellow, IEEE*

***Abstract*— In many applications it is of interest to analyze and recognize phenomena occurring at different scales. The recently introduced wavelet transforms provide a time-and-scale decomposition of signals that offers the possibility of such an analysis. Until recently, however, there has been no corresponding statistical framework to support the development of optimal, multiscale statistical signal processing algorithms. A recent work of some of the present authors and co-authors proposed such a framework via models of "stochastic fractals" on the dyadic tree. In this paper we investigate some of the fundamental issues that are relevant to system theories on the dyadic tree, both for systems and signals.**

I. Introduction

Multiresolution signal processing has been and continues to be an extremely active area of research in both theory and applications. In part, with the development of the wavelet transform [18]–[20], [25], [29], [30], pyramidal image representation schemes [10], and multirate digital filtering [17], activity in this area has increased dramatically in the past few years. For the most part the home for this research has been the signal processing community, but, in our opinion, there is a significant role that should be played by the system theory community. In particular, in [7] we present an overview of some of our work toward this objective, and we explain why linear models of systems and stochastic processes on the *dyadic tree* provide a natural and powerful setting for multiscale modeling and processing. Roughly speaking, all multiresolution methods involve a process of successive operations of filtering-and-decimation which associate with any signal a collection of successively decimated waveforms. For example orthonormal wavelet and wavelet packets [16] transforms proceed by successively splitting a signal into low- and high-pass components that are decimated before the next stage of filtering. In this way we form the representation of the signal in a wavelet basis in which the different components of the transform are very naturally indexed using a dyadic tree representation. Such a procedure clearly describes a dynamic process in which the index set is the dyadic tree and the basic dynamic operations can be viewed as recursions in *scale,* consisting of filtering and decimation.

Although these observations might be viewed simply as an interesting interpretation of multiresolution signal representations, we believe (and our recent work confirms) that they provide much more than that. Specifically, the description of multiresolution representations as dynamic systems on trees provides a setting for the multiresolution *modeling* of signals and phenomena which, in turn, lead directly to powerful methods for statistically optimal multiresolution signal and image processing. In particular, these observations have led us to examine scale-recursive models for stochastic processes, leading to multiscale generalizations of Schur-Levinson techniques [4] and Kalman filtering [13],[14],[7]. Moreover, in [7],[13]–[15],[26]–[28] we have demonstrated that these and related algorithms for likelihood calculation lead to new methods for a variety of important signal and image processing problems, including multiscale data fusion, motion estimation in image sequences, and texture discrimination. Furthermore, these new methods offer considerable advantages over previous methods in terms of computational efficiency, statistical optimality, explicit calculation of error statistics, estimates at multiple resolutions, and so forth.

In large part it is these successes that provide the motivation for this paper. In particular in the applications we have considered, the multiscale models that have been used were either provided or obvious. However, given the promise of these methods as well as the demonstrated richness of this multiscale framework [27], it is clear that there is a need for a theory for multiscale modeling, and, in particular, for a realization theory for multiscale models. In this paper we introduce and develop the basic ideas for such a multiscale system theory. The key to this development is the identification of a transform concept that is as naturally associated with multiscale systems as the z-transform is with usual discrete-time systems. The basis for this ransform is the definition of elementary "dynamic shift" operators associated with signals on the dyadic tree. That is, while in the usual signals and systems framework the basic operation is the time-shift of signals, usually represented using the variable z, the basic operation for us will be *filtering-and-decimation,* which will play the role of our abstract shifts on the dyadic tree.

In particular, as we develop in Section 2, these abstract shifts are identified by examining the purely algebraic structure of Quadrature Mirror Filter (QMF) banks, which are the key building block in defining pyramidal multiscale signal representations. QMF banks are also central to the construction and design of orthonormal wavelet bases, although in this paper we exclusively use only the basic algebraic properties

Manuscript received September 12, 1991; revised July 1, 1993. This work was supported in part by an INRIA-NSF agreement, as well as by the Air Force Office of Scientific Research under Grant AFOSR-92-5-0002, the National Science Foundation under Grants MIP-9 015 281 and INT-9 002 393, and by Office of Naval Research under Grant No. N00 014-91-J-1004. This paper was recommended by Associate Editor Joseph Nossek.

A. Benveniste is with IRISA-INRIA, Campus de Beaulieu, 35 042 Renes CEDEX, France.

R. Nikoukhah is with INRIA, Rocquencourt, BP 105, 78 153 Le Chesnay CEDEX, France.

A. Wilsky is with LIDS, Massachusetts Institute of Technology, Cambridge, MA 02139.

IEEE Log Number 9214331.

1057-7122/94$04.00 © 1994 IEEE

of QMF banks and not any of the analytic properties used either in filter bank design [31] or wavelet construction [19]. Specifically, we first show that QMF banks can be naturally interpreted as linear operators from $l^2(\mathcal{T})$ into itself, where $\mathcal{T}$ denotes the dyadic tree. These operators satisfy a key algebraic condition that is a direct counterpart of the usual QMF property, and which amounts to defining an orthonormal decomposition of $l^2(\mathcal{T})$. Once we have these basic shift operators we can immediately introduce the concept of a *system* on the tree in exactly the same way that we define linear discrete-time systems in terms of z-transforms. That is, roughly speaking, a system is a linear combination of products of the above mentioned basic shift operators. An important point here is that while for z-transforms we in essence make use of only two operators, namely the forward and backward shift z and z^{-1}, which are inverses of each other, so that $zz^{-1} = z^{-1}z = 1$, for systems on dyadic trees we will need to define *four* operators that satisfy somewhat more complex conditions abstracted from those for QMF's (indeed these operators do not commute and are not invertible). This, in turn, makes the study of such systems much more than a straightforward exercise. In particular, Section 3 is devoted to the study of such systems on the dyadic tree: rational systems, their realizations, and associated state-space forms are introduced.

Following this, the geometry of the dyadic tree is deeply exploited in Section 4 to properly define the concept of a *stationary system* in this context. Stationary systems are defined as systems commuting with translations, where the notion of a "translation" on $\mathcal{T}$ must be carefully defined. In particular, for usual discrete-time systems we typically abuse notation and use the operator z to represent both an operation on signals (i.e. $z(x)(n) = x(n+1)$) *and* as a translation of the index set itself ($z : n \longrightarrow n+1$). For the dyadic tree these two notions are rather different, and thus we must provide a precise notion of the concept of translation. Once we have this definition, a very simple characterization of stationary systems is given, and similar results are presented for stationary stochastic processes, i.e., processes with a covariance that is left invariant by translations (the isotropic processes studied in [4] are thus particular cases of the notion of stationary process we introduce in this paper). In particular, we show that stationary systems driven by white noise produce stationary stochastic processes as outputs, and we provide a "spectral calculus" for such processes. Finally, in Section 5 we summarize the results of the paper and point to several questions for the future.

II. Multiscale Representations And Systems On The Dyadic Tree

The diagram in Fig. 1 depicts a maximally decimated filter bank that still produces alias-free and perfect reconstruction of signals [17],[31],[32]. In this picture, the symbol $\downarrow 2$ denotes the decimation by a factor of two, i.e., the linear map $(x_n) \mapsto (u_n) = (x_{2n})$. Hence, in the z-transform domain, one has $U(z^2) = \frac{1}{2}(X(z) + X(-z))$. Similarly, the symbol $\uparrow 2$ denotes the interpolation by a factor of two, defined in

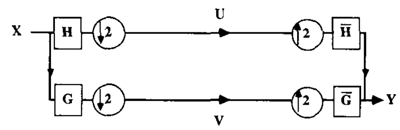


Fig. 1. Maximally decimated filter bank: the $\downarrow 2$ denotes a downsampling operator with rate 2 (samples with even time index are deleted).

the z-transform domain by the map $U(z) \mapsto Y(z) = U(z^2)$. Accordingly, the corresponding input-output map is written as follows:

$$
\begin{aligned}
Y(z) &= \overline{H}(z)U(z^2) + \overline{G}(z)V(z^2) \\
&= \frac{1}{2}\left[\overline{H}(z)H(z) + \overline{G}(z)G(z)\right]X(z) \\
&+ \frac{1}{2}\left[\overline{H}(z)H(-z) + \overline{G}(z)G(-z)\right]X(-z).
\end{aligned} \tag{1}
$$

In this formula, $\frac{1}{2}\left[\overline{H}(z)H(-z) + \overline{G}(z)G(-z)\right]X(-z)$ represents the *aliasing* component, whereas $\frac{1}{2}\left[\overline{H}(z)H(z) + \overline{G}(z)G(z)\right]X(z)$ represents the linear transfer component. For this map to be both alias-free (i.e., $\overline{H}(z)H(-z) + \overline{G}(z)G(-z) = 0$) and providing perfect reconstruction of signals (i.e., $\overline{H}(z)H(z) + \overline{G}(z)G(z) = 2$), a well-known method [31,32,8] is to construct a pair (H, G) satisfying the so-called *power complementary* condition:

$$
\begin{aligned}
H(z)H(z^{-1}) + H(-z)H(-z^{-1}) = 2 \\
z^{-1}H(-z^{-1}) = G(z)
\end{aligned} \tag{2}
$$

and then to select (up to a delay) the pair $(\overline{H}, \overline{G})$ as follows:

$$
\overline{H}(z) = H(z^{-1}), \overline{G}(z) = G(z^{-1}) \tag{3}
$$

This yields the "paraunitary maximally decimated filter banks". Simple examples are given now:

- The pair
$$
H(z) = 1, G(z) = z^{-1} \tag{4}
$$
together with its proper dual $\overline{H}(z) = 1, \overline{G}(z) = z$. This leads to pure sampling-and-interpolation operations (without any filtering) in which one channel keeps the even samples and the other channel the odd samples.
- The pair
$$
H(z) = 1 + z^{-1}, G(z) = 1 - z^{-1} \tag{5}
$$
which corresponds to the Haar transform [19].

A pair (H, G) satisfying the condition (2) is called a *Quadrature Mirror Filter (QMF)* pair.

To make the notion of a *scale* more apparent, we shall slightly modify this classical setting. For ξ real, we denote by $\xi\mathbf{Z}$ the set $\{\xi n | n \in \mathbf{Z}\}$. Successive scales will be figured by the sets $2^{-n}\mathbf{Z}$, where n varies from $-\infty$ (the "coarsest" scales") to $+\infty$ (the "finest" scales"). Then we simply redefine the symbol $\uparrow 2$ as the natural embedding $l^2(2\mathbf{Z}) \hookrightarrow l^2(\mathbf{Z})$ defined by $(x_n) \mapsto (y_n)$, where $y_{2n} = x_n$ and $y_{2n+1} = 0$. Similarly, $\downarrow 2$ is the natural projection $l^2(\mathbf{Z}) \mapsto l^2(2\mathbf{Z})$, which consists of deleting the samples whose index is not even. In doing so, filtering by H and then decimation by a rate of two induces an operator

$$\mathcal{H} : l^2(\mathbf{Z}) \longrightarrow l^2(2\mathbf{Z})$$

and similarly for $\mathcal{G}$. Vice-versa interpolation by a rate of 2 and then filtering by $\overline{H}$ induces an operator

$$\overline{\mathcal{H}} : l^2(2\mathbf{Z}) \longrightarrow l^2(\mathbf{Z})$$

and similarly for $\overline{\mathcal{G}}$. Now, the QMF conditions (2), (3) induce the following identities [19], [8]:

$$\begin{aligned} I &= \mathcal{H}\overline{\mathcal{H}} = \mathcal{G}\overline{\mathcal{G}} \\ I &= \overline{\mathcal{H}}\mathcal{H} + \overline{\mathcal{G}}\mathcal{G} \\ 0 &= \mathcal{H}\overline{\mathcal{G}} = \mathcal{G}\overline{\mathcal{H}} \\ \overline{\mathcal{H}} &= \mathcal{H}^*, \overline{\mathcal{G}} = \mathcal{G}^* \end{aligned} \tag{6}$$

To verify these identities, note first that the second one is equivalent to the perfect reconstruction condition so that it is satisfied. To prove the first or the third ones, we note that elements of $l^2(2\mathbf{Z})$ are obtained via decimation of elements of $l^2(\mathbf{Z})$, i.e., they have z-transforms of the form $U(z^2) = X(z) + X(-z)$ where X is the z-transform of some element of $l^2(\mathbf{Z})$. Thus applying to U the operator $\overline{\mathcal{H}}$ yields, in the z-transform domain corresponding to $l^2(\mathbf{Z})$, $\overline{H}(z)(X(z) + X(-z))$. Then applying $\mathcal{H}$ finally yields, again in the z-transform domain corresponding to $l^2(\mathbf{Z})$

$$\begin{aligned} &\frac{1}{2}[H(z)\overline{H}(z)(X(z) + X(-z)) \\ &+H(-z)\overline{H}(-z)(X(-z) + X(z))] \\ &= \frac{1}{2}[H(z)\overline{H}(z) + H(-z)\overline{H}(-z)](X(z) + X(-z)) \\ &= U(z^2) \end{aligned}$$

where the last equality follows directly from QMF conditions. Other properties are verified in the same way.

Indeed, since decimation and interpolation are generally applied successively several times, it is convenient to consider *all* scales $2^{-n}\mathbf{Z}$. Then for each scale n, filtering by H and then decimation by a rate of 2 induces an operator

$$\mathcal{H} : l^2(2^{-n}\mathbf{Z}) \longrightarrow l^2(2^{-(n-1)}\mathbf{Z})$$

and similarly for $\mathcal{G}$. Vice-versa, for each scale n, interpolation by a rate of two, and then filtering by $\overline{H}$ induces an operator

$$\overline{\mathcal{H}} : l^2(2^{-n}\mathbf{Z}) \longrightarrow l^2(2^{-(n+1)}\mathbf{Z})$$

and similarly for $\overline{\mathcal{G}}$. To avoid using this floating scale n, it is convenient to glue all $2^{-n}\mathbf{Z}$'s together and consider instead their disjoint union $\coprod_{-\infty}^{+\infty} 2^{-n}\mathbf{Z}$. But this latter set is equivalently represented by the *homogeneous dyadic tree* $\mathcal{T}$, defined as follows: the nodes of $\mathcal{T}$ are the truncated binary expansions of real numbers, and $t \rightarrow s$ is a branch of $\mathcal{T}$ if and only if t is obtained via cancelling the last bit in s. Then $2^{-n}\mathbf{Z}$, i.e., the n-th scale, identifies with the set of all binary numbers that are multiples of 2^{-n}. Fig. 4 shows a picture of this dyadic tree. Furthermore, the process of bit cancelling corresponds to

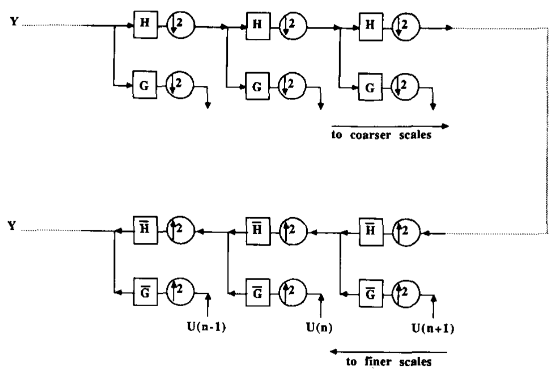


Fig. 2. Cascading a QMF analysis-synthesis bank.

moving to the next coarser scale (i.e. increasing decimation). Hence, the pair $(\mathcal{H}, \mathcal{G})$ is now considered as a pair of operators

$$\mathcal{H}, \mathcal{G} : l^2_{\mathrm{loc}}(\mathcal{T}) \longmapsto l^2_{\mathrm{loc}}(\mathcal{T}) \tag{7}$$

where $l^2_{\mathrm{loc}}(\cdots)$ refers to signals that are locally l^2-summable, and identities (6) are still valid. Now consider a QMF pair $(\mathcal{H}, \mathcal{G})$ as discussed above and let us consider what we could mean by a system involving such a pair and, in particular, how the filter branch of Fig. 2 can be represented as a dynamic object on $\mathcal{T}$. As in the standard Hankel operator approach to linear system theory, we wish to look at input-output maps in which applied inputs finish at some point and we look for outputs "after" this point. In our case, this corresponds to considering the reference scale $\mathbf{Z}$ (rather than a time origin in the usual case) and considering inputs $u \in l^2_{\mathrm{loc}}(\mathcal{T})$ which are zero for scales "after" this one, i.e., for all finer scales, which means that the support of u is contained in the disjoint union $\coprod_{n=-\infty}^{-1} 2^{-n}\mathbf{Z}$. Thus, we can think of u as a collection of signals $U(n)$, where $U(n)$ is the restriction of u at the n-th level of resolution. Thus each $U(n)$ can be interpreted as an ordinary signal and $U(n-1)$ has a sampling rate that is half that of $U(n)$. In Fig. 2, each component $U(n)$ is processed first via a single application of $\overline{\mathcal{G}}$ and then by $(n-1)$ applications of $\overline{\mathcal{H}}$. That is, if we define the following input-output map on $\mathcal{T}$,:

$$y = \left(\overline{\mathcal{G}} + \overline{\mathcal{H}}\overline{\mathcal{G}} + \overline{\mathcal{H}}^2\overline{\mathcal{G}} + \overline{\mathcal{H}}^3\overline{\mathcal{G}} + \ldots\right)u = \left(1 - \overline{\mathcal{H}}\right)^{-1}\overline{\mathcal{G}}u \tag{8}$$

then the desired signal Y is simply the restriction of y to the zero scale. Thus we see that the QMF reconstruction procedure can be thought of *exactly* as a particular Hankel map on $l^2(\mathcal{T})$ that is both causal with respect to scale (i.e., coarse-to-fine) and, as indicated in (8), rational. A similar but dual interpretation holds for the pair $(\mathcal{H}, \mathcal{G})$, which defines an anticausal (i.e., fine-to-coarse) system on $\mathcal{T}$. In the next section we present a precise framework for describing such dynamic systems and for studying the concepts of rationality and realizability.

III. System Theory and Realizations

3.2 Systems on the Tree

In this section we consider some "abstract" QMF pair, i.e., a 4-tuple of operators $\{\alpha, \beta, \overline{\alpha}, \overline{\beta}\}$ on $l^2(\mathcal{T})$ satisfying the following QMF properties that are just the translation of properties (6):

$$\alpha\overline{\alpha} = \beta\overline{\beta} = 1 \tag{9}$$

$$\overline{\alpha}\alpha + \overline{\beta}\beta = 1 \tag{10}$$

$$\beta\overline{\alpha} = \alpha\overline{\beta} = 0. \tag{11}$$

The class of operators we consider is the multiplicative algebra of linear combinations of these primitive ones: this is a noncommutative algebra. The class of abstract systems we consider are matrices whose entries are elements of this algebra. Thanks to these rules, any system can be expressed as follows:

$$H = \sum_{\substack{w^\uparrow \in \mathcal{W}^\uparrow \\ w^\downarrow \in \mathcal{W}^\downarrow}} h_{w^\uparrow w^\downarrow} w^\uparrow w^\downarrow \tag{12}$$

where

$$\mathcal{W}^\uparrow = \left\{\overline{\alpha}^{i_1}\overline{\beta}^{i_2} | i_1, i_2 \in \mathbf{N}\right\}, \mathcal{W}^\downarrow = \left\{\alpha^{j_1}\beta^{j_2} | j_1, j_2 \in \mathbf{N}\right\}$$

are the family of monomials generated by the operators $\overline{\alpha}, \overline{\beta}$ and α, β, respectively, and the $h_{w^\uparrow w^\downarrow}$'s are matrix coefficients. In this writing we implicitly assume that all simplifications (9), (10), and (11) have been performed. We shall call the *support* of H the set of monomials in (12) with nonzero coefficient and shall call the *degree* of the monomial $w^\uparrow$ or $w^\downarrow$ the length of the considered word, i.e., degree$(w^\uparrow)$ = degree$(\overline{\alpha}^{i_1}\overline{\beta}^{i_2}) = i_1 + i_2$, and similarly for $w^\downarrow$.

Examples of Abstract QMF Pairs on the Tree: We already discussed the case of abstract QMF pairs $(\mathcal{H}, \mathcal{G})$ originating from some actual QMF filter bank (H, G) in the classical setting. Of particular interest in the sequel will be the simplest such case in which $H(z) = 1$ and $G(z) = z^{-1}$,: this corresponds to performing pure decimation and interpolation without any filtering, and will be widely studied in Section 4. The particular feature of QMF pairs $(\mathcal{H}, \mathcal{G})$ originating from some actual QMF filter bank (H, G) is that they map signals with support contained in a given scale into signals with support contained in the next scale. But this does not need to be the case in general for our abstract QMF pairs,: in particular, it may be the case that, for $t \in \mathcal{T}$, $(\alpha x)_t$ is a linear combination of x_s's for some nodes s located in *different* scales.

Examples of Systems on the Tree. Referring to our discussion in (8), the system $(1 - \overline{\mathcal{H}})^{-1}\overline{\mathcal{G}}$ considered there can be put in the generic form (12) by setting

$$\begin{aligned} &\alpha = \mathcal{H}, \beta = \mathcal{G}, \overline{\alpha} = \overline{\mathcal{H}}, \overline{\beta} = \overline{\mathcal{G}} \\ &\forall n \geq 0 : h_{\overline{\alpha}^n\overline{\beta}} = 1, \\ &\text{otherwise } h_{w^\uparrow w^\downarrow} = 0. \end{aligned} \tag{13}$$

When the underlying QMF pair (H, G) of filters is low-pass/high-pass, this system corresponds to a filter bank with logarithmically equal frequency bands. Similarly, the choice

$$\alpha = \mathcal{H}, \beta = \mathcal{G}, \overline{\alpha} = \overline{\mathcal{H}}, \overline{\beta} = \overline{\mathcal{G}} \tag{14}$$

$\forall n \geq 0 : h_{w^\uparrow} = 1$ if $w^\uparrow$ is of degree n, otherwise $h_{w^\uparrow w^\downarrow} = 0$

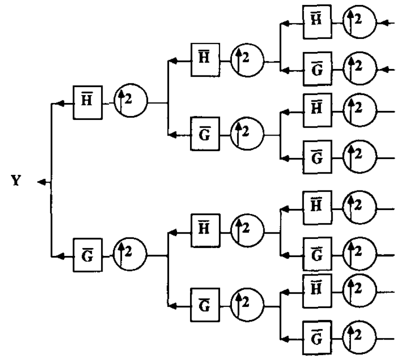


Fig. 3. A wavelet packet synthesis bank.

which is shown in Fig. 3, corresponds to the case of a filter bank with equal frequency bands.

Causality. We shall say that a monomial $w^\uparrow w^\downarrow$ is *causal* if

$$\text{degree}(w^\uparrow) \geq \text{degree}(w^\downarrow) \tag{15}$$

and we say that the system H is *causal* if, in expression (12), $h_{w^\uparrow w^\downarrow} = 0$ whenever $w^\uparrow w^\downarrow$ is noncausal. Strict causality is defined accordingly. Referring to the particular case where our pair $\{\alpha, \beta\}$ corresponds to a usual QMF pair, we see that causal monomials involve only decimation. Hence, in this case, causal systems are just systems such that their output at a particular level of resolution depends only on the values of their input at the same or at coarser levels of resolution: *causality thus may be interpreted as referring to a dynamics oriented from coarse to fine resolutions.* In particular, the synthesis QMF bank of Fig. 2 specifies a (purely) causal system, while the analysis QMF bank specifies a (purely) anticausal one. Causal systems may be written as follows:

$$H = \sum_{\substack{w^\uparrow \in \mathcal{W}^\uparrow \\ \vec{w} \in \vec{\mathcal{W}}}} h_{w^\uparrow \vec{w}} w^\uparrow \vec{w} \tag{16}$$

where $\vec{\mathcal{W}}$ is the set of monomials $w^\uparrow w^\downarrow$ such that

$$\text{degree}(w^\uparrow) = \text{degree}(w^\downarrow).$$

Referring again to the synthesis QMF bank of Fig. 2, the $h_{w^\uparrow \vec{w}}$ coefficients are zero unless $\vec{w}$ is the empty word: such a system will be referred to as "zero-depth," where the notion of depth is captured in the definition to follow. In Section 4 we consider a particular concrete set of variables $\alpha, \beta, \overline{\alpha}, \overline{\beta}$ that are the correct ones for the study of the notion of stationarity on $\mathcal{T}$ and we will find that the correct class of stationary models of the form of (16) will include those that are not restricted to be zero depth but are in fact finite depth.

Definition: We define the **depth** of a causal monomial $w = w^\uparrow \vec{w}$ (cf. formula (16)) as one-half the degree of $\vec{w}$. A system H is called **finite depth** if it can be expressed as a (possibly infinite) sum of monomials with uniformly bounded depth. The **depth** of H is the minimum among such bounds over all possible representations of H.

At this point, it will be useful to provide a representation of the space of all monomials of depth $\leq k$. Consider the space

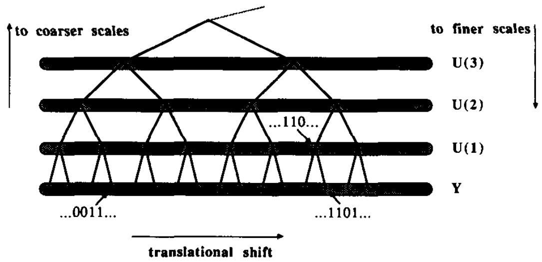


Fig. 4. The dyadic homogeneous tree. The various scales, i.e., $2^n\mathbf{Z}$, are vizualized as the horizontal alignments of nodes. Moving one step upward corresponds to moving to the next coarser scale, i.e., removing the least bit in the binary coding of the node. The Y and $U(n)$ signals are also visualized for $n = 1, 2, 3$.

$\vec{\mathcal{W}}_k$ spanned by the monomials $\vec{w}$ of degree $\leq 2k$. Recall that $\overline{\alpha}\alpha + \overline{\beta}\beta = 1$ so that the family of these monomials is not a basis of $\vec{\mathcal{W}}_k$. However, it is easily checked that monomials with a degree *exactly* equal to $2k$ form a basis for $\vec{\mathcal{W}}_k$. Denote by $\{\phi_1, ..., \phi_{n_k}\}$ such a basis and set

$$\underline{\phi}_k = \begin{bmatrix} \phi_1 \\ \cdot \\ \cdot \\ \cdot \\ \phi_{n_k} \end{bmatrix}, \Phi_k = \mathrm{diag}(\underline{\phi}_k, \underline{\phi}_k, \cdots, \underline{\phi}_k).$$

If H is finite depth we can decompose it as follows:

$$H = H^{\uparrow}\vec{H} \tag{17}$$

where $H^{\uparrow}$ is a system with support in $\mathcal{W}^{\uparrow}$ and $\vec{H}$ a *finite degree* system with support in $\vec{\mathcal{W}}$. This is shown as follows. Consider the form (16) for the causal systems. Write each monomial $\vec{w}$ using the above basis: $\vec{w} = M_{\vec{w}}\Phi_k$ for a suitable row matrix $M_{\vec{w}}$, then H is rewritten as

$$\begin{aligned} H &= \sum_{\substack{w^{\uparrow} \in \mathcal{W}^{\uparrow} \\ \vec{w} \in \vec{\mathcal{W}}}} h_{w^{\uparrow}\vec{w}} w^{\uparrow}\vec{w} \\ &= \left(\sum_{\substack{w^{\uparrow} \in \mathcal{W}^{\uparrow} \\ \vec{w} \in \vec{\mathcal{W}}}} h_{w^{\uparrow}\vec{w}} M_{\vec{w}} w^{\uparrow} \right) \Phi_k \\ &= H^{\uparrow}\Phi_k. \end{aligned}$$

Next, consider two such decompositions

$$H = H_1^{\uparrow}\Phi_k = H_2^{\uparrow}\Phi_k.$$

Since monomials $\{\phi_1, ..., \phi_{n_k}\}$, $\overline{\alpha}^i$, and $\overline{\beta}^j$ form a free system, it follows that

$$H_1^{\uparrow} = H_2^{\uparrow}$$

must hold in the above decomposition. Thus, for any k such that H of finite depth can be decomposed into

$$H = H^{\uparrow}\Phi_k \tag{18}$$

where the system $H^{\uparrow}$ is unique and has a depth of zero. Thus we shall call the *depth* of H the minimum integer k for which decomposition (18) holds.

Note that there is a deep reason for restricting ourselves to finite rather than infinite depth systems. In particular, we are interested in defining the class of *rational* systems as in (16). It is well known in language and automata theory that the language $\{a^n b^n | n \geq 0\}$ is not rational, i.e., it cannot be generated by a finite-state automaton. However, any finite restriction of this language, e.g., $\{a^n b^n | 0 \leq n \leq N\}$ *is* rational. The language $\vec{\mathcal{W}}$ is exactly of this form (with more letters than simply a and b, namely $\{\alpha, \beta\}$ and $\{\overline{\alpha}, \overline{\beta}\}$), and this rationality demands the restriction to the finite depth case.

3.2 Realizations

We now investigate some aspects of a system theory for the notion of a system introduced above. We shall see that the theory of general systems is related to realization theory for automata [22,9] rather than linear system theory even though we are considering linear operators on signals.

Definition: A finite depth system H as in (17) is *realizable* if there exist constant matrices C, A_α, A_β and a system $\vec{H}$ as in (17) such that

$$H = C\left(I - \overline{\alpha}A_\alpha - \overline{\beta}A_\beta\right)^{-1}\vec{H}. \tag{19}$$

A *state-space realization* of (19) is

$$\begin{cases} x &= \overline{\alpha}A_\alpha x + \overline{\beta}A_\beta x + \vec{H}u \\ y &= Cx. \end{cases} \tag{20}$$

Realization in the Zero Depth Case: According to (17), a zero depth system may be expressed as

$$H = \sum h_{w^{\uparrow}} w^{\uparrow}.$$

As is usually done in automata and noncommutative formal power series theories, we associate with H the following *Hankel matrix:*

$$\mathrm{Hank}(H)_{ij} = h_{w_i^{\uparrow} w_j^{\uparrow}}$$

where the monomials $\left(w_i^{\uparrow}\right)_{i \geq 0}$ are ordered according to the increasing degree with priority given to $\overline{\alpha}$. Note that this Hankel matrix is just the matrix representation of the Hankel operator we discussed in Section 2. Then the following results may be borrowed from noncommutative formal power series theory [22], [9].

Theorem: H is realizable if and only if $\mathrm{Hank}(H)$ has finite rank. Moreover, the dimension of minimal realizations equals this rank, i.e.,

$$H = C\left(I - \overline{\alpha}A_\alpha - \overline{\beta}A_\beta\right)^{-1}B$$

where the dimensions of A_α and A_β equals the rank of $\mathrm{Hank}(H)$.

By writing

$$A_w = \text{coefficient of } w \text{ in } \left(I - \overline{\alpha}A_\alpha - \overline{\beta}A_\beta\right)^{-1}$$

we also have the following theorem.

Theorem: A realization $(C, A_\alpha, A_\beta, B)$ is minimal if and only if

$$\bigvee_{|w|<n} \mathrm{Im}(A_w B) = \mathbf{R}^n$$

$$\bigcap_{|w|<n} \mathrm{Ker}(C A_w) = \{0\}$$

where n is the dimension of the state and where $|w|$ denotes the total degree of w.

As a corollary, we know that all minimal realizations are related by similarity transformations. To conclude, the realization theory for the zero-depth case has been tied back to the classical theory of noncommutative formal power series, which is in contrast to the realization theories for classical $1D$- and $2D$- linear systems.

Realization in the Finite Depth Case: The above procedure has to be modified for this case.

Theorem: Consider again decomposition (18) where $k = k_o$ is taken to be the depth of H:

1) H is realizable if and only if $H^\uparrow$ in (18) is realizable.
2) If $(C, A_\alpha, A_\beta, B)$ is a minimal realization of $H^\uparrow$ then $(C, A_\alpha, A_\beta, B\Phi_k)$ is a minimal realization of H.

Proof: We just have to prove the second statement. For this consider two decompositions

$$H = H_1^\uparrow \Phi_{k_o} = H_2^\uparrow \Phi_k$$

where k_o is the depth and $k \geq k_o$. Since k_o is minimal and the diagonal elements of Φ_k form a basis of $\mathcal{W}_k$, there exists a surjective constant matrix G_1 such that $G_1 \Phi_k = \Phi_{k_o}$. Since the left factor in the decomposition (18) is unique for k fixed, we must have

$$H_1^\uparrow G_1 = H_2^\uparrow.$$

Assume the conclusion of the theorem is false, i.e., there exists a realization $(C_2, A_{2,\alpha}, A_{2,\beta}, B_2)$ of $H_2^\uparrow$ of degree less than the minimal realization of $H_1^\uparrow$. But in this case $(C_2, A_{2,\alpha}, A_{2,\beta}, B_2 G_1^\dagger)$ is a realization of $H_1^\uparrow$ where $G_1^\dagger$ is any right inverse of G_1, which is a contradiction. □

The realization procedure for the k-depth case is:

1) Express H as

$$H = H^\uparrow \Phi_k.$$

2) Realize $H^\uparrow$ as

$$H^\uparrow = C\left(I - \overline{\alpha} A_\alpha - \overline{\beta} A_\beta\right)^{-1} B. \tag{21}$$

IV. Stationary Causal And Noncausal Systems And Stochastic Processes

In this section we investigate the notion of stationarity. An important objective of this investigation is to introduce a notion of stationarity for both systems and stochastic processes with the desirable property that *the output of a stationary system driven by a stationary stochastic process be itself a stationary stochastic process.* The usual notion of stationarity for systems indexed by integers is the following: translating the input would provide a translated version of the output. This is the notion of stationarity we shall consider here consequently we must first understand what is a translation on the tree, and what are isometries.

4.1 The Homogeneous Tree and its Geometry

Homogeneous trees, and their structure, have been the subject of some work [1], [2], [12], [21], [11] in the past on which we build and which we now briefly review. A *homogeneous tree $\mathcal{T}$ of order q* is an infinite acyclic, undirected, connected graph such that every node of $\mathcal{T}$ has exactly $(q+1)$ branches. Note that $q = 1$ corresponds to the usual integers with the obvious branches from one integer to its two neighbors. The case of $q = 2$, illustrated in Figs. 4 and 5, corresponds, as we will see, to the dyadic tree on which we focus in this paper. In 2-D signal processing, it would be natural to consider the case of $q = 4$ leading to a pyramidal structure on the indexing set of the 2-D processes.

Isometries: The tree $\mathcal{T}$ has a natural notion of distance: $d(s,t)$ is the number of branches along the shortest path between the nodes $s, t \in \mathcal{T}$ (by abuse of notation we use $\mathcal{T}$ to denote both the tree and its collection of nodes). One can then define the notion of an isometry on $\mathcal{T}$, which is simply a one-to-one map of $\mathcal{T}$ onto itself that preserves distance. For the case of $q = 1$, the group of all possible isometries corresponds to translations of the integers $(t \mapsto t + k)$, the reflection operation $(t \mapsto -t)$, and concatenations of the two. For $q \geq 2$ the group of isometries of $\mathcal{T}$ is significantly larger and more complex. The following classification of isometries may be found in [12], see Appendix A for a proof and related lemmas on the geometry of the homogeneous tree:

Lemma Classification of Isometries: Given an isometry f of the homogeneous tree $\mathcal{T}$, three cases are possible, namely

$$\exists s \in \mathcal{T} : f(s) = s \tag{22}$$

$$\exists s, t \in \mathcal{T} : d(s,t) = 1 \text{ and } f(s) = t, f(t) = s \tag{23}$$

$$\exists (s_n)_{n \in \mathbf{Z}} \in \mathcal{T}, \exists i > 0 : d(s_n, s_{n+1}) = 1 \text{ and } f(s_n) = s_{n+i}. \tag{24}$$

For obvious reasons, isometries of type (24) will be called *translations.*

Boundary Points and Horocycles: An important concept here is the notion of a *boundary point* [2], [11] of a tree. Consider the set of infinite sequences of $\mathcal{T}$ where any such sequence consists of a sequence of distinct nodes $t_1, t_2, \ldots$ where $d(t_i, t_{i+1}) = 1$. A boundary point is an equivalence class of such sequences where two sequences are equivalent if they differ by a finite number of nodes. For $q = 1$, there are only two such boundary points corresponding to sequences increasing toward $+\infty$ or decreasing toward $-\infty$. For $q = 2$ the set of boundary points is uncountable. In this case, let us choose one boundary point, which we denote by $-\infty$.

Once we have distinguished this boundary point, we can identify a partial order on $\mathcal{T}$. In particular, note that from any node t there is a unique path in the equivalence class defined by $-\infty$ (i.e.\ a unique path from t "toward" $-\infty$). Then if we take any two nodes s and t, their paths to $-\infty$ must differ

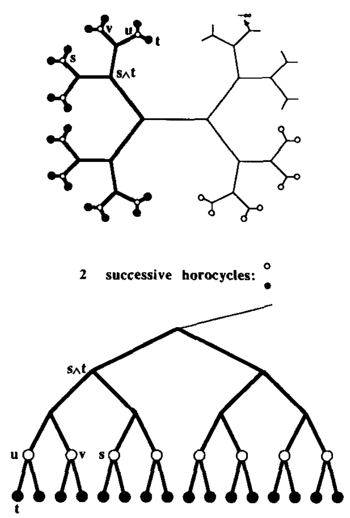


Fig. 5. The dyadic homogeneous tree: showing two ways of looking at it.

only by a finite number of points and thus must meet at some node, which we denote by $s \wedge t$ (see Fig. 5). Thus, we can define a notion of *relative distance* of two nodes to $-\infty$

$$\delta(s,t) = d(s, s \wedge t) - d(t, s \wedge t) \tag{25}$$

so that

$s \preceq t$ ("s is at least as close to $-\infty$ as t") if $\delta(s,t) \leq 0$

$s \prec t$ ("s is closer to $-\infty$ than t") if $\delta(s,t) < 0$.

This also yields an equivalence relation on nodes of $\mathcal{T}$

$$s \asymp t \leftrightarrow \delta(s,t) = 0$$

For example, the points s, v, and u in Fig. 5 are all equivalent. The equivalence classes of such nodes are referred to as *horocycles*. In this case the horocycles appear as points on the same horizontal level and $s \preceq t$ means that s lies on a horizontal level above or at the level of t. Note that in this way we make explicit the dyadic structure of the tree. With regard to multiscale signal representations, a shift on the tree toward $-\infty$ corresponds to a shift from a finer to a coarser scale and points on the same horocycle correspond to the points at different translational shifts in the signal representation at a single scale.

Translations and Primitive Translations. In the remainder of this article, some $-\infty$ is selected and fixed: *only translations associated with paths originating from* $-\infty$ *will be considered.* Translations will play an important role in the definition of stationarity. Translations certainly are the isometries of the third class (see (24)) according to the classification of lemma 1. However, for the sequel, we shall need *primitive translations* encoding "moving away from $-\infty$", i.e. the counterpart of the shift operator z on $\mathbf{Z}$. These are defined as follows:

1) Select an infinite path $(t_n)_{n \in \mathbf{Z}}$ originating from $-\infty$, call it the *skeleton* of the primitive translation.

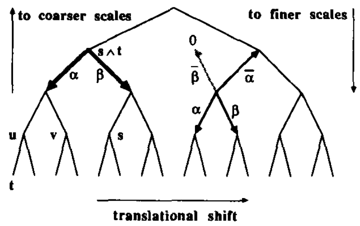


Fig. 6. Showing shifts: very thick lines show the moves on the tree, and thick lines show the operators on signals (the value at the origin of each arrow is picked at the corresponding end).

2) Denote by s_n the unique point outside the skeleton such that $d(s_n, t_n) = 1$.
3) Denote by $\mathcal{T}^+_{s_n}$ the semiinfinite dyadic tree with root s_n composed of the semiinfinite paths originating at s_n and moving away from $-\infty$.
4) Then the *primitive translation with skeleton* (t_n) is the unique isometry τ such that (see Fig. 7)

$$\tau(t_n) = t_{n+1}, \tau(\mathcal{T}^+_{s_n}) = \mathcal{T}^+_{s_{n+1}}. \tag{26}$$

4.2 Shifts on $\mathcal{T}$

We shall call *shifts* the most elementary pair of $\{\alpha, \beta\}$ operators; shifts on the tree will be the counterpart of the "z" shift for the usual case of systems indexed by integers. Indeed, these elementary shifts are obtained by considering the simplest QMF pair (H, F) as in Fig. 1, namely $(H, G) = (1, z^{-1})$ [for an obvious QMF pair see (4)], and then by applying the construction of Section 2 to get the corresponding abstract QMF pair on the tree. We introduce now these shift operators in a detailed way. In Fig. 6, two shifts are first described that act on the nodes of the tree:

- 1 is the identity operator (no move)
- α is the left down-shift (move one step away from $-\infty$ toward the left)
- β is the right down-shift (move one step away from $-\infty$ toward the right)

These shifts act on the right (if t is any node on the tree, $t\alpha$ is its left offspring). Note that α and β are one-to-one but not onto; they are *not* isometries. From these shifts on nodes we can derive shift operators on signals. By "signal" we mean a family y_t of scalars or vectors indexed by the vertices of the tree. The primitive operators that we consider are "dual" of the shifts on $\mathcal{T}$, namely (see Fig. 6):

- 1 is the identity operator (no move)
- α is the left down-shift operator:[1]

$$y = \alpha u \Leftrightarrow \forall t : y_t = u_{t\alpha}$$

- β is the right down-shift operator:

$$y = \beta u \Leftrightarrow \forall t : y_t = u_{t\beta}$$

[1]The value of y at a given node is obtained by picking the value of u at the corresponding *left down* node.

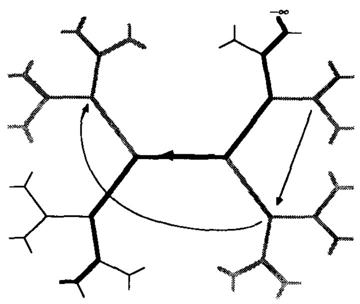


Fig. 7. Translations: we show how the $\mathcal{T}^+_{s_n}$ (in grey) are succesively mapped

- $\overline{\alpha}$ is the right up-shift operator:[2]

$$y = \overline{\alpha} u \Leftrightarrow \forall t : \begin{cases} y_{t\alpha} &= u_t \\ y_{t\beta} &= 0 \end{cases}$$

- $\overline{\beta}$ is the left up-shift operator:

$$y = \overline{\beta} u \Leftrightarrow \forall t : \begin{cases} y_{t\beta} &= u_t \\ y_{t\alpha} &= 0 \end{cases}$$

It is easy to verify that these elementary shifts satisfy the QMF relations (11) and correspond to choice (4) for QMF pairs. When state-space forms for systems are considered with *these* particular shifts following the preceding section, the moves on the tree that we just introduced can be used to rewrite these state-space forms as follows:

$$\begin{cases} x_{t\alpha} &= A_\alpha x_t + \alpha \vec{H} u_t \\ x_{t\beta} &= A_\beta x_t + \beta \vec{H} u_t \\ y_t &= C x_t \end{cases} \tag{27}$$

and

$$\begin{cases} x_{t\alpha} &= A_\alpha x_t + \alpha B \Phi_k u_t \\ x_{t\beta} &= A_\beta x_t + \beta B \Phi_k u_t \\ y_t &= C x_t \end{cases} \tag{28}$$

would replace (20) and (21), respectively. Such notations will be heavily used in the remainder of the paper.

It is clear that these elementary shifts cannot be considered as "stationary" in any reasonable sense. For instance, the relation $y = \overline{\alpha} u$ where $u \equiv 1$ yields $y_{t\alpha} = 1$ but $y_{t\beta} = 0$. This means that, to develop a theory of stationary systems and processes, we need to constrain the class of systems that we have considered so far. This will be the subject of the next section.

4.3 Characterization of Stationary Systems

Throughout the remainder of this section, the symbols $(\alpha, \beta, \overline{\alpha}, \overline{\beta})$ will denote the *specific* shift operators we introduced in the preceding subsection. Given a translation τ of $\mathcal{T}$, by abuse of notation, we also denote by τ its action on signals defined by

$$\tau(y)_t = y_{\tau(t)}.$$

[2]The value of y at a given node is obtained by picking the value of u at the corresponding *right up* node if available, or by setting 0 otherwise.

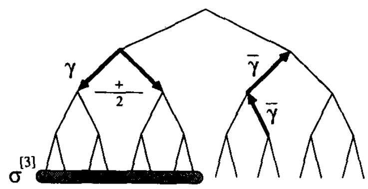


Fig. 8. Shifts for stationary transfer functions: the value at the origin of each arrow is picked at the corresponding end and the grey cigar replaces each value by the corresponding average.

Definition Stationary Systems: A linear operator H acting on signals is said to be stationary if[3]

$$H \circ \tau = \tau \circ H \tag{29}$$

holds for any translation τ.

The following fundamental result is proved in Appendix B.

Theorem: Let H be a linear operator acting on signals.

1) If H satisfies (29) for any primitive translation τ, then it must be a system of the form

$$H = \sum_{i,j \geq 0} h_{i,j} \overline{\gamma}^i \gamma^j \tag{30}$$

where

$$\gamma = \frac{1}{2}(\alpha + \beta), \overline{\gamma} = \overline{\alpha} + \overline{\beta}. \tag{31}$$

2) Conversely, any H of the form (30) is stationary, i.e., satisfies (29) for any translation τ.

These two operators generate two semigroups. The action of these semigroups is depicted in Fig. 8: $\overline{\gamma}$ is a "backward" shift toward $-\infty$, whereas γ is a "forward-and-average" shift (the "Haar smoother"). In fact, the pair $\{\frac{1}{2}(\alpha+\beta), \frac{1}{2}(\alpha-\beta)\}$ form a QMF pair with adjoint operators given by $\{\overline{\alpha} + \overline{\beta}, \overline{\alpha} - \overline{\beta}\}$ respectively: this is just the QMF pair associated with the Haar transform. In particular, the γ and $\overline{\gamma}$ operators obey the following simplification rule

$$\gamma \overline{\gamma} = 1 \tag{32}$$

To encode causal stationary systems, it will be useful to introduce the following family of operators, which perform a smoothing of data on the same horocycle as shown in the Fig. 8:

$$\sigma^{[i]} = \overline{\gamma}^i \gamma^i \tag{33}$$

All $\sigma^{[i]}$'s are *idempotent* operators. These operators may be used to provide the following counterpart of formula (16) for the stationary case.

Theorem: If H is stationary and causal, it can be expressed as follows:

$$H = \sum_{i,j \geq 0} h_{i,j} \overline{\gamma}^i \sigma^{[j]} \tag{34}$$

Obviously, the matrix coefficients $h_{i,j}$ are different in formulae (30) and (34).

[3]$\circ$ denotes the composition of maps.

4.4 Realization of Stationary Systems

Both (30) and (34) may be interpreted as standard $2D$-systems that are *causal in the two variables.* Hence, standard $2D$ realization theories may be applied to both cases. We shall briefly investigate the two cases.

Noncausal Systems: If we interpret γ as the row operator and $\overline{\gamma}$ as the column operator, then it is natural to consider the row-by-row scanning to define a total ordering on the $2D$ index space. This corresponds to decomposing the system H according to the following two steps:

1) A bottom-up (i.e. fine-to-coarse) smoothing, followed by
2) A top-down (i.e. coarse-to-fine) propagation.

$2D$-system theory for systems having separable denominator [3] may be applied here. Rational systems in this latter case are of the following form [24]:

$$H = C(I - \overline{\gamma}A_{\overline{\gamma}})^{-1}P(I - \gamma A_{\gamma})^{-1}B \tag{35}$$

which yields the following state space form [we use here notations similar to those of (27) and (28)]

$$\begin{cases} v_t &= A_{\gamma}\left(\frac{v_{t\alpha}+v_{t\beta}}{2}\right) + Bu_t \\ z_t &= P_2 v_t \\ x_{t\alpha} &= A_{\overline{\gamma}}x_t + P_1 z_{t\alpha} \\ x_{t\beta} &= A_{\overline{\gamma}}x_t + P_1 z_{t\beta} \\ y_t &= Cx_t \end{cases} \tag{36}$$

where $P = P_1P_2$. The first two equations define a purely "anticausal" system, whereas the last three equations define a causal zero depth system.

Causal Systems: Here we interpret the sequence $\sigma^{[i]}$ as the powers of the row operator and $\overline{\gamma}$ as the column operator. Then again we consider the row-by-row scanning to define a total ordering of the $2D$ index space. This corresponds to decomposing the system H according to the following two steps:

1) A smoothing along the considered horocycle (i.e. constant scale smoothing), followed by
2) A top-down (i.e. coarse-to-fine) propagation.

$2D$-system theory for systems with a separable denominator [3] may again be applied here. Rational systems in this latter case are of the following form [24]:

$$H = C(I - \overline{\gamma}A_{\overline{\gamma}})^{-1}P(I - \sigma A_{\sigma})^{-1}B \tag{37}$$

where it is understood that, in expanding such a formula into a power series, σ^i should be replaced by $\sigma^{[i]}$. As a consequence, the latter has an unusual feature in that no tractable time-domain translation of the "frequency domain" formula (37) is available. The *finite depth* case, however, yields

$$\begin{cases} x_{t\alpha} &= A_{\overline{\gamma}}x_t + B\left(1,\sigma,...,\sigma^{[i]}\right)u_{t\alpha} \\ x_{t\beta} &= A_{\overline{\gamma}}x_t + B\left(1,\sigma,...,\sigma^{[i]}\right)u_{t\beta} \\ y_t &= Cx_t \end{cases} \tag{38}$$

where $B\left(1,\sigma,...,\sigma^{[i]}\right)$ is a linear combination of the listed operators. This corresponds to the case where A_σ is nilpotent. Thus stationary finite depth scalar systems are of the form

$$H = \frac{1}{\det(I - \overline{\gamma}A_{\overline{\gamma}})}K\left(\overline{\gamma}; 1, \sigma, ..., \sigma^{[i]}\right)$$

where $K(...)$ is a linear combination of the $1, \sigma, ..., \sigma^{[i]}$'s with coefficients that are polynomials in $\overline{\gamma}$. It is easily shown that H may be equivalently expressed in the following *ARMA* form

$$H = A^{-1}B \tag{39}$$

where A is a causal system of *finite support* involving the operators $\overline{\gamma}$ and $1, \sigma, ..., \sigma^{[i]}$ whereas $B = B\left(1,\sigma,...,\sigma^{[k]}\right)$ is as in (38). The AR modeling filters for isotropic processes introduced in [4]–[6] are in fact ARMA systems in the above sense.

4.5 Stationary Stochastic Processes

To simplify the presentation, we concentrate here on scalar processes.

Definition: A zero mean stochastic process y is said to be stationary if its covariance function is translation invariant, i.e.

$$\mathbf{E}(y_s y_t) = \mathbf{E}\left(y_{\tau(s)}y_{\tau(t)}\right)$$

for any primitive translation τ.

The following theorem shows that this definition of stationarity for processes is consistent with that of stationarity for transfer functions (this theorem is proved in Appendix B).

Theorem: Let y be a stochastic process.

1) The process y is stationary if and only if

$$\mathbf{E}(y_s y_t) = r[d(s, s\wedge t), d(t, s\wedge t)]$$

where $s \wedge t$ is defined in (25).

2) If the process u and the transfer function H are both stationary, so is the process Hu.

Note that the second statement is an immediate consequence of the first one.

Remark: Theorem 6 has the following interesting result as a consequence. Pick a point $t_o \in \mathcal{T}$ and order the words[4] $w \in \{\alpha, \beta\}^*$ of length n according to lexicographic order with priority to α: the corresponding set of nodes $t_o w$ is exactly the left-to-right ordered horocycle "segment" in Fig. 6, collect the values $y_{t_o w}$ into a vector Y. Then the covariance matrix Σ_Y of Y has the following recursively defined structure:

$$\Sigma(r_0) = r_0$$
$$\Sigma(r_0, ..., r_m) = \begin{bmatrix} \Sigma(r_0, ..., r_{m-1}) & r_m U_{m-1} \\ r_m U_{m-1} & \Sigma(r_0, ..., r_{m-1}) \end{bmatrix}$$
$$\Sigma_Y = \Sigma(r_0, ..., r_n)$$

where U_m is a $2^m \times 2^m$-matrix whose entries are 1. It is then easy to show that *the eigenvectors of Σ_Y are the discrete Haar basis* (see [4], [13] for more details). □

More generally, x and y are said to be *jointly* stationary if we have

$$\mathbf{E}(x_s y_t) = r^{xy}[d(s, s\wedge t), d(t, s\wedge t)]. \tag{40}$$

[4]The notation to follow denotes the language of the words on the alphabet $\{\alpha, \beta\}$.

4.6 Spectral Calculus

The purpose of this subsection is to provide a formula to encode how stationary transfer functions modify covariance sequences. In the case of time series, the product of usual transfer functions and spectra is used for this purpose. Consider two jointly stationary processes x and y. Pick a pair (s,t) and set $d[s, s\wedge t] = i, d[s\wedge t, t] = j$. We have

$$\begin{aligned}\mathbf{E}(x_s y_t) &= 2^{\min\{1-j,0\}} \sum_{t':d[s\wedge t',t']=j} \mathbf{E}(x_s y_{t'}) \\ &= \mathbf{E}\big(x_s (\overline{\gamma}^i \gamma^j . y)_s\big)\end{aligned}$$

This suggests to define the *cross-spectrum* of x and y as the following power series:

$$R^{xy} \triangleq \sum_{i,j\geq 0} r^{xy}[i,j]\overline{\gamma}^i\gamma^j$$

where $r^{xy}[i,j]$ is the cross-covariance sequence of x and y, see (40). Finally, given a stationary transfer function of the form $H = \sum h_{i,j}\overline{\gamma}^i\gamma^j$ (cf. (30)), it will be useful to introduce the following notion of an "adjoint" :

$$H^* \triangleq \sum h_{j,i}\overline{\gamma}^i\gamma^j$$

Then the following formula yields the cross-spectrum of two stationary processes Hx and Ky, where H and K are stationary transfer functions and x, y are jointly stationary processes :

$$R^{(Hx)(Ky)} = H^* R^{xy} K \tag{41}$$

See Appendix B for a proof. This formula generalizes a well-known result of the case of standard stationary time series.

V. Conclusion

In this paper we have introduced and developed the basic concepts of a system theory for the multiscale modeling and processing of signals. The starting point for this theory is an examination of the abstract properties defining QMF banks and the identification of the filtering and decimation operations performed in such banks as signal transformations from one scale of representation to another. This led directly to the idea of viewing QMF filtering and decimation operations in terms of a pair of operators acting on $l^2(\mathcal{T})$ where $\mathcal{T}$ is the homogeneous dyadic tree and where these operators satisfy algebraic properties directly inherited from the algebraic QMF conditions. Using these abstract QMF operators as basic "shifts" on the tree, we developed a system theory on the homogeneous dyadic tree as a foundation for a *multiscale system theory*. We have shown that the homogeneous tree possesses critical geometric properties that have the following consequences: the double role played by the classical z-transform, namely: (1) Encoding systems as weighted sums of products of the basic shift operators; and (2) defining stationarity, must be split into two separate objects—the shifts (which are *not* invertible) to encode systems, and the translations to define stationarity (which are not easily expressed via shifts). We sketched two system theories, each with an emphasis on one of these two different objects. Finally a notion of stationary stochastic processes has been introduced, and the transformation of the second-order "spectral" characterization of a process when it is passed through a stationary multiscale system has been characterized via a simple "spectral calculus" formula. The main results of this paper are summarized below and we also suggest a number of directions for further work:

1) The usual notion of QMF pairs in multirate digital filtering generalizes naturally to abstract QMF pairs $\{\alpha, \beta\}$ of operators acting on $l^2(\mathcal{T})$. Thus, system theory for abstract QMF pairs on $l^2(\mathcal{T})$ is a natural framework for fundamental studies on multiscale signal processing. We established the foundations for such a system theory, further work has to be pursued to exploit it. In particular, we have shown in (13) and (14) how the wavelet and wavelet packet filter banks are encoded as rational systems on the tree. A natural question to look at would then be to characterize and completely parametrize *all* rational systems that are associated with orthonormal decompositions of l^2-spaces of signals.
2) We also see that multiscale signal processing within this framework has (at least) *two* levels of flexibility as compared to usual discrete-time processing. In particular, in addition to the flexibility in defining the system on the tree (as in the previous point), we also have flexibility in the choice of the concrete specification of what the variables on the tree actually represent. That is, we have the flexibility in the specific choices of QMF filters corresponding to the abstract pair of literature on filter banks [31] as well as the analytical properties of wavelets inherited from the QMF's that define them [19].
3) There is a unique natural way to define stationarity for both systems and stochastic processes on the homogeneous tree. Such a notion emphasizes *"stochastic fractalness"*, as discussed at length in [7], [5], [4]. An important direction for further research is the stochastic realization problem, which is the construction of multiscale models driven by white noise and which produce outputs whose second-order characteristics match (either exactly or approximately) those of a given process. In [4,5] we provide some results along these lines for the construction of so-called multiscale autoregressive models whose outputs have second-order statistics that are or a theory of multiscale state space stochastic realization.

Finally, we note that the results of this paper immediately generalize to homogeneous trees with more than 3 branches originating from each node. For instance, multiscale system theory for images would require a homogeneous tree with 5 branches at each node, i.e., a quadtree (1 branch toward the coarser scale, and 4 for the pyramid going toward the finer scale). Here again our framework may provide new views and insights on how to model, analyze, and design multiresolution systems and representations for signals indexed by multidimensional sets.

ACKNOWLEDGMENT

The authors are indebted to a sizeable group for fruitful discussions and stimulations, namely M. Basseville, K.C. Chou, B. Claus, A. Cohen, J.-P. Conze, B.C. Levy, Y. Meyer, and O. Zeitouni.

We also gratefully thank the reviewers for helpful comments and corrections to an earlier version of the manuscript.

APPENDIX A. SOME USEFUL RESULTS ON THE GEOMETRY OF THE HOMOGENEOUS TREE

We collect here all results we need on the geometry of the homogeneous tree.

Classifying the Isometries: Here we prove lemma 1. The three types of isometries have just the *identity* in their intersection. In the sequel we consider *nontrivial* isometries only. First, it is clear that the cases (22) and (23) are distinct. Hence, it remains to prove that an isometry f, which is neither of type (22) nor of type (23), must be of type (24). This is what we do now. Let f denote such an isometry. Introduce

$$d_{\min} = \min_{t \in \mathcal{T}} d[t, f(t)].$$

We know that $d_{\min} > 0$ due to the assumption. Hence, consider

$$\mathcal{T}_{\min} = \{t : d[t, f(t)] = d_{\min}.\}$$

Note that, for $t \in \mathcal{T}_{\min}$, we have

$$d_{\min} = d[t, f(t)] = d\big[f(t), f^2(t)\big]$$

so that $\mathcal{T}_{\min}$ is f-invariant. Next, denote by $[\![s, t]\!]$ the path linking s and t. Considering that $t' \in [\![t, f(t)]\!]$, we claim that $t' \in \mathcal{T}_{\min}$ as well. Indeed

$$\begin{aligned} d_{\min} &= d[t, t'] + d[t', f(t)] \\ &= d[f(t), f(t')] + d[f(t'), f^2(t)] \end{aligned}$$

whence

$$\begin{aligned} d_{\min} &= d[t', f(t)] + d[f(t), f(t')] \\ &= d[t', f(t')] \end{aligned}$$

Considering the infinite sequence $t, f(t), f^2(t), \ldots$, we have

$$\mathcal{T}_{\min} \subseteq \bigcup_{k \geq 0} [\![t, f^k(t)]\!] \triangleq \Gamma.$$

Assume Γ to be a *finite* path. Then we must have, for some k_o, $f^{k_o+1}(t) \in [\![t, f^{k_o}(t)]\!]$, so that the points $f^{k_o}(t)$ and $f^{k_o+1}(t)$ would be exchanged by f. This latter fact turns out to contradict our assumption that f is not of type (23). Hence $\mathcal{T}_{\min}$ must contain an infinite path. If $\mathcal{T}_{\min}$ contains more than a unique path, f must be the identity. This proves the lemma. □

Translations Commute with $\wedge$:

Lemma: Let τ be a translation. Then we have

$$\tau(s \wedge t) = \tau(s) \wedge \tau(t) \tag{42}$$

$$d[s, s \wedge t] = d[\tau(s), \tau(s) \wedge \tau(t)]. \tag{43}$$

Proof: (43) is a consequence of (42), so that we just prove the first formula. Denote by Γ the infinite path left invariant by τ. We have

$$d[s \wedge t, \Gamma] = d[\tau(s \wedge t), \Gamma].$$

Hence, taking a point s_o on Γ sufficiently close to $-\infty$, we have $d[s_o, s \wedge t] \leq \min\{d[s_o, s], d[s_o, t]\}$ so that $d[\tau(s_o), \tau(s \wedge t)] \leq \min\{d[\tau(s_o), \tau(s)], d[\tau(s_o), \tau(t)]\}$. Hence $\tau(s \wedge t)$ is closer to $-\infty$ than $\tau(s) \wedge \tau(t)$. In particular

$$d[\tau(s), \tau(s \wedge t)] \geq d[\tau(s), \tau(s) \wedge \tau(t)]. \tag{44}$$

and similarly for t. From (44) we get

$$\begin{aligned} d[s, t] &= d[s, s \wedge t] + d[s \wedge t, t] \\ &= d[\tau(s), \tau(s \wedge t)] + d[\tau(s \wedge t), \tau(t)] \\ &\geq d[\tau(s), \tau(s) \wedge \tau(t)] + d[\tau(s) \wedge \tau(t), \tau(t)] \\ &= d[\tau(s), \tau(t)] \end{aligned}$$

so that the inequality in (44) must in fact be an equality. This and the fact that $\tau(s \wedge t)$ is closer to $-\infty$ than $\tau(s) \wedge \tau(t)$ together prove the lemma. □

Translations Generate Vertical Symmetries: We call *vertical symmetries* and denote by ν the isometries of the following type:

1) Pick a point $t_o \in \mathcal{T}$, call it the *pivot.*
2) Partition $\mathcal{T}$ according to $\mathcal{T} = \mathcal{T}_{t_o}^{\downarrow} \cup \mathcal{T}_{t_o}^{\uparrow}$, where

$$\mathcal{T}_{t_o}^{\downarrow} = \{s = t_o w^{\downarrow} | w^{\downarrow} \in \{\alpha, \beta\}^*\}.$$

3) ν satisfies

$$\nu(s) = \begin{cases} s & \text{if} \quad s \in \mathcal{T}_{t_o}^{\uparrow} \\ t_o \check{w}^{\downarrow} & \text{if} \quad s = t_o w^{\downarrow}. \end{cases}$$

where $\check{w}^{\downarrow}$ is the word obtained by exchanging α and β in $w^{\downarrow}$.

Lemma: The closure (w.r.t. the weak topology) of the group spanned by the primitive translations contains the vertical symmetries.

Proof: consider a vertical symmetry ν with pivot t_o. Introduce $\Gamma_\alpha = [\![-\infty, t_o]\!] \cup \{t_o \alpha^n, n > 0\}$ and similarly for Γ_β. $f \triangleq \tau_{\Gamma_\beta} \circ \tau_{\Gamma_\alpha}^{-1}$ is an isometry, which keeps $\mathcal{T}_{t_o}^{\uparrow}$ invariant and exchanges $t_o\alpha^n$ and $t_o\beta^n$. Since f is an isometry, it must also exchange $t_o\alpha^{n-1}\beta$ and $t_o\beta^{n-1}\alpha$. Next consider

$$\begin{aligned} &t_1 = \text{the } s = t_o w^{\downarrow} \text{ closest to } t_o \text{ such that} \\ &\quad s\alpha \text{ is not exchanged with } s\beta. \end{aligned}$$

Certainly, it holds that $t_1 \neq t_o$. So we apply the procedure above (definition of f, etc...) to t_1 instead of t_o. Given any finite subset of the tree, after finitely many steps, the resulting composition of primitive translations coincides with ν on this finite subset. This proves the lemma. □

More on Primitive Translations: It will be convenient to re-encode definition (26) of primitive translations using the shift operators on $\mathcal{T}$. Let $\Gamma = \{t_n\}_{n\in\mathbf{Z}}$ be the skeleton of the considered primitive translation denoted by τ_Γ, and denote by s_n the unique point outside the skeleton such that $d(t_n, s_n) = 1$. Then τ_Γ is encoded by the following formulae:

$$\begin{aligned}\tau_\Gamma(t_n) &= t_{n+1} \\ \tau_\Gamma(s_n w^\downarrow) &= s_{n+1} w^\downarrow\end{aligned} \tag{45}$$

Given two skeletons Γ and Γ', we define their *composition*

$$\Gamma'' \triangleq \Gamma \circ \Gamma'$$

by the following formulae, where we label the two skeletons in such a way that they exactly bifurcate after t_0, i.e., $t_0 = t'_0, t_1 \neq t'_1$, and n denotes an arbitrary nonnegative integer:

$$\begin{aligned}t''_{-n} &= t_{-n} \\ t''_1 &= t_1 \\ t''_{2+n} &= s_1 w^\downarrow \text{ if } t'_{1+n} = t'_1 w^\downarrow.\end{aligned} \tag{46}$$

We have the following result:

$$\tau_\Gamma \circ \tau_{\Gamma'} = \tau^2_{\Gamma\circ\Gamma'}. \tag{47}$$

A nice consequence of formula (47) is that *the family of powers of primitive translations is a semigroup.*

Appendix B. Proofs of the Results on Stationary Systems andStochastic Proccesses.

Proof of the Characterization of Stationary Systems Here we prove theorem 4.

Proof of 1. Since H is a linear operator we have, for any primitive translation τ

$$[H.u]_t = \sum_{s\in\mathcal{T}} h_{t,s} u_s$$

whence

$$[H.\tau_\Gamma(u)]_t = \sum_{s\in\mathcal{T}} h_{t,s} u_{\tau(s)}$$

and

$$\begin{aligned}[\tau_\Gamma(H.u)]_t &= \sum_{s\in\mathcal{T}} h_{\tau(t),s} u_s \\ &= \sum_{s\in\mathcal{T}} h_{\tau(t),\tau(s)} u_{\tau(s)}\end{aligned}$$

Since u is arbitrary, H must satisfy

$$h_{t,s} = h_{\tau(t),\tau(s)}$$

which proves 1.

Proof of 2. Denoting by $\mathbf{N}$ the set of the nonnegative integers, we can write H in the form

$$H = \sum_{i,j\in\mathbf{N}} h_{i,j} \sum_{\substack{|w^\uparrow| = i \\ |w^\downarrow| = j}} w^\uparrow w^\downarrow$$

so that

$$(Hu)_t = \sum_{i,j} h_{i,j} \sum_{s:\begin{cases} d[t, t\wedge s] = i \\ d[t\wedge s, s] = j\end{cases}} u_s$$

which proves part two of the theorem, thanks to (43) in lemma 2. □

Proof of the Characterization of Stationary Processes Here we prove theorem 6. The sufficiency of the condition of the theorem follows triviallly from (43) in lemma 2. To prove the necessity, select a path Γ_o containing the segment $[s\wedge t, t]$, and take $m > d[s\wedge t, t]$. For any such path we have

$$\mathbf{E}(y_s y_t) = \mathbf{E}\Big(y_{\tau^m_{\Gamma_o}(s)} y_{\tau^m_{\Gamma_o}(t)}\Big). \tag{48}$$

Now, take any path Γ such that

$$\Gamma \cap \Gamma_o \supseteq (-\infty, \tau^m_{\Gamma_o}(s\wedge t)]$$

then (48) holds for such a path as well. But $\tau^m_\Gamma(s\wedge t) = \tau^m_{\Gamma_o}(s\wedge t)$ and $\tau^m_\Gamma(s) = \tau^m_{\Gamma_o}(s)$ both hold true, whereas $\tau^m_\Gamma(t)$ *ranges over all points t' such that*

$$d[\tau^m_{\Gamma_o}(s\wedge t), t'] = d[\tau^m_{\Gamma_o}(s\wedge t), \tau^m_{\Gamma_o}(t)].$$

This proves that $\mathbf{E}(y_s y_t)$ is invariant for those points t that are at a given distance of $s\wedge t$. The necessity of the condition of the theorem 6 follows easily. □

Proof of the Spectral Calculus Formula: It is enough to verify (41) in the following cases:

$$H = \overline{\gamma}, K = 1 \tag{49}$$
$$H = \gamma, K = 1 \tag{50}$$
$$H = 1, K = \overline{\gamma} \tag{51}$$
$$H = 1, K = \gamma. \tag{52}$$

Pick a pair (s, t) with $d[s, s\wedge t] = i, d[s\wedge t, t] = j$. Consider (49). Then

$$\begin{aligned}\mathbf{E}((Hx)_s(Ky)_t) &= \mathbf{E}((\overline{\gamma}.x)_s y_t) \\ &= \mathbf{E}\big((\overline{\gamma}.x)_s(\overline{\gamma}^i\gamma^j.y)_s\big) \\ &= \mathbf{E}\big(x_s(\gamma\overline{\gamma}^i\gamma^j.y)_s\big)\end{aligned}$$

holds true, where the last equality is readily verified. Noting that $\overline{\gamma}^* = \gamma$ we get (41) in case (49). For (50), write

$$\begin{aligned}\mathbf{E}((Hx)_s(Ky)_t) &= \mathbf{E}((\gamma.x)_s y_t) \\ &= \mathbf{E}\big((\gamma.x)_s(\overline{\gamma}^i\gamma^j.y)_s\big) \\ &= \mathbf{E}\big(x_s(\overline{\gamma}\overline{\gamma}^i\gamma^j.y)_s\big)\end{aligned}$$

which also yields (41), since $\overline{\gamma} = \gamma^*$. For the case (51) we write

$$\begin{aligned}\mathbf{E}((Hx)_s(Ky)_t) &= \mathbf{E}(x_s(\overline{\gamma}.y)_t) \\ &= \mathbf{E}\big(x_s(\overline{\gamma}^i\gamma^j.\overline{\gamma}y)_s\big)\end{aligned}$$

which proves (41) in this case. Case (52) is handled in a similar way. □

REFERENCES

[1] J. P. Arnaud, "Fonctions sphériques et fonctions définies positives sur l'arbre homogène," *C. R.A. S.* t. 290, série A (14 Jan. 1980), p. 99–101.

[2] J. P. Arnaud, G. Letac, "La formule de représentation spectrale d'un processus gaussien stationnaire sur un arbre homogène," *Publi. Labo. Stat. and Proba.*, UA 745, Toulouse, France.

[3] S. Attasi, "Systèmes linéaires homogènes à deux indices," *Rapport Laboria*, no. 31, Rocquencourt, 1973.

[4] M. Basseville, A. Benveniste, and A. S. Willsky, "Multiscale autoregressive processes, part I and part II," *IEEE Trans. Signal Processing,* vol. 40, no. 8, pp. 1915–1955, 1992.

[5] M. Basseville, A. Benveniste, K. C. Chou, and A. S. Willsky, "Multiscale statistical signal processing: Stochastic processes indexed by trees," in *MTNS 89*, June 19-23, 1989, Amsterdam.

[6] M. Basseville and A. Benveniste, "Multiscale statistical signal processing," *Proc. of the ICASSP-89,* 2065-2068, Glasgow, 1989.

[7] M. Basseville, A. Benveniste, K. C. Chou, S. A. Golden, R. Nikoukhah, and A. S. Willsky, "Modeling and estimation of multiresolution stochastic processes," *IEEE Trans. Inform. Theory,* special issue on Wavelets and Multiresolution Analysis, vol. 38 no. 2, 766–784, 1992.

[8] A. Benveniste, "Multiscale signal processing: from QMF to wavelet," to appear in the International Series on *Advances in Control and Dynamic Systems,* C. T. Leondes Ed., San Diego, CA: Academic Press, 1993.

[9] J. Berstel, C. Reutenauer, "Les séries rationnelles et leurs langages," in collection *"Etudes et Recherches en Informatique,"* Masson, 1984.

[10] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code,"*IEEE Trans. Comm.*, vol. 31, pp. 482-540, 1983.

[11] P. Cartier, "Harmonic analysis on trees," *Proc. Sympos. Pure Math.*, vol. 26, Amer. Math. Soc., Providence, RI, pp. 419–424, 1974.

[12] P. Cartier, "Géométrie et analyse sur les arbres". *Séminaire Bourbaki*, no. 407, Feb. 1972.

[13] K. C. Chou, A. S. Willsky, and A. Benveniste, "Multiscale recursive estimation, data fusion, and regularization,"*IEEE Trans. on Automat. Cont.,* to appear.

[14] K. C. Chou, A. S. Willsky, and R. Nikoukhah, "Multiscale systems, Kalman filters, and Riccati equations,"*IEEE Trans. on Automat. Cont.,* to appear.

[15] K. C. Chou, S. A. Golden, and A. S. Willsky, "Wavelet transform-based algorithms for the fusion of multiresolution data," *Signal Processing*, to appear.

[16] R. R. Coifman and M. V. Wickehauser, "Entropy-based algorithms for best basis selection" *IEEE Trans. Inform. Theory,* special issue on Wavelets and Multiresolution Analysis, vol. 38, no. 2, 1992.

[17] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1983.

[18] I. Daubechies, A. Grossmann, and Y. Meyer, "Painless nonorthogonal expansions," *J. Math. Phys.*, vol. 27, p. 1271–1283, 1986.

[19] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. on Pure and Appl.Math.*, vol. 91, pp. 909–996, 1988.

[20] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inform. Theory,* vol. 36, 1990, 961-1005.

[21] J. L. Dunau, "Etude d'une classe de marches aléatoires sur l'arbre homogène," in *Arbres homogènes et Couples de Gelfand*, J. P. Arnaud, J. L. Dunau, G. Letac, Eds. Publications du Laboratoire de Statistique et Probabilités, Univ. Paul Sabatier, Toulouse, France, no. 02-83, June 1983.

[22] I. Gohberg, M. A. Kaashoek, and L. Lerer, "Nodes and realization of rational matrix functions: minimality theory and applications,"*Operator Theory: Advances and Applicat.,* vol. 29, 1988 Basel: Birkhauser Verlag.

[23] J. Kovacevic and M. Vetterli, "Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for $\mathbf{R}^n$," *IEEF Trans. Inform. Theory,* special issue on Wavelets and Multiresolution Analysis, vol. 38 no. 2, 1992.

[24] T. Lin, M. Kawamata, and T. Higuchi, "New necessary and sufficient conditions for local controllability and observability of 2-D separable denominator systems," *IEEE Trans. Automat. Cont.*, vol. 32, pp. 254-256, 1987.

[25] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. and Mach. Intell.,* vol. 11, July 1989, 674-693.

[26] M. R. Luettgen, W. C. Karl, and A. S. Willsky, "An efficient algorithm for multiscale regularization with application to the computation of optical flow," *IEEE Trans. Image Processing,* to appear.

[27] M. R. Luettgen, W. C. Karl, A. S. Willsky, and R. R. Tenney, "Multiscale representations of Markov random fields," to appear in the special issue on wavelet transformation of the *IEEE Trans. Signal Processing* .

[28] M. R. Luettgen, and A. S. Willsky, "Likelihood calculation for a class of multiscale stochastic models with applications to texture discrimination," submitted for publication.

[29] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Trans. Acoust., Speech and Signal Processing,* vol. 37, Dec. 1989, 2091–2110.

[30] Y. Meyer, *Ondelettes et Opérateurs.* Hermann actualités mathématiques, Paris: Hermann, 1990.

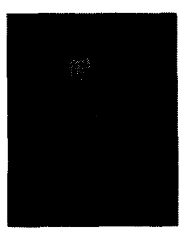
[31] P. P. Vaidyanathan, *Multirate Systems and Filter Banks.* Signal Processing Series, Englewood Cliffs, NJ: Prentice–Hall, 1993.

[32] P. P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect reconstruction techniques,"*IEEE-ASSP Mag.,*vol. 4, no. 3, pp. 4-20, 1987.

Albert Benveniste (M'81, SM'89, F'91) graduated from the Ecole des Mines de Paris in 1971. He performed his Thèse d'Etat in Mathematics, probability theory, in 1975.

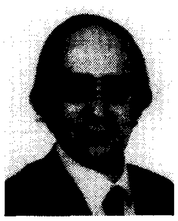
From 1976 to 1979 he was an associate professor in mathematics at the Universite de Rennes I. From 1979 to the present he has been Directeur de Recherche at INRIA. For his thesis, he worked on probability theory, stochastic processes, and ergodic theory. In parallel, he pursued work on automatic control and signal processing in the area of adaptive and change detection algorithms for time-varying systems, with applicatiojns to data communications, vibration mechanics, and fault detection in process control. Since 1987, he has been interested in computer science, in the area of real-time languages and systems. Cooperating with P. Le Guernic, he participated in the definition of and and theoretic studies on the SIGNAL language. He has co-authored with M. Métivier and P. Priouret the book *Adaptive Algorithms and Stochastic Approximations*, and has been an editor, jointly with Michèle Basseville of the collective monograph *Detection of Abrupt Changes in Signals and Systems.*

In 1980, Albert Benveniste was co-winner of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL Best Transactions Paper Award for his paper on blind deconvolution in data communications. In 1990, he received the CNRS silver medal, and in 1991, he was elected an IEEE fellow. From 1986 to 1990, he was Vice-Chairman of the IFAC committee on Theory, and was Chairman of this committee for 1991–1993. From 1987 to 1990, he was an Associate Editor for IEEE TRANSACTIONS ON AUTOMATIC CONTROL . He is currently an Associate Editor for the INTERNATIONAL JOURNAL OF ADAPTIVE CONTROL AND SIGNAL PROCESSING, INTERNATIONAL JOURNAL OF DISCRETE EVENT DYNAMICAL SYSTEMS, and Associate Editor at Large for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.



Ramine Nikoukhah (S'87, M'90) received the B.S. degree from the University of California, Davis in 1982; the degree of Maitrise from the Université de Paris-Sud, Orsay, France, in 1983; and the M.S. and the Engineer's Degrees in 1986, and the Ph.D. degree in 1988 from the Massachusetts Institute of Technology, Cambridge, MA, all in electrical engineering and computer science.

During 1983 he was a research engineer at TAU Co., Los Gatos, CA, where he worked in the area of flight control. Since 1988, he has been with the Institut National de Recherche en Informatique et en Automatique (INRIA), Rocquencourt, France. He has also done consulting for Matra Co., and has held teaching positions at various engineering schools in France. His research interests are in the areas of control, signal processing, and algebraic system theory. He is currently involved in the development of the control and signal processing software package Scilab at INRIA.



Alan S. Willsky (S'70, M'73, SM'82, F'86) received both the S.B. degree and the Ph.D. degree from the Massachusetts Institute of Technology in 1969 and 1973, respectively.

He joined the M.I.T. faculty in 1973, and his present position is Professor of Electrical Engineering. From 1974 to 1981, Dr. Willsky served as Assistant Director of the M.I.T. Laboratory for Information and Decision Systems. He is also founder and a member of the board of directors of Alphatech, Inc. Dr. Willsky has held visiting positions at Imperial College, London; L'Université de Paris-Sud; and the Institut de Recherche en Informatique et Systèmes Aléatoires in Rennes, France. He is the author of the research monograph *Digital Signal Processing and Control and Estimation Theory* and is co-author of the undergraduate text *Signals and Systems.* His present research interests are in problems involving multidimensional and multiresolution estimation and imaging, discrete-event systems, and the asymptotic analysis of control and estimation systems.

Dr. Willsky was program chairman for the 17th IEEE Conference on Decisions and Control, has been an Associate Editor of several journals, including IEEE TRANSACTIONS ON AUTOMATIC CONTROL, has served as a member of the Board of Governors and Vice President for Technical Affairs of the IEEE Control Systems Society, was Program Chairman for the 1981 Bilateral Seminar on Control Systems held in the People's Republic of China, and was Special Guest Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY on wavelet transforms and multiresolution signal analysis. Also in 1988, he was made a Distinguished Member of the IEEE Control Systems Society. In addition, Dr. Willsky has given several plenary lectures at major scientific meetings, including the 20th IEEE Conference on Decision and Control, the 1991 IEEE International Conference on Systems Engineering, the SIAM Conference on Applied Linear Algebra, 1991, and the 1992 Inaugural Workshop for the National Centre for Robust and Adaptive Systems, Canberra, Australia. He received the Donald P. Eckman award from the American Automatic Control Council in 1975. Dr. Willsky has also been awarded the Alfred Noble Prize by the ASCE and the 1980 Browder J. Thompson Memorial Prize by the IEEE for a paper excerpted from his monograph.