Intrinsic wavelet and frame applications

John J. Benedetto and Travis D. Andrews

Norbert Wiener Center
Department of Mathematics
University of Maryland at College Park

ABSTRACT

There are *intrinsic* wavelet applications, by which we mean mathematical modeling of a physical phenomenon in which wavelet theory is the most natural quantitative means of explaining the phenomenon. This is not the same as the invaluable use of dyadic wavelets, say, as a tool with which to zoom-in or -out with regard to multi-scale phenomena. An example of an intrinsic wavelet application is wavelet auditory modeling (WAM). WAM is analyzed herein, and a natural excursion, one of many possibilities, is taken from WAM to applications of finite frames. This path includes the role of the Discrete Fourier Transform (DFT) in WAM, the emergence of DFT frames, and their use in analyzing $\Sigma\Delta$ quantization, which itself is a staple in audio engineering as well as in a host of other applications.

Keywords: Wavelet auditory modeling, DFT frames, $\Sigma\Delta$ quantization, perfect-correlation sequences

1. INTRODUCTION

1.1 Theme

Wavelet systems and their siblings, such as Gabor systems, and their extensions, such as frames, that are sometimes geometrically unstructured, can be harmonics for various signal decompositions. In the best of all worlds these harmonics are *intrinsic* to the mathematical modeling at hand. For example, in wavelet auditory modeling (WAM), the structure of the cochlea effectively dictates the use of wavelet systems for compression and noise reduction algorithms. Often the properties of the harmonics are not intrinsic, but provide *guidelines* for appropriate types of applicability, whether it be criteria for stability, periodicity detection, or classification algorithms for hyper-spectral imagery in terms of wavelet packets, see SPIE 2010 (San Jose) by Czaja, Ehler, Flake, Hirn, and the first named author for the last topic.

In the case of finite decomposition systems, i.e., finite frames, these guidelines are less obvious. On the other hand, the setting of finite frames is the most natural, and one can say intrinsic, for applicable mathematics in which viable computation is essential.

We shall begin with a specific intrinsic wavelet application, viz., WAM.

There are, of course, other intrinsic wavelet applications, but not nearly as many as those we categorized above in terms of "guidelines."

Our theme will be to follow a natural wavelet and frame path, from WAM in this case, to other recent techniques and ideas associated with wavelet theory. Our particular path will go in the direction of finite frames because of the importance of effective computation and the inherent stability associated with frames.

Before describing this path in our outline (Section 1.2), we want to make the following remark. It deals with the plethora of data and accompanying computational challenges which are a stimulant for developing new techniques that could possibly be based in a wavelet point of view.

REMARK 1.1. The early successes of wavelet theory, in its role as a technique and guideline for applicability, were rooted in the structure of multiresolution analysis (MRA), and, for the dyadic case, in its "zooming" property, e.g., [1], [2], [3]. In fact, for the latter, by zooming in and out by powers of 2 (or by any $n \geq 2$), we can fathom multi-scale phenomena in a signal and/or control computational costs vis-à-vis signal resolution in reconstruction. This zooming property can be thought of as a mathematical tool to accompany techniques such as non-linear spectral methods for dimension reduction in dealing with the challenge of managing massive data sets. These sets are generated by data-creating devices such as LI-DAR, hyper-spectral imagery, supercomputers, internet traffic, CAT scanners, and digital cameras. IDC estimates that the world generated 487 billion gigabytes of information in 2008. This creates formidable problems for obtaining digital representations suitable for storage, transmission, and/or recovery, as well as for handling informaiton accurately, efficiently, and robustly.

As is well-known, dyadic MRAs of $L^2(\mathbb{R}^d)$ require $2^d - 1$ wavelets, a fact which dampens the hope of using dyadic MRAs to deal with data sets such as the aforementioned. However, a theoretical solution

to this predicament is based in the fact that there are single dyadic orthonormal wavelets for $L^2(\mathbb{R}^d)$. Further, if rapid decay of the wavelet is desired, there are frame preserving smoothing results, see [4] for references dealing with single orthonormal or frame wavelets for $L^2(\mathbb{R}^d)$.

1.2 Outline

We shall describe our patented wavelet auditory model (WAM) developed with Teolis [5]. This is the subject of Section 2. The first goal of Section 2 is to illustrate the instrinsic role of wavelet theory in auditory modeling. The second goal is to illustrate some of the technical steps required in order to transition the wavelet auditory model into a computationally effective method for addressing certain compression and noise reduction problems. We shall emphasize the role of the discrete Fourier transform (DFT) in this second goal, even though the DFT is well-traveled and is well-understood at several levels.

This leads to Section 3 where we provide some details about the DFT and the computation and quantitative estimation of Fourier transforms in terms of DFTs. We do this, not only to see what is happening with regard to the computations of Section 2, but to introduce the notion of DFT frames for $\mathbb{C}^d$. Frames are a generalization of orthonormal bases (ONBs). DFT frames are a particularly auspicious frame for several applications. In Section 4 we use DFT frames for $\mathbb{C}^d$ to develop and compare quantization theories for low-bit Sigma-Delta ($\Sigma\Delta$) modulation and PCM, see [6], which builds on previous work of Powell, Yilmaz, and the first named author. The $\Sigma\Delta$ Theory is also the basis for our deterministic sparse representation theory, which can be useful for signal reconstruction in noisy environments.

Of course, there are wavelet and Gabor frames for $L^2(\mathbb{R}^d)$, general frames for $L^2(\mathbb{R}^d)$, and frames for other infinite dimensional spaces. For $\mathbb{C}^d$, DFT frames are a special case of finite unit norm tight frames (FUNTFs), which have a host of applications. These include the following: Robust transmission of data over erasure channels such as the internet (Casazza, Goyal, Kelner, Kovačevic); Multiple antenna code design for wireless communications (Hochwald, Marzetta, T. Richardson, Sweldens, Urbanke); Multiple description coding (Goyal, Heath, Kovačevic, Strohmer, Vetterli); Quantum detection (Bolcskei, Eldar, Forney, Oppenheim, Kebo, Benedetto).

The Appendix provides the requisite notation, definitions and mathematical background.

2. WAVELET AUDITORY MODEL (WAM)
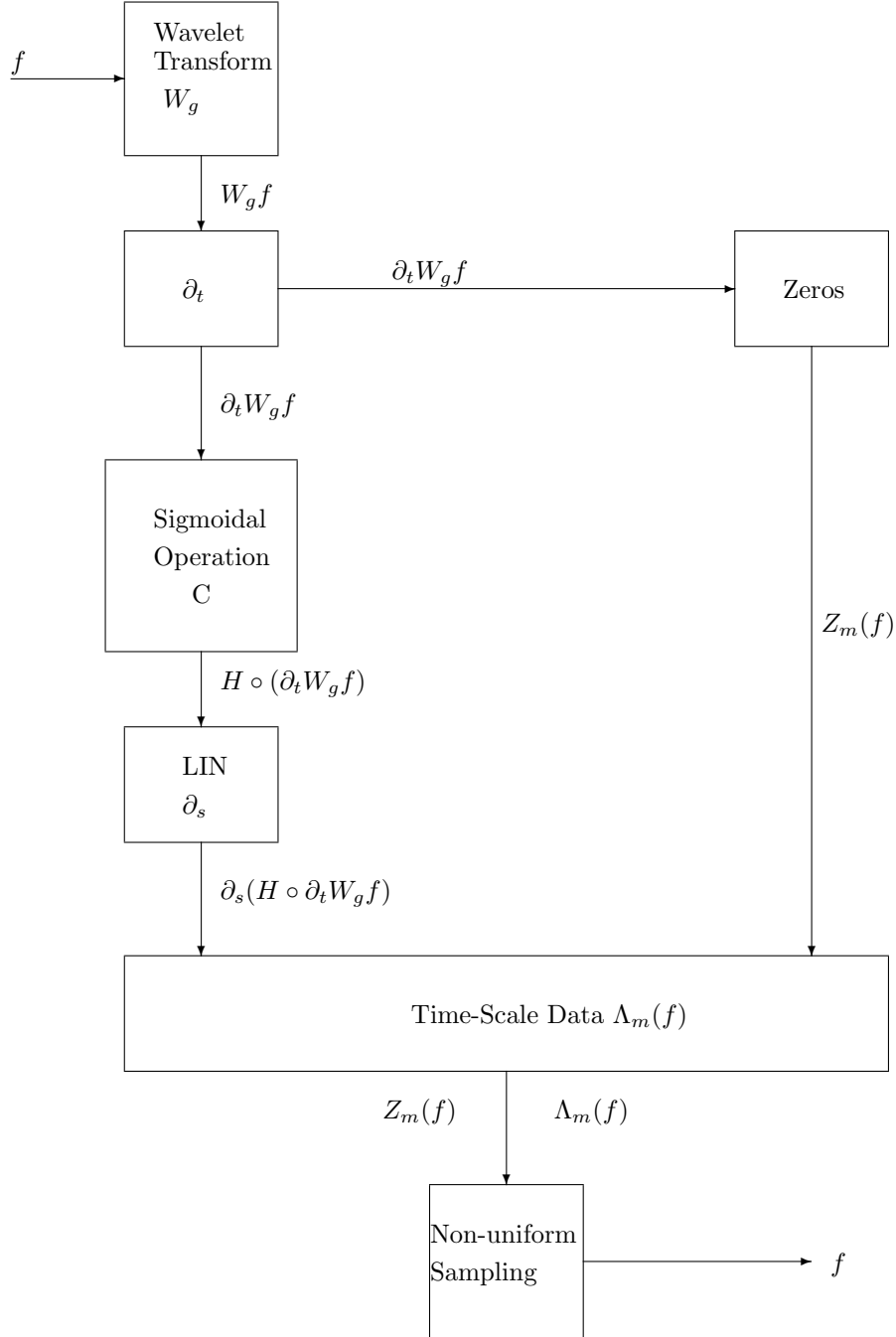
2.1 Setting for WAM

The mammalian auditory system possesses excellent abilities to detect, separate, and recognize speech and environmental sounds. In recent decades, these capabilities have been the subject of theoretical and experimental research, particularly with a view towards applying auditory functional principles in the design of man-machine communication links, e.g., [7], [8], [9], [10], [11]. We use frames and non-uniform sampling methods to construct a wavelet and Fourier frame based mathematical model for the mammalian auditory system. It is called the Wavelet Auditory Model (WAM).

A critical feature in the success of mammalian auditory systems is the fact that the cochlea has the equivalent of a sophisticated and effective filter bank on its basilar membrane.

In a mammalian auditory system, an acoustic signal or sound wave f induces vibrations in the ear drum, which travel through the middle ear to the cochlear fluid of the inner ear. These vibrations then cause traveling waves on the basilar membrane of the cochlea. As the waves propogate into the spiral shaped cochlea, they produce a pattern of displacements W of the basilar membrane at different locations for different frequencies. Displacements for high frequencies occur at the basal end; for low frequencies they occur at the wider apical end inside the spiral, see e.g., [9]. The basilar membrane records frequency responses between 200 and 20,000 Hz. For comparison, telephone speech bandwidth deals with the range 300-4,000 Hz. The cochlea analyzes sound in terms of these traveling waves much like a parallel bank of filters - in this case a band with 30,000 channels. The impulse responses of these filters along most of the interior length of the cochlea are related by dilation. Consequently, their transfer functions are invariant except for frequency translation along the approximately logarithmic axis of the cochlea, see e.g., [12], [13].

Mathematically, this dilational relationship between impulse responses can be expressed by the assertion that there is a function $g : \mathbb{R} \mapsto \mathbb{C}$ such that the set of impulse responses is of the form $\{D_s g : s \in [s_1, s_2] \subseteq (0, \infty)\}$, where $D_s g(t) = \sqrt{s} g(st)$. Thus, we identify the displacements W, due to the stimulus f, with the output of the cochlear filter bank having the impulse responses $\{D_s g\}$, i.e., we set $W = W_g f(t, s)$, where g is a fixed causal impulse response. $W_g f$ is the wavelet transform of f.

Figure 1. Wavelet Auditory Model

$$Z_m(f) = \{(t(n; s_m) : \partial_t W_g f(t(n; s_m)) = 0\}$$
$$\Lambda_m(f) = \{\partial_s \partial_t W_g f(t(n; s_m), s_m) : t(n; s_m) \in Z_m(f)\}$$

Because of the structure of the cochlear filter bank, we fix $a > 1$ and set $s_m = a^m, m \in \mathbb{Z}$. As such, in WAM the signal f is first transformed into a pattern of displacements,

$$W_g f(t, s_m), \quad m \in \mathbb{Z},$$

for a discrete set of points (t, s_m) in the time-scale plane, where t is written more explicitly in (1) below, see the box labelled Wavelet Transform in Fig. 1.

The shape of $|G| \equiv |\hat{g}|$ is critical for the effectiveness of the auditory model. Generally, G should be a causal filter that has a "shark-fin" shaped amplitude. There are solvable design problems for such filters. In the case of properly designed filters, the high frequency edges of the cochlear filters $D_{1/s}G$ act as abrupt "scale" delimiters. Thus, a sinusoidal stimulus will propogate up to the appropriate scale and die out beyond it.

The auditory system does not receive the wavelet transform directly, but rather a substantially modified version of it. In fact, the output of each cochlear filter is effectively highpassed by the velocity coupling between the cochlear membrane and the cilia of the hair cell transducers that initiate the electrical nervous activity by a shearing action on the tectorial membrane. Hence, the mechanical motion of the basilar membrane is converted to a receptor potential in the inner hair cells. It is reasonable to approximate this stage by a time derivative, obtaining the ouput $\partial_t W_g f(t, s_m)$, see the box labelled ∂_t in Fig. 1. The extrema of the wavelet transform $W_g f(t, s_m)$ become the zero-crossings of the new function $\partial_t W_g$; and so one output of the auditory process is

$$\forall m \in \mathbb{Z}, \quad Z_m = \{t(n; s_m) : \partial_t W_g f(t(n; s_m), s_m) = 0\}, \tag{1}$$

where n is used to index the domain of the zero-crossings for a given m, i.e., the box labelled Zeros in Fig. 1.

2.2 Sigmoidal Operation

At the next step in the auditory process, an instantaneous sigmoidal non-linearity R is applied, followed by a low pass filter with impulse response h. These operations model the threshold and saturation that occur in the hair cell channels and the leakage of electrical current throughout the membranes of these cells [14], [12]. The resulting cochlear output,

$$C_{h,R}(t, s) = (R \circ \partial_t W_g f(\cdot, s)) * h(t),$$

where "$\circ$" denotes composition of functions and where convolution $*$ is with respect to time, is a planar auditory nerve pattern sent to the brain. Typically, the composition by R can be represented by functions

$$R_T(x) = \frac{e^{Tx}}{1 + e^{Tx}} ,$$

parameterized by T. Obviously, $\lim_{T \to \infty} R_T = U$, the Heaviside step function. Approximations to the Heaviside step function are reasonable since the nerve fibers from the inner hair cells to the auditory nervous system fire at positive rates, and since this action cannot process above a certain limit, i.e., the aforementioned saturation. For computational convenience in WAM, we take R to be U and set $h = \delta$, the Dirac δ-measure (impulse), even though δ does not give rise to a low pass filter. Thus, $C_{h,R}(t, s)$ above is replaced by the *cochlear output,*

$$C(t, s) = U \circ \partial_t W_g f(t, s),$$

i.e., the output of the box labelled Sigmoidal Operation C in Fig. 1.

2.3 Lateral inhibitory network

The mammalian auditory nerve patterns determined by $W, \partial W$, and Z_m are now processed by the brain in ways that are not completely understood. One processing model, the lateral inhibitory network (LIN), has been closely studied with a view to extracting the spectral pattern of the acoustic stimulus [14], [13], [15]; and we shall implement it in our alogorithm. Scientifically, it reasonably reflects proximate frequency channel behavior, and mathematically it is relatively simple.

For a given acoustic signal f, constant $a > 1$, and properly designed causal filter G, we generate Z_m and the set

$$\Lambda_m = \{\partial_s \partial_t W_g f(t(n; s_m), s_m) : t(n; s_m) \in Z_m\} \tag{2}$$

for each $m \in \mathbb{Z}$, i.e., the box labelled Time-Scale Data $\Lambda_m(y)$ in Fig. 1. The elements of Λ_m are the coefficients of a natural linear expansion of $\partial_s(U \circ \partial_t W_g y)$, see page 5 of [5] for details. The scaling partial ∂_s reflects the action of LIN, indicated by the box in Fig. 1 called LIN ∂_s.

2.4 The WAM problem and solution

Let G be a properly designed finite energy causal filter and let $a > 1$. Suppose an unknown acoustic signal f has generated the set

$$\Lambda_m = \{\partial_s \partial_t W_g f(t(n; s_m), s_m) : t(n; s_m) \in Z_m\}$$

and that the receiver has knowledge of this set or of some subset. This non-uniformly spaced array in the $t - s$ plane is called *WAM data*. It is natural for the receiver to attempt to reconstruct the signal f in terms of this data. The *WAM problem* is to effect this reconstruction by means of non-uniform sampling formulas. The WAM problem can be solved theoretically by means of such formulas, see [5]. However, this leaves open the problem of effective implementation.

   With effective implementation in mind, we have revisited WAM with Baccala and Criner. One of our techniques is the tried and true DFT, but refurbished by using fine, computable error estimates in evaluating the DFT approximants of the Fourier transform. The background for this approach is given in Section 3. More substantively, and with regard to signal decomposition, we have gone in the direction of DFT frames. Their subtlety is reflected in a host of recent applications and theoretical results. For the latter, we mention the beautiful result of Hirn [16]. Our own interest in DFT frames is in the context of deterministic sparse coefficient signal representation in order to provide reliable signal representation in noisy environments. The background for this material is the $\Sigma\Delta$ theory of Section 4.

3. DISCRETE FOURIER TRANSFORM (DFT)

3.1 WAM and the DFT

The computations required to solve the WAM problem include the computation of $W_g f(t, s)$, $C(t, s)$, Z_m, and Λ_m defined in Section 2 and the Appendix. These computations are optimally made, both with regard to speed and precision, by means of the FFT in which the corresponding DFT can be formulated so that the infinite sums giving the coefficients of the DFT can be approximated accurately and in a straightforward way. The following subsection gives the background for this approximation.

3.2 DFT Computation of Fourier transforms

The following Classical Sampling Theorem goes back to papers by Cauchy (1840s), see [17], Theorem 3.10.10. It had a significant impact on various topics in mathematics, including number theory and interpolation theory, long before Shannon's application of it in communications.

THEOREM 3.1. *Classical Sampling Theorem*

   Let $T, \Omega > 0$ satisfy the condition that $0 < 2T\Omega \leq 1$, and let $s \in PW_{1/(2T)}$ satisfy the condition that $\hat{s} \equiv S = 1$ on $[-\Omega, \Omega]$ and $S \in L^\infty(\widehat{\mathbb{R}})$. Then

$$\forall f \in PW_\Omega, \qquad f = T \sum f(nT) \tau_{nT} s, \tag{3}$$

where the convergence in (3) is in $L^2(\mathbb{R})$ norm and uniformly in $\mathbb{R}$. A possible sampling function s is

$$s(t) = \frac{\sin 2\pi\Omega t}{\pi t}. \tag{4}$$

THEOREM 3.2. *DFT computation of Fourier transforms*

   Let $T, \Omega > 0$ satisfy the property that $2T\Omega = 1$, let $N \geq 2$ be an even integer, and let $f \in PW_\Omega \cap L^1(\mathbb{R})$. Consider the dilation $f_T(t) = Tf(Tt)$ as a continuous function on $\mathbb{R}$, as well as a function on $\mathbb{Z}$ defined by

$$f_T \quad : \quad \mathbb{Z} \to \mathbb{C}$$
$$m \to f_T[m],$$

where $f_T[m] \equiv f_T(m)$. Assume $f_T \in \ell^1(\mathbb{Z})$. For every integer $n \in \left[-\frac{N}{2}, \frac{N}{2}\right]$, we have

$$\widehat{f}\left(\frac{2\Omega n}{N}\right) = \widehat{f}\left(\frac{n}{NT}\right) = \sum_{m=0}^{N-1} (f_T)_N^0 [m]\, W_N^{mn}, \tag{5}$$

where $W_N = e^{-2\pi i/N}$ and $(f_T)_N^0[m] = \sum_{k \in \mathbb{Z}} f_T[m - kN]$.

   Proof. By the Classical Sampling Theorem with sampling function s in (4), we have

$$f = T \sum f(kT)\, \tau_{kT} s, \tag{6}$$

where the convergence is in $L^2(\mathbb{R})$ norm and uniformly on $\mathbb{R}$. Taking the Fourier transform of both sides of (6), we obtain

$$\widehat{f} = T \sum f(kT)\, e_{-kT} \mathbb{1}_{[-\Omega,\Omega]}, \tag{7}$$

where the convergence is in $L^2(\widehat{\mathbb{R}})$ and uniformly on $\widehat{\mathbb{R}}$.

   The integration required to obtain (7) from (6) requires justification.

   Because of (7) we have

$$\widehat{f}\left(\frac{2\Omega n}{N}\right) = T \sum_k f(kT)\, e^{-2\pi i kT 2\Omega n/N} \tag{8}$$

for every integer $n \in \left[-\frac{N}{2}, \frac{N}{2}\right]$. Thus, for such n,

$$
\begin{aligned}
\widehat{f}\left(\frac{2\Omega n}{N}\right) &= T \sum_k f(kT) e^{-2\pi i k n/N} \\
&= T \sum_k \sum_{p=kN}^{kN+N-1} f(pT)\, e^{-2\pi i pn/N} \\
&= T \sum_k \sum_{m=0}^{N-1} f((m+kN)T)\, e^{-2\pi i\left(\frac{mn}{N}+kn\right)} \\
&= \sum_{m=0}^{N-1} \left(\sum_k Tf((m+kN)T) \right) e^{-2\pi i mn/N},
\end{aligned}
\tag{9}
$$

and this is the right side of (5). The first equation in (9) follows from (8) and the hypothesis $2T\Omega = 1$. The second equality follows from regrouping in blocks of length N. The third equality is a consequence of the change of variables $p = m + kN$ for each fixed k. The final equality results since $f_T \in \ell^1(\mathbb{Z})$. □

   The effective use of Theorem 3.2 requires error estimates on each of the infinite series $(f_T)_N^0[m]$. This is the subject of a paper by Jacob Criner and the first named author, cf., [18] and [19].

3.3 DFT frames

DEFINITION 3.3. *A set $F = \{e_j\}_{j \in J} \subseteq \mathbb{F}^d$ is a frame for $\mathbb{F}^d$, $\mathbb{F}^d = \mathbb{R}$ or $\mathbb{C}$, if $\exists A, B > 0$ such that*

$$\forall x \in \mathbb{F}^d,\ A\|x\|^2 \le \sum_{j \in J} |\langle x, e_j \rangle|^2 \le B\|x\|^2$$

F is tight if $A = B$. A finite unit-norm tight frame F is a FUNTF.

EXAMPLE 3.4. If we form an $N \times d$ matrix using any d columns of the $N \times N$ Discrete Fourier Transform (DFT) matrix $(e^{2\pi ijk/N})_{j,k=0}^{N-1}$, then the rows of this $N \times d$ matrix, up to a multiplication by a proper constant, constitute a FUNTF for $\mathbb{C}^d$. In other words, if $\{k_1, \dots, k_d\} \subseteq \{0, 1, \dots, N-1\}$ and if we let

$$e_j = \frac{1}{\sqrt{d}} \left[e^{-2\pi ij \cdot k_1/N}, e^{-2\pi ij \cdot k_2/N}, \dots, e^{-2\pi ij \cdot k_d/N} \right], \tag{10}$$

then $\{e_j\}_{j=0}^{N-1}$ is a FUNTF for $\mathbb{C}^d$. In fact, for any $x \in \mathbb{C}^d$,

$$\sum_{j=0}^{N-1} |\langle x, e_j \rangle|^2 \;=\; \sum_{j=0}^{N-1} \sum_{l,l'=0}^{N-1} x[l+1]\overline{x[l'+1]}e_j[l]\overline{e_j[l']}$$

$$=\; \sum_{l,l'=0}^{N-1} x[l+1]\overline{x[l'+1]}\frac{1}{d}\sum_{j=0}^{N-1} e^{2\pi i j \cdot k_l/N} e^{-2\pi i j \cdot k_{l'}/N}$$

$$=\; \frac{N}{d}\sum_{l=0}^{N-1} |x[l+1]|^2$$

$$=\; \frac{N}{d}\|x\|^2.$$

In particular, when $k_l = l$ for $l = 1, \ldots, d$, we shall follow the notation of [20], and denote this particular harmonic frame by H_N^d, although the terminology "harmonic frame" is not specifically used there. These play a major role in the finite frame $\Sigma\Delta$ quantization of Section 4. An illustration for $N = 8$ and $d = 5$ follows.

$$\frac{1}{\sqrt{5}}\begin{pmatrix}
* & * & \cdot & \cdot & * & * & * & \cdot \\
* & * & \cdot & \cdot & * & * & * & \cdot \\
* & * & \cdot & \cdot & * & * & * & \cdot \\
* & * & \cdot & \cdot & * & * & * & \cdot \\
* & * & \cdot & \cdot & * & * & * & \cdot \\
* & * & \cdot & \cdot & * & * & * & \cdot \\
* & * & \cdot & \cdot & * & * & * & \cdot \\
* & * & \cdot & \cdot & * & * & * & \cdot
\end{pmatrix}$$

$$e_j = \frac{1}{5}\left[e^{2\pi i j \frac{1}{8}}, e^{2\pi i j \frac{2}{8}}, e^{2\pi i j \frac{5}{8}}, e^{2\pi i j \frac{6}{8}}, e^{2\pi i j \frac{7}{8}} \right]$$

$$j = 1, \ldots, 8.$$

PROPOSITION 3.5. *If F is a FUNTF for $\mathbb{F}^d$, then*

$$\forall x \in \mathbb{F}^d, \; x = \frac{d}{N}\sum_{j=1}^{N} \langle x, e_j \rangle e_j.$$

REMARK 3.6. Frames have many redeeming qualities. They offer redundant representation, can compensate for hardware errors, are inexpensive to implement, provide numerical stability, and can minimize effects of noise.

The vertices of platonic solids are FUNTFs; however, points that constitute FUNTFs do not have to be equidistributed, e.g., ONBs and Grassmanian frames. FUNTFs can be characterized as minimizers of a frame potential function analogous to Coulomb's Law. This characterization is due to Fickus and the first named author. Frame potential energy optimization has basic applications dealing with classification problems for hyperspectral and multi-spectral (biomedical) image data.

4. $\Sigma\Delta$ QUANTIZATION

4.1

4.2

Let $K \in \mathbb{N}$ and $\delta > 0$. Given the quantization alphabet

$$\mathcal{A}_K^\delta = \left\{ \left(k + \frac{1}{2} \right)\delta + i\ell\delta : k = -K, \ldots, K-1 \text{ and } \ell = -K, \ldots, K \right\},$$

consisting of $2K(2K+1)$ elements, we define the $2K(2K+1)$-*level uniform scalar quantizer* with stepsize δ by

$$Q(u) = \arg \min_{q \in \mathcal{A}_K^\delta} |u - q|_{\max},$$

where $|\cdot|_{\max}$ is defined by $|z|_{\max} = \max\{|Re(z)|, |Im(z)|\}$ for all $z \in \mathbb{C}$. Thus, $Q(u)$ is the element of the alphabet which is closest to u in the norm $|\cdot|_{\max}$. If there are at least two elements of $\mathcal{A}_K^\delta$ that are equally close to u, then we let $Q(u)$ be the element with the largest real part. If these equally closest elements have the same real parts, then we let $Q(u)$ be the element with the largest imaginary part.

DEFINITION 4.1. *Given $K \in \mathbb{N}, \delta > 0$, and the corresponding quantization alphabet $\mathcal{A}_K^\delta$ and the scalar quantizer Q with stepsize δ, let $\{x_n\}_{n=1}^N \subseteq \mathbb{C}$, and let p be a permutation of $\{1, \ldots, N\}$. The associated first order $\Sigma\Delta$ quantization is defined by the iteration*

$$
\begin{aligned}
u_n &= u_{n-1} + x_{p(n)} - q_n, \\
q_n &= Q(u_{n-1} + x_{p(n)}),
\end{aligned}
\tag{11}
$$

where u_0 is a specified constant. The first order $\Sigma\Delta$ quantizer produces the quantized sequence $\{q_n\}_{n=1}^N$, and an auxiliary sequence $\{u_n\}_{n=0}^N$ of state variables.

DEFINITION 4.2. *Let $F = \{e_n\}_{n=1}^N$ be a finite frame for $\mathbb{C}^d$, and let p be a permutation of $\{1, \ldots, N\}$. We define the variation of the frame F with respect to p as*

$$
\sigma(F, p) = \sum_{n=1}^{N-1} \|e_{p(n)} - e_{p(n+1)}\|.
$$

THEOREM 4.3. *Given the $\Sigma\Delta$ scheme of Definition 4.1. Let $F = \{e_n\}_{n=1}^N$ be a unit-norm tight frame for $\mathbb{C}^d$ and assume that F satisfies the zero sum condition*

$$
\sum_{n=1}^N e_n = 0.
\tag{12}
$$

Let p be a permutation of $\{1, \ldots, N\}$ and let $x \in \mathbb{C}^d$ satisfy $\|x\| \leq (K - 1/2)\delta$. Additionally, set $u_0 = 0$ in (11). Then the approximation error $\|x - \widetilde{x}\|$ satisfies

$$
\|x - \widetilde{x}\| \leq
\begin{cases}
\sqrt{2}\dfrac{\delta d}{2N}\sigma(F, p), & \text{if } N \text{ is even;} \\
\sqrt{2}\dfrac{\delta d}{2N}\left(\sigma(F, p) + 1\right), & \text{if } N \text{ is odd.}
\end{cases}
\tag{13}
$$

EXAMPLE 4.4. We shall show that the harmonic frames of Example 3.4 have uniformly bounded frame variation. Let $\{k_1, \ldots, k_d\} \subseteq \{0, 1, \ldots, N-1\}$, and let $\{e_j\}_{j=0}^{N-1}$ be as in (10). Then, for $j = 0, \ldots, N-2$, we have

$$
\begin{aligned}
\|e_j - e_{j+1}\|^2 &= \frac{1}{d}\sum_{l=1}^d |e^{-2\pi i k_l j/N} - e^{-2\pi i k_l (j+1)/N}|^2 \\
&= \frac{1}{d}\sum_{l=1}^d \left(e^{-2\pi i k_l j/N} - e^{-2\pi i k_l (j+1)/N}\right)\left(e^{2\pi i k_l j/N} - e^{2\pi i k_l (j+1)/N}\right) \\
&= \frac{1}{d}\sum_{l=1}^d 2 - \left(e^{2\pi i k_l/N} + e^{-2\pi i k_l/N}\right) \\
&= \frac{2}{d}\sum_{l=1}^d 1 - \cos\frac{2\pi k_l}{N} = \frac{4}{d}\sum_{l=1}^d \sin^2\frac{\pi k_l}{N} \\
&\leq \frac{4}{d}\sum_{l=1}^d \left(\frac{\pi k_l}{N}\right)^2 = \frac{4\pi^2}{dN^2}\cdot\sum_{l=1}^d k_l^2
\end{aligned}
$$

Thus, if p is the identity permutation, then

$$
\sigma(\{e_j\}_{j=0}^{N-1}, p) = \sum_{j=0}^{N-2} \|e_j - e_{j+1}\| \leq (N-1)\frac{2\pi}{N\sqrt{d}}\left(\sum_{l=1}^d k_l^2\right)^{1/2} \leq \frac{2\pi}{\sqrt{d}}\left(\sum_{l=1}^d k_l^2\right)^{1/2}.
\tag{14}
$$

In particular,

$$\sigma(H_N^d, p) \leq \frac{2\pi}{\sqrt{d}} \left(\sum_{l=1}^{d} l^2 \right)^{1/2} = \frac{2\pi}{\sqrt{d}} \sqrt{\frac{d(d+1)(2d+1)}{6}} \leq \frac{2\pi}{\sqrt{3}}(d+1).$$

Harmonic frames also satisfy (12), if none of the k_l is equal to zero. In fact, since $k_l \neq 0$, $e^{-2\pi i k_l/N} \neq 1$ for each $l = 1, \dots, d$. Then,

$$\sum_{j=0}^{N-1} e_j[l] = \sum_{j=0}^{N-1} e^{-2\pi i k_l j/N} = \frac{1 - e^{-2\pi i k_l j}}{1 - e^{-2\pi i k_l/N}} = 0$$

for every $l = 1, \dots, d$.

We can now derive error estimates for $\Sigma\Delta$ quantization of harmonic frames for the identity permutation. If we set $u_0 = 0$ and assume that $x \in \mathbb{C}^d$ satisfies $\|x\| \leq (K - 1/2)\delta$, then combining Theorem 4.3, (14), and the fact that harmonic frames satisfy (12) gives

$$\|x - \widetilde{x}\| < \begin{cases} \sqrt{2}\frac{\delta d}{2N}\frac{2\pi}{\sqrt{d}}C_d, & \text{if } N \text{ is even}, \\ \sqrt{2}\frac{\delta d}{2N}\left(\frac{2\pi}{\sqrt{d}}C_d + 1\right), & \text{if } N \text{ is odd}, \end{cases}$$

where $C_d = \sqrt{\sum_{l=1}^{d} k_l^2}$

In [21], Wang proves the following theorem, and asserts that the proof leads to an algorithm for finding an ordering that reduces the frame variation for frames for $\mathbb{R}^d$.

THEOREM 4.5. *Let $S = \{v_j\}_{j=1}^{N} \subseteq [-\frac{1}{2}, \frac{1}{2}]^d$ with $d \geq 3$. There exists a permutation p of $\{1, \dots, N\}$ such that*

$$\sum_{j=1}^{N-1} \|v_{p(j)} - v_{p(j+1)}\| \leq 2\sqrt{d+3}N^{1-\frac{1}{d}} - 2\sqrt{d+3}.$$

COROLLARY 4.6. *Let $F = \{e_j\}_{j=1}^{N}$ be a unit norm frame for $\mathbb{F}^d$, $d \geq 3$. There exists a permutation p of $\{1, \dots, N\}$ such that*

   i. *if $\mathbb{F} = \mathbb{R}$, then $\sigma(F, p) \leq 4\sqrt{d+3}\ N^{1-\frac{1}{d}} - 4\sqrt{d+3}$,*

   ii. *if $\mathbb{F} = \mathbb{C}$, then $\sigma(F, p) \leq 4\sqrt{2d+3}\ N^{1-\frac{1}{2d}} - 4\sqrt{2d+3}$.*

THEOREM 4.7. *Let $F = \{e_n\}_{n=1}^{N}$ be a FUNTF for $\mathbb{C}^d$, $|u_0|_{\max} \leq \delta/2$, and let $x \in \mathbb{C}^d$ satisfy $\|x\| \leq (K - 1/2)\delta$. Then, there exists a permutation p of $\{1, \dots, N\}$ such that the approximation error $\|x - \widetilde{x}\|$ satisfies*

$$\|x - \widetilde{x}\| \leq \sqrt{2}\delta d \left((\frac{1}{2} - 2\sqrt{2d+3})N^{-1} + 2\sqrt{2d+3}N^{-\frac{1}{2d}} \right).$$

APPENDIX A.

The Fourier Transform of the function $f : \mathbb{R}^d \to \mathbb{C}$ is formally defined by

$$F(\gamma) = \int f(x)e^{-2\pi i x\gamma}\, dx,$$

where $\int$ denotes integration over $\mathbb{R}^d$; and the inverse Fourier transform $F^\vee$ of F: $\widehat{\mathbb{R}}^d \to \mathbb{C}$ is formally defined by

$$F^\vee(x) = \int F(\gamma)e^{2\pi i x\gamma}\, d\gamma, \text{ for } x \in \mathbb{R}^d.$$

For a fixed $g \in L^2(\mathbb{R})$, the *wavelet transform* of $f \in L^2(\mathbb{R})$ is the function

$$W_g f(t, s) = (f * D_s g)(t) \tag{15}$$

defined on the *time-scale plane $t \in \mathbb{R}$, $s > 0$.* By a straightforward calculation, we obtain

$$W_g f(t,s) = \langle f, \tau_t D_s \tilde{g} \rangle,$$

where $\tilde{g}$ is the *involution* of g defined as $\tilde{g}(u) = \overline{g}(-u)$. If the derivative $\partial_t g$ is an element of $L^2(\mathbb{R})$, we define $W_{\partial_t g} f$ analogously to the definition of $W_g f$ in (15). In this case, if $W_{\partial_t g} f$ converges uniformly on time intervals for each fixed scale $s > 0$, and if a mild smoothness condition is satisfied, then

$$\partial_t W_g f(t,s) = s W_{\partial_t g} f(t,s). \tag{16}$$

(16) is true for the causal filters G and signals f under consideration in this paper.

In particular, $L^2[-\Omega, \Omega]$ is the space of finite energy signals defined on the interval $[-\Omega, \Omega]$, and PW_Ω is the *Paley-Wiener space* (bandlimited finite energy signals), defined as

$$PW_\Omega = \{ f \in L^2(\mathbb{R}) : supp\, F \subseteq [-\Omega, \Omega] \},$$

where $supp\, F$ is the support of F.

ACKNOWLEDGMENTS

REFERENCES

[1] Y. Meyer, *Ondelettes et Opérateurs*, Hermann, Paris, 1990.

[2] I. Daubechies, *Ten lectures on wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.

[3] S. Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*, Academic Press, 2008.

[4] J. J. Benedetto and R. L. Benedetto, "The Construction of Wavelet Sets," in *Wavelets and Multiscale Analysis: Theory and Applications*, J. Cohen and A. I. Zayed, eds., *Applied and Numerical Harmonic Analysis*, ch. 2, Birkhäuser, Boston, 2011.

[5] J. J. Benedetto and A. Teolis, "A wavelet auditory model and data compression," *Applied and Computational Harmonic Analysis* **1**(1), pp. 3 – 28, 1993.

[6] J. J. Benedetto and O. Oktay, "Pointwise comparison of PCM and $\Sigma\Delta$ quantization," *Constructive Approximation* **32**, pp. 131–158, 2010. 10.1007/s00365-010-9083-x.

[7] D. Childers et al., "The past, present, and future of speech processing," *IEEE Signal Processing Magazine*, pp. 24–48, May 1998.

[8] J. Cohen, "Application of an auditory model to speech recognition," *J. Acoust. Soc. Amer* **85**, pp. 2623–2629, 1989.

[9] S. Greenberg, "Acoustic transduction in the auditory periphery," *Journal of Phonet* **16**, pp. 3–18, 1988.

[10] S. Shamma, "Auditory cortical representation of complex acoustic spectra as inferred from the ripple analysis method," *Network: Computation in Neural Systems* **7**, pp. 439–476, 1996.

[11] A. Waibel and K. F. Lee, *Readings in Speech Recognition*, Morgan Kaufmann Publishers, Inc, San Mateo, CA, 1990.

[12] S. Shamma, "Speech processing in the auditory system, i. the representation of speech sounds in the responses of the auditory nerve," *J. Acoust. Soc. Amer* **78**, pp. 1612–1621, 1985.

[13] S. Shamma, "Speech processing in the auditory system, ii. lateral inhibition and the central processing of speech evoked activity in the auditory nerve," *J. Acoust. Soc. Amer* **78**, pp. 1622–1632, 1985.

[14] I. Morishita and A. Yajima, "Analysis and simulation of networks of mutually inhibiting neurons," *Kybernetik* **11**, pp. 154–165, 1972.

[15] X. Yang, K. Wang, and S. A. Shamma, "Auditory representations of acoustic signals," *IEEE Trans. Inform. Theory* **38**, pp. 824–839, March 1992.

[16] M. Hirn, "The number of harmonic frames of prime order," *Linear Algebra and its Applications* **432**(5), pp. 1105 – 1125, 2010.

[17] J. J. Benedetto, *Harmonic Analysis and Applications*, CRC Press, 1996.

[18] L. Auslander and F. A. Grünbaum, "The Fourier transform and the discrete Fourier transform," *Inverse Problems* **5**, pp. 149–164, 1989.

[19] C. Epstein, "How well does the finite Fourier transform approximate the Fourier transform?," *Communications on Pure and Applied Mathematics* **58**, pp. 1421–1435, 2005.

[20] G. Zimmermann, "Normalized tight frames in finite dimensions," in *Recent Progress in Multivariate Approximation*, K. Jetter, W. Haussmann, and M. Reimer, eds., Birkhäuser, Boston, 2001.

[21] Y. Wang, "Sigmadelta quantization errors and the traveling salesman problem," *Advances in Computational Mathematics* **28**, pp. 101–118, 2008. 10.1007/s10444-006-9016-1.