Subband coding for sigmoidal nonlinear operations

John J. Benedetto
Department of Mathematics
University of Maryland
College Park, MD 270742

Sandra Saliani
Department of Mathematics
University of Potenza
Potenza, Italy

ABSTRACT

A theory of waveletpackets is developed for nonlinear operators consisting of a composition, generalizing a sigmoidal operation, followed by convolutions with filter pairs H_0 and H_1. The pyramidal waveletpacket structure is defined by bit reversal trees. The reconstruction theorem, from which the original signal is obtained from frequency localized data at other nodes of the tree, requires fixed point theory as well as conditions on H_0 and H_1 resembling those defining quadrature mirror filter pairs. Applications will be to biological systems and neural networks where such nonlinearities occur.

1. INTRODUCTION

Sigmoidal nonlinearities arise in mathematical models of biological systems[1,3,6,13] and in neural networks. They are a special case of superpositions[2] Φ, viz., $(\Phi f)(t) = \phi(t, f(t))$; and superpositions occur everywhere, e.g., differential and integral equations, variational calculus, probability, etc. The theory of waveletpackets[8] arises in addressing data compression problems, and is a tool in a variety of frequency localization problems. The history of waveletpackets includes engineering efforts to deal with channel crosstalk at the turn of the century, as well as significant advances in subband coding by the speech and image processing communities[7] in the late 1970s and early 1980s. There is also related recent work by Mallat and Zhang on their concept of *matching pursuits*[11]. In this paper we consider nonlinear systems with a subband coding structure. The goal is to provide *local* and *self-similar* low pass and high pass filtering, and accompanying signal reconstruction. A consequence of this type of filtering is computable frequency localization.

There is a more general and mathematical theory of nonlinear waveletpackets that we have developed[5]. In this paper, as well as in the more general theory, the setting is the *real Paley-Wiener space* $PW_{\Omega,r}$. The elements f of $PW_{\Omega,r}$ are real valued, Ω-bandlimited, finite energy signals, i.e.,

$$PW_{\Omega,r} = \{f\colon \mathbb{R} \to \mathbb{R}\colon f \in L^2(\mathbb{R}) \text{ and } \operatorname{supp} \widehat{f} \subseteq [-\Omega, \Omega]\},$$

where $L^2(\mathbb{R}) = \{f\colon \|f\|_2 = (\int |f(t)|^2\, dt)^{1/2} < \infty\}$, $\widehat{f}(\gamma) = \int f(t)e^{-2\pi i t\gamma}\, dt$, integration is over $\mathbb{R}$, and $\operatorname{supp} \widehat{f} \subseteq [-\Omega, \Omega]\}$ signifies that $\widehat{f} = 0$ outside of $[-\Omega, \Omega]$. There are new constructive results[12] on the structure of $PW_{\Omega,r}$ in terms of the notion of weak coercivity; the analysis involves Caccioppoli's Global Inverse Mapping Theorem.

In *Section 2* we motivate and describe a general problem. We solve a refinement of the problem recursively in *Section 3* by means of fixed point theory. This part is inspired by the work of Landau and Miranker[10] Finally, in *Section 4* we formulate the problem as a local and self-similar low pass and high pass subband coding pyramidal scheme. Using the results from *Section 3*, we show how to reconstruct the given signal from frequency localized nonlinear versions of it.

2. MOTIVATION AND GENERAL PROBLEM

In some mammalian auditory models[13] the output of a speech signal $f \in PW_{\Omega,r}$ on the basilar membrane filter bank can be thought of as the *wavelet transform*,

$$W_g f(s,t) = (f * D_s g)(t),$$

where t is time, $*$ is convolution, g is the impulse response for the cochlear filter[1] $\widehat{g}$, $s = a^m$ is a scale channel for some $a > 1$, and $D_s g(t) = s^{1/2}\, g(s,t)$. Next, an instantaneous sigmoidal nonlinearity ϕ is applied to $W_g f$ followed by a low pass filter with impulse response h. These processes model the threshold and saturation that occur in the hair cell channels, and the leakage of electrical current through the membranes of these cells[1]. The *cochlear output* is

$$C(s,t) = (\phi \circ W_g f(s,t)) * h(t) \tag{1}$$

for each channel s. The nonlinearity ϕ can be of the form

$$\phi(y) = \phi_T(y) = \frac{e^{Ty}}{1 + e^{Ty}},$$

noting that ϕ is an increasing function and $\lim_{T\to\infty} \phi_T$ is the Heaviside function. From this point there are a number of hypotheses, e.g., the lateral inhibitory network, and a number of techniques for speech signal reconstruction, e.g., alternating projections[13] and irregular sampling[3,6]. Our goal in this paragraph has been to establish the importance of systematically studying models such as Equation (1).

Let L be the function

$$\begin{aligned} L\colon PW_{\Omega,r} &\longrightarrow PW_{\Omega,r} \\ f &\longmapsto Lf = (\phi \circ f) * h, \end{aligned} \tag{2}$$

where $\phi\colon \mathbb{R} \to \mathbb{R}$ is differentiable and monotonic, $0 < m \le |\phi'| \le M$ on $\mathbb{R}$, $h \in PW_{\Omega,r}$ is the impulse response of the filter $\widehat{h} = H$, and $0 \le |H| \le \beta$ on $[-\Omega, \Omega]$. The *general problem* is to reconstruct f from frequency localized data. We are more precise and technical in *Sections 3* and *4*.

Assume only that $|\phi'| \le M$ on $\mathbb{R}$ and $\phi(0) = 0$. Then for all real valued elements $f \in L^2(\mathbb{R})$, we have $\phi \circ f \in L^2(\mathbb{R})$. With this assumption and the hypothesis that $|H| \le \beta$ on $[-\Omega, \Omega]$, it is also easy to see that L defined by (2) is a Lipschitz function; in fact,

$$\forall f_1, f_2 \in PW_{\Omega,r}, \quad \|Lf_1 - Lf_2\|_2 \le M\beta \|f_1 - f_2\|_2.$$

3. A FIXED POINT PROBLEM AND SOLUTION FOR L

3.1 Beurling's uniqueness theorem

It is not difficult to prove the following result by means of the Parseval-Plancherel theorem and showing a little care about the signs of the functions involved. The idea is due to Beurling, and was used by Landau and Miranker[10] for the case that $H = 1_{[-\Omega,\Omega]}$, the characteristic function of $[-\Omega, \Omega]$.

3.1 Theorem. *Let $L\colon PW_{\Omega,r} \longrightarrow PW_{\Omega,r}$ be defined by (2), where $\phi' > 0$ on $\mathbb{R}$, $h \in PW_{\Omega,r}$, and $|H| > 0$ a.e. on $[-\Omega, \Omega]$ for $\widehat{h} = H$. Then L is injective.*

3.2 The Banach fixed point theorem

Let X be a complete metric space with metric d. A function $A\colon X \mapsto X$ is a *contracting map* if there is $q \in [0,1)$ such that

$$\forall f_1, f_2 \in X, \quad d(Af_1, Af_2) \leq q d(f_1, f_2).$$

This is a Lipschitz condition. *Theorem 3.2* is the Banach fixed-point theorem[9]; and the proof is well-known and elementary.

3.2 Theorem. *Let $A\colon X \mapsto X$ be a contracting map on the complete metric space X with metric d.*

a. *If $f_0 \in X$ and $f_{n+1} = Af_n$ for each $n \geq 0$ then*

$$\forall k,\ n \geq 0, \quad d(f_{n+k}, f_n) \leq \frac{q^n}{1-q}\, d(f_1, f_0).$$

b. *The sequence $\{f_n\}$ converges to some f in X,*

$$\forall n \geq 1, \quad d(f, f_n) \leq \frac{q^n}{1-q}\, d(f_1, f_0),$$

and f is the unique solution of the equation $Ag = g$, i.e., f is the unique fixed point of A.

3.3 Existence theorem

The main result of *Section 3* is the following theorem.

3.3 Theorem. *Let $L\colon PW_{\Omega,r} \longrightarrow PW_{\Omega,r}$ be defined by (2), where $0 < m \leq \phi' \leq M$ on $\mathbb{R}$, $h \in PW_{\Omega,r}$, and $0 < \alpha \leq |H| \leq \beta$ a.e. on $[-\Omega, \Omega]$ for $\widehat{h} = H$. Then L is continuous[2], and L^{-1} is Lipschitz; in particular, if $\phi(0) = 0$ then there is $K > 0$ such that L and L^{-1} are Lipschitz:*

$$\forall f_1, f_2, \in PW_{\Omega,r}, \quad \|Lf_1 - Lf_2\|_2 \leq M\beta \|f_1 - f_2\|_2 \leq KM\beta \|Lf_1 - Lf_2\|_2.$$

In *Section 4* we consider the case $0 \leq |H| \leq \beta$ a.e. The difficult case, where $0 < |\phi'| \leq M$ and $\alpha \leq |H| \leq \beta$, is due to Saliani[12].

3.4 Outline of proof

The outline of our proof of *Theorem 3.3* seems a little idiosyncratic, but it works.

L is injective by *Theorem 3.1*, and the continuity follows from the general theory of superposition[2] or a direct calculation. We then make the hypotheses that H is real and

$$\frac{M-m}{M+m} < \frac{\alpha}{\beta}. \tag{3}$$

Assuming $\beta < 1$, we can use *Theorem 3.2* to see that L is surjective for this case. In fact, we show that there is a constant $c(\phi)$ such that

$$\forall f_0 \in PW_{\Omega,r} \text{ and } \forall S \in PW_{\Omega,r}, \quad Lf = S,$$

where

$$\lim_{n \to \infty} (f_n - c(\phi)(\phi \circ f_n) * h + c(\phi)\, S) = f \text{ in } PW_{\Omega,r}, \tag{4}$$

$$\forall n \geq 0, \quad f_{n+1} = Bf_n,$$

and

$$Bg = g - c(\phi)(\phi \circ g) * h + c(\phi)\, S.$$

A technical modification then allows us to obtain surjectivity for $\beta \geq 1$ in the case of (3). $c(\phi)$ is computable and (4) is an iterative means of computing $L^{-1}S$. The fact that L^{-1} is Lipschitz follows from an elementary calculation using (4).

This argument has as a consequence the Landau-Miranker theorem[10], which asserts the bijectivity of L for the filter $H = 1_{[-\Omega,\Omega]}$.

Finally, we use the Landau-Miranker theorem and another elementary calculation to complete the proof of *Theorem 3.3*.

4. A NONLINEAR SUBBAND CODING PROBLEM AND SOLUTION

4.1 Bit reversal

Bit reversal ordering arises and is required in the FFT algorithm because of decompositions of DFT computations into smaller and smaller DFT computations. Let $M = 2^m$, where $m \geq 0$. At level $m = 0$ consider the set $\{0\}$. At level $m = 1$ the bit reversal ordering of the set $\{0,1\}$ is the ordered 2^1-tuple $(0,1)$. At level $m = 2$ the bit reversal ordering of the set $\{0,1,2,3\}$ is the ordered 2^2-tuple $(0,2,1,3)$. Inductively, at level m suppose the set $\{0,1,\cdots,M-1\}$ has as its bit reversal ordering the ordered 2^m-tuple,

$$(b_0, b_1, \cdots, b_{M-1}).$$

Then, by definition, at level $m+1$, the *bit reversal ordering* of the set $\{0, 1, \cdots, 2M-1\}$ is the ordered 2^{m+1}-tuple

$$(2b_0, 2b_1, \cdots, 2b_{M-1}, 2b_0+1, 2b_1+1, \cdots, 2b_{M-1}+1).$$

For example, the bit reversal orderings at levels 3 and 4 are

$$0, 4, 2, 6, 1, 5, 3, 7$$

and

$$0, 8, 4, 12, 2, 10, 6, 14, 1, 9, 5, 13, 3, 11, 7, 15,$$

respectively. The terminology "bit reversal" is appropriate since the coefficients of the binary expansion of integers are reversed at the critical step in the above process.

Let $\{X_n^m\}$ be a tree of Banach spaces, where m designates the level, and where, for each fixed $m \geq 0$, there are $M = 2^m$ elements X_n^m, indexed by n. Using the binary expansion

$$n = \sum_{j=1}^{m} 2^{j-1} \epsilon_j, \quad \epsilon_j \in \{0, 1\}, \tag{5}$$

we write

$$X_n^m = X_{(\epsilon_1, \cdots, \epsilon_m)}^m;$$

and, using the bit reversal ordering, the tree $\{X_n^m\}$ has the form

$$X_0^0$$

$$X_0^1 \qquad X_1^1$$

$$X_{(0,0)}^2 \; X_{(0,1)}^2 \; X_{(1,0)}^2 \; X_{(1,1)}^2.$$

At level $m-1$ the space $X_{(\epsilon_1, \cdots, \epsilon_{m-1}, 0)}^{m-1}$ is the (single) "parent" of

$$X_{(\epsilon_1, \cdots, \epsilon_{m-1}, 0)}^m \text{ and } X_{(\epsilon_1, \cdots, \epsilon_{m-1}, 1)}^m.$$

For a given m and n, the generational synapse between "parent" and "child" is a surjective mapping denoted by

$$L_n^{m-1} = L_{(\epsilon_1, \cdots, \epsilon_m)}^{m-1} : X_{(\epsilon_1, \cdots, \epsilon_{m-1})}^{m-1} \longrightarrow X_n^m.$$

We define the tree $\{X_n^m\}$ and double sequence $\{L_n^m\}$ of mappings in the next section.

4.2 Frequency localization

Let $X_0^0 = PW_{\Omega,r}$ and suppose $\phi\colon \mathbb{R} \to \mathbb{R}$ is a differentiable function which satisfies the condition,

$$\exists m,\ M > 0 \text{ such that } \forall y \in \mathbb{R}, \quad m \le \phi'(y) \le M.$$

Let $\widehat{h}_0 = H_0$ and $\widehat{h}_1 = H_1$ be two filters, where $h_0, h_1 \in PW_{\Omega,r}$, and suppose

$$\exists \alpha,\ \beta > 0 \text{ such that } \alpha \le |H_0(\gamma)|^2 + |H_1(\gamma)|^2 \le \beta \text{ a.e.} \tag{6}$$

Generally, H_0 is a low pass filter and H_1 is a high pass filter on the band $[-\Omega, \Omega]$. The inequalities (6) extend the QMF condition, and play a critical role in the theory of multiresolution analysis frames[4].

The spaces X_0^1, X_1^1 at level 1 are determined by the mappings L_0^0, L_1^0, which are defined as

$$L_0^0 f = (\phi \circ f) * h_0 \text{ and } L_1^0 f = (\phi \circ f) * h_1,$$

respectively, i.e.,

$$X_0^1 = L_0^0\left(X_0^0\right) \text{ and } X_1^1 = L_1^0\left(X_0^0\right).$$

The goal is to decompose the frequency band $[-\Omega, \Omega]$ so that locally, for every sufficiently large level m, there is a "band pass" interval $I \subseteq [-\Omega, \Omega]$ which can be associated with the low pass filter H_0, and there are corresponding intervals which bound I and which can be associated with the high pass filter H_1. We proceed to quantify this point of view.

First, we define the inner and outer dilations D_i and D_o as follows for $f \in PW_{\Omega,r}$ and $c > 1$:

$$D_i\widehat{f} = D_c\widehat{f} \text{ and } D_o\widehat{f} = D_{\frac{2c}{c-1}}\left(\tau_{\Omega\left(\frac{c+1}{c-1}\right)}\widehat{f} + \tau_{-\Omega\left(\frac{c+1}{c-1}\right)}\widehat{f}\right),$$

where D_c is the L^2-normalized dilation defined in *Section 2*. D_i squeezes the support of $\widehat{f}$ into $[-\Omega/c, \Omega/c]$, and $D_o\widehat{f}$ squeezes the support of $\widehat{f}$ and sends copies "far away". Next, if $n \in \{0, 1, \cdots, 2^m - 1\}$ has the binary representation (5) and the subspace,

$$X^{m-1}_{(\epsilon_1,\cdots,\epsilon_{m-1})} \subseteq PW_{\Omega,r},$$

is specified, then we define X_n^m as

$$X_n^m = L^{m-1}_{(\epsilon_1,\cdots,\epsilon_m)}\left(X^{m-1}_{(\epsilon_1,\cdots,\epsilon_{m-1})}\right)$$

where

$$\left(L^{m-1}_{(\epsilon_1,\cdots,\epsilon_m)} f\right)^{\wedge} = (\phi \circ f)^{\wedge}(D_{\epsilon_1} D_{\epsilon_2} \cdots D_{\epsilon_{m-1}} H_{\epsilon_m}),$$

$$H_{\epsilon_j} = \begin{cases} H_0 & \text{if } \epsilon_j = 0 \\ H_1 & \text{if } \epsilon_j = 1, \end{cases}$$

and

$$D_{\epsilon_j} = \begin{cases} D_i & \text{if } \epsilon_j = 0 \\ D_o & \text{if } \epsilon_j = 1. \end{cases}$$

Using this splitting device, we proceed along various branches by means of the mappings L_n^m to obtain a *hemline* of subspaces corresponding to a desired bandsplitting of $[-\Omega, \Omega]$. We then reconstruct $f \in X_0^0$ from its frequency localized components along the hemline. Because of the nonlinearities in the mappings L_n^m, the results of *Section 3* are used in this reconstruction.

Before stating a special case of the reconstruction theorem we give the following example to illustrate the frequency localization procedure.

4.3 Example

Let $\widehat{h}_0 = H_0 = 1_{[-\Omega/2,\Omega/2]}$ and $\widehat{h}_1 = H_1 = 1_{[-\Omega,\Omega]} - H_0$, and let $c = 2$ in the definition of D_i and D_o. Thus, $D_i\widehat{f}(\gamma) = \sqrt{2}\widehat{f}(2\gamma)$ and

$$D_o\widehat{f}(\gamma) = 2(\widehat{f}(4\gamma - 3\Omega) + \widehat{f}(4\gamma + 3\Omega)).$$

The level $m = 0$ corresponds to the band $[-\Omega, \Omega]$ since $X_0^0 = PW_{\Omega,r}$. At level $m = 1$, we have the following correspondences between function space and the closed band of elements in it:

$$X_0^1 \left[-\frac{\Omega}{2}, \frac{\Omega}{2}\right], \quad X_1^1 \left[-\Omega, -\frac{\Omega}{2}\right] \bigcup \left[\frac{\Omega}{2}, \Omega\right]. \tag{7}$$

The union of the bands in (7) is all of $[-\Omega, \Omega]$. At level $m = 2$ we have the correspondences:

$$X_{(0,0)}^2 \left[-\frac{\Omega}{4}, \frac{\Omega}{4}\right], \quad X_{(0,1)}^2 \left[-\frac{\Omega}{2}, -\frac{\Omega}{4}\right] \bigcup \left[\frac{\Omega}{4}, \frac{\Omega}{2}\right] \tag{8}$$

and

$$X_{(1,0)}^2 \left[-\frac{3\Omega}{4} - \frac{\Omega}{8}, -\frac{3\Omega}{4} + \frac{\Omega}{8}\right] \bigcup \left[\frac{3\Omega}{4} - \frac{\Omega}{8}, \frac{3\Omega}{4} + \frac{\Omega}{8}\right],$$

$$X_{(1,1)}^2 \left[-\Omega, -\frac{3}{4}\Omega - \frac{\Omega}{8}\right] \bigcup \left[-\frac{3\Omega}{4} + \frac{\Omega}{8}, -\frac{\Omega}{2}\right] \bigcup \left[\frac{\Omega}{2}, \frac{3\Omega}{4} - \frac{\Omega}{8}\right]$$
$$\bigcup \left[\frac{3\Omega}{4} + \frac{\Omega}{8}, \Omega\right],$$

where once again the union of the bands in (8) is all of $[-\Omega, \Omega]$. The splitting in (8) is due to the fact that if $f \in X_0^1$ then

$$\left(L_{(0,0)}^1 f\right)^\wedge = (\phi \circ f)^\wedge D_i H_0 \quad \text{and} \quad \left(L_{(0,1)}^1 f\right)^\wedge = (\phi \circ f)^\wedge D_i H_1,$$

and if $f \in X_1^1$ then

$$\left(L^1_{(0,0)} f\right)^\wedge = (\phi \circ f)^\wedge D_o H_0 \text{ and } \left(L^1_{(1,1)} f\right)^\wedge = (\phi \circ f)^\wedge D_o H_1.$$

Thus, the *frequency support* I_n^m of the elements of X_n^m is a sequence of intervals. At the $m+1$ level, each subinterval J of I_n^m, i.e., each element of the union (of intervals) I_n^m, is split into two sets J_1 and J_2. J_1 is an interval which is a symmetric contraction of J. J_2 is the union of two intervals which "bound" this contraction (on the left and the right) and which themselves are contractions of J. Further, $J \subseteq J_1 \cup J_2$. This procedure provides locally high pass filters corresponding to locally low pass filters, and is the frequency localization of *Section 4.2.*

4.4 Frequency localization reconstruction

The following theorem gives signal reconstruction from the hemline X_0^1, X_1^1 consisting of the first level. (An example of a hemline at the second level is $X^2_{(0,0)}, X^2_{(0,1)}, X^1_1$.) There is a general reconstruction theorem[5] for arbitrary hemlines, and it is formulated in terms of our theory of nonlinear wavelet packets.

4.1 Theorem. *Let $\phi\colon \mathbb{R} \longrightarrow \mathbb{R}$ be a differentiable function for which*

$$\exists m,\ M > 0 \text{ such that } \forall y \in \mathbb{R}, \quad m \le \phi'(y) \le M.$$

Suppose $h_0, h_1 \in PW_{\Omega,r} = X_0^0$ have the properties that

$$\operatorname{supp} H_0 \subseteq \left[-\frac{\Omega}{a}, \frac{\Omega}{a}\right] \text{ and } \operatorname{supp} H_1 \subseteq \left[-\Omega, \Omega\right] \setminus \left(-\frac{\Omega}{b}, \frac{\Omega}{b}\right),$$

where $\widehat{h}_0 = H_0,\ \widehat{h}_1 = H_1,$ and $b \ge a > 1$, and that

$$\exists \alpha, \beta > 0 \text{ such that } \alpha \le H(\gamma) = |H_0(\gamma)|^2 + |H_1(\gamma)|^2 \le \beta \text{ a.e.}$$

a. *$H = \widehat{h}$ for some $h \in PW_{\Omega,r}$.*
b. *If we define the mapping,*

$$L\colon X_0^0 \longrightarrow X_0^0$$
$$f \mapsto (\phi \circ f) * h,$$

then L is a continuous bijection, L^{-1} is Lipschitz, and

$$\forall f \in X_0^0, \quad f = L^{-1}\left(\left(\left(L_0^0 f\right)^\wedge \bar{H}_0 + \left(L_1^0 f\right)^\wedge \bar{H}_1\right)^\vee\right),$$

where "$\vee$" designates the inverse Fourier transform.

REFERENCES

1. J. Allen, "Cochlear modeling," *IEEE-ASSP Magazine*, Vol. **2**(1), pp. 3-29, January 1985.

2. J. Appell and P. Zabrejko, *Nonlinear Superposition Operators*, Cambridge University Press, Cambridge, 1990.

3. J. Benedetto, "Wavelet auditory models and irregular sampling," *Prometheus Inc. Tech. Report*, January 1990.

4. J. Benedetto and S. Li, "The theory of multiresolution analysis frames and applications to filter construction," preprint.

5. J. Benedetto and S. Saliani, "Nonlinear wavelet packets," preprint.

6. J. Benedetto and A. Teolis, "A wavelet auditory model and data compression," *Applied and Computational Harmonic analysis*, Vol. **1**(1), December 1993.

7. P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, Vol. **31**, pp. 532–540, 1983.

8. R. Coifman, Y. Meyer and V. Wickerhauser, "Wavelet analysis and signal processing," *Wavelets and their Applications*, (M. B. Ruskai et al, eds.), Jones and Bartlett Publishers, Boston, 1992, pp. 153–178.

9. H. Heuser, *Functional Analysis*, John Wiley and Sons, New York, 1982.

10. H. Landau and W. Miranker, "The recovery of distorted band-limited signals," *J. Math. Analysis and Appl.*, Vol. **2**, pp. 97–104, 1961.

11. S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *Tech. Report*, 1992.

12. S. Saliani, "On the inversion of a Hammerstein operator in the space of bandlimited functions," preprint.

13. X. Xang, K. Wang and S. Shamma, "Auditory representations of acoustic signals," *IEEE Trans. on Information Theory*, Vol. **38**(2), pp. 824–839, March 1992.