Lossless Data Compression of Monitored Power Signals Using the PLHaar Transform

A Thesis Presented to
the Faculty of the Graduate School
at the University of Missouri–Columbia

In Partial Fulfillment
of the Requirements for the Degree
Master of Science

By
Nicholas J. Hudson

Dr. Michael J. Devaney, Thesis Supervisor
July 2015
The undersigned have examined the thesis entitled

**Lossless Data Compression of Monitored Power Signals Using the PLHaar Transform**

presented by Nicholas Hudson and determined it worthy of acceptance.

__________________________
Professor Michael J. Devaney

__________________________
Professor Zhihai "Henry" He

__________________________
Professor Jeffrey Uhlmann
for the old man
Acknowledgements

I’ve unmeasurable gratitude to my advisor Dr. Michael Devaney. Without his boundless patience, breadth of knowledge, excellent advice, and boundless patience not only would this work have been not just impossible but unimaginable to me. I’m both deeply honored and fantastically fortunate to be his student.

I’d also like to thank Dr. James Keller, who not only introduced me to the non-Euclidean spaces fundamental to this work, but also extended my education with a philosophical element of value nigh-impossible to overstate.

And I’m deeply indebted to my mother and brother, whose faith in me served as bedrock to this research.
Contents

Acknowledgements ii

Contents iii

List of Figures v

List of Tables vi

1 Introduction 1

1.1 Motivation ....................................................... 1
1.2 Relevance ...................................................... 2
1.3 Proposed Approach ............................................ 4
1.4 Organization .................................................... 4

2 Background 5

2.1 History .......................................................... 5
2.1.1 Fourier ...................................................... 5
2.1.2 Windowed Transforms ...................................... 5
2.1.3 First-Generation Wavelets .................................. 6
2.1.4 Second-Generation Wavelets ............................... 7
2.2 Literature Review ............................................... 8
2.2.1 Algebraic Wavelets ........................................ 8
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.2</td>
<td>Perfect Reconstruction Wavelets</td>
<td>9</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Dynamic Range</td>
<td>10</td>
</tr>
<tr>
<td>2.2.4</td>
<td>Infinity-norm Rotation</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>Theory</td>
<td>12</td>
</tr>
<tr>
<td>3.1</td>
<td>The Infinity Norm</td>
<td>12</td>
</tr>
<tr>
<td>3.2</td>
<td>Additional Techniques</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>Experiments</td>
<td>18</td>
</tr>
<tr>
<td>4.1</td>
<td>Data</td>
<td>18</td>
</tr>
<tr>
<td>4.2</td>
<td>Validation</td>
<td>19</td>
</tr>
<tr>
<td>4.3</td>
<td>Transformed</td>
<td>19</td>
</tr>
<tr>
<td>4.4</td>
<td>Results</td>
<td>20</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Signal 1 (Voltage)</td>
<td>20</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Signal 2 (Amps)</td>
<td>22</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Signal 3 (Amps)</td>
<td>22</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Signal 4 (Amps) S-Transform Failure</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>Conclusions and Summary</td>
<td>25</td>
</tr>
<tr>
<td>5.1</td>
<td>Further Musings</td>
<td>25</td>
</tr>
<tr>
<td>5.2</td>
<td>Conjectures and Open Questions</td>
<td>26</td>
</tr>
<tr>
<td>5.3</td>
<td>Conclusion</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Reference List</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Appendices</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>A Mathematical Preliminaries</td>
<td>44</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Unit Circles for $L^2, L^4, L^\infty$, rotation by $\frac{\pi}{6}$ radians on $L^2$, rotation by $2 - \frac{\sqrt{3}}{2}$ from $\frac{\sqrt{3}}{2}$ on $L^\infty$.</td>
<td>12</td>
</tr>
<tr>
<td>4.1</td>
<td>Signal 1 Native (in green) and several bands of PLHaar (left,black) and S-Transform(right,black)</td>
<td>20</td>
</tr>
<tr>
<td>4.2</td>
<td>Signal 1 Decomposed: green encoded natively, black with Exponential-Golomb, PLHaar left, S-Transform right</td>
<td>21</td>
</tr>
<tr>
<td>4.3</td>
<td>Signal 1 Native (in red) and several bands of PLHaar (left,black) and S-Transform(right,black)</td>
<td>21</td>
</tr>
<tr>
<td>4.4</td>
<td>Signal 1 Decomposed Green encoded natively, black with Exponential-Golomb, PLHaar left, S-Transform right</td>
<td>22</td>
</tr>
<tr>
<td>4.5</td>
<td>Signal 3 Native (in red) and several bands of PLHaar (left,black) and S-Transform(right,black)</td>
<td>23</td>
</tr>
<tr>
<td>4.6</td>
<td>Signal 3 Decomposed: green encoded natively, black with Exponential-Golomb, PLHaar left, S-Transform right</td>
<td>23</td>
</tr>
<tr>
<td>4.7</td>
<td>Signal 4 Native (in blue) and several bands of PLHaar (left,black) and S-Transform(right,black)</td>
<td>24</td>
</tr>
<tr>
<td>4.8</td>
<td>Signal 4 Decomposed: blue encoded natively, black with Exponential-Golomb, PLHaar left, S-Transform right</td>
<td>24</td>
</tr>
</tbody>
</table>
List of Tables

4.1 Dataset Summary .................................................. 19
4.2 Compression Format .............................................. 20
4.3 Signal 1 Results Table ........................................... 20
4.4 Signal 1 Results Table ........................................... 21
4.5 Signal 3 Results Table ........................................... 22
4.6 Signal 4 Results Table, * is invalid because S Transform Failed ........ 23
Chapter 1    Introduction

Abstract

We develop a lossless data compression technique for power quality monitoring signals, using the PLHaar and S transforms and Exponential-Golomb encoding. We also prove that rotations in $l^2$ may be embedded into $l^\infty$ even in the discrete case, and thus all orthonormal wavelets have of $l^2$ have an equivalent under the infinity norm (max).

1.1 Motivation

Monitoring power signals to ensure has great importance to industry; high-power devices operating with sub-par Voltage or Amperage waveforms can cause damage, and a machine’s interaction with the network can provide information relevant to the need for maintenance \[137\]. In larger networks common to industrial operations a multitude of monitoring systems may often be invoked, with one or more centralized monitoring stations which can react to unexpected problems. Unfortunately, these extensive power networks often fail in a cascading fashion, which can put immense strain on the corresponding communications network of monitoring devices. Additionally, in situations where longterm data retention is required, the storage requirements add up quickly particularly since high-powered equipment commonly runs in three-phase.

For lossless data compression of signals, one may use the minimum description length principle (MDL) \[73\][107] to simultaneously encode a model as well as the data. Wavelet
analysis with MDL achieved a reasonable measure of success in lossy compression of power quality disturbance data by means of a dictionary of wavelets, decomposition structures, and quantization [117]. A typical (orthonormal) wavelet transform acts as a linear change of basis for the space of finitely square-Lebesgue integrable functions, denoted \( L^2 \) for continuous spaces and \( l^2 \) when discrete, and this measure remains invariant (up to a multiplicative normalization constant usually applied during the reverse) in transform space [104,132]. This invariance condition potentially swells the dynamic range of a signal, such as the scenario in which an \( n \)-dimensional vector of identical values \( x \) is transformed to a vector with exactly one nonzero entry \( |y| = \sqrt{n}x = \sqrt{n}|x| \), clearly expanding absolute maximum of the signal by \( \sqrt{n} \). As noted in [147] this can cause difficulty in systems with severe constraints on bit depth, but the motivating problem is related to the models used to encode the high frequency bands often quantized away as noise.

In many cases, the high frequencies generally modeled very well as a zero mean, IID symmetric probability distribution, with the exception of a few outlying regions or samples. A common situation is a distribution approximately \( \mathcal{N}(0,\sigma^2) \), but with outliers \( > 5\sigma \) stretching the capabilities of double-precision arithmetic to approximate the value, necessitating escape codes. Though a common phenomenon, these codes complicate the MDL optimization step, and seemed a rather inelegant bandage over a deeper algorithmic misdesign.

We expand upon the infinity rotations of Yang and Hao [147] by demonstrating its deep connection with orthonormal wavelets.

### 1.2 Relevance

Harmonic analysis buttresses much of signal processing, and blends splendidly with information theory providing numerous tools highly suited to data compression [53]. Fourier techniques provide excellent bases for steady-state data, and by though local windows may be extended and tailored to transient events. Classical wavelets consist of orthogonal families of
functions exhibiting (potentially infinitely) recursive, self-similar subspaces with translated and dilated kernels forming a Riez basis over the $L^2(\mathbb{C}^n)$. So called “Second-generation wavelets” extend this family to more generalized spaces, and the lifting scheme greatly facilitated development of adaptive transforms. Algebraic structure of discrete wavelets has been of some small interest lately, providing opportunities to study perfect reconstruction, which is this work’s primary focus.

Whereas classical wavelets arose from analysis, the lifting formulation of wavelet transformations is decidedly algebraic making it a more natural fit for computational applications. In addition to yielding an efficient filtering technique perfectly compatible with the Mallat algorithm, through judicious application of parameters and rounding any classical or second-generation wavelet may be lifted into the set of integers achieving perfect bit-fidelity after recomposition by the (lifted) dual transform without increasing the dynamic range of the transformed representation. This overcomes a drawback of classical wavelets: the floating point error arising from the filter coefficients precluded the application of most wavelets to lossless data compression. Even further generalizations of lifting have been developed, and there’s been steady progress towards a rigorous linear algebraic theory of the techniques.

Although far friendlier to the fidelity-minded than the floating point predecessors, integer-to-integer transformations lifted or expanded from wavelets on $l^2(\mathbb{R}^n)$ inherit a subtle pitfall: as isometries of square integrable functions spaces, orthonormal wavelets run the risk of integer overflow. Even worse, lifting an orthogonal wavelet to integers almost certainly yields a biorthogonal wavelet expanding the domain of overflow. The straightforward approach of operating at a redundant bit depth suffices more often than not, but in extreme conditions such non-terrestrial systems every bit costs. I also assert with neither evidence nor argument that transformations mathematically incapable of overflow of interest to critical domains such as vehicle control, medical devices, secure software, and hardware design.

Subsequently is the introduction a class of finite dimensional $l^\infty$-isometries, demonstration these transforms exhibit isomorphism to normalized Givens rotations, and hence may
be realized through lifting. Further, we show in the discrete case these transforms host an embedding from every normalized Euclidean rotation, and how this makes more efficient use of the bit depth. Then experimental evidence of its applicability to one dimensional lossless signal compression is presented.

1.3 Proposed Approach

We prove isomorphism between \( l^\infty \) and \( l^2 \) (Givens) rotations, and subsequently homomorphism between cascades of \( l^\infty \)-rotations and orthogonal wavelets. Then we compare the results of a simple data compression algorithm using the \textit{Piecewise-Linear Haar (PLHaar)} \cite{125} transform and the \textit{S-Transform} \cite{26}.

1.4 Organization

In the subsequent chapter \cite{2} we cover much of the history of wavelets, with special attention paid to the development of perfect reconstruction techniques and the recent developments in algebraic representations of the transforms. In chapter \cite{3} we prove a number of assertions about the relationship between \( l^2 \)-orthonormal wavelets and \( l^\infty \)-isometries, and introduce some techniques used in chapter \cite{4} for experiments in lossless compression of power quality monitoring signals. Finally, chapter \cite{5} considers the greater picture surrounding this work.
2.1 History

The development of modern wavelet theory has hundreds of years of history, yet only in the last couple decades have the techniques bloomed into the powerful and flexible multiresolution tool chest we have today.

2.1.1 Fourier

In the 1807, after investigating the the physical properties of heat diffusion, Joseph Fourier asserted by appeal to natural philosophy that an arbitrary function converged from a trigonometric series [87]. Though overzealous, the conjecture proved true for a large set of functions as demonstrated by Direchlet [98] and Riemann [119]. In 1966 Carlson demonstrated convergence for $L^p$ [27] followed 2 years later by Hunt extending the results for $L^p : p \geq 1$ [81]. The series and transform engendered by this observation still bear Fourier’s name, and over 200 years later remain indispensible tools in a plethora of fields. The efficient implementation yielded by the Fast Fourier Transform [36,37] solidified Fourier’s beliefs an exalted place in the computational analysis tool chest for the foreseeable future.

2.1.2 Windowed Transforms

Although pervasive, Fourier in the strict sense abandons any time-localized information, a weakness which Gabor remedied in 1946 using a parameterized Gaussian envelope yielding
the transform which still bears his name to this day \cite{66}. The greater class of functions generated by application of localized functions to oscillatory functions are known as \textit{windowed transforms}. This includes cosine-modulated filter banks, of which the discrete cosine transform introduced by Ahmed et al. \cite{9} achieved ubiquity through inclusion in original JPEG standard \cite{116}.

\section{First-Generation Wavelets}

Rather than induce a time-frequency response through selection and combination of local and non-local functions, wavelets are traditionally generated by a \textit{scaling function} orthogonal to a \textit{mother wavelet}, and the translations and dyadic dilations of these functions are orthogonal kernel functions forming a complete basis of the space of square-integrable functions $L^2$; the self-similarity and recursive nature of the functions provide a natural framework for multiresolution decomposition. Bits and pieces of wavelet theory have appeared throughout the last hundred years or so, including work as old as Haar’s series of square waves published in 1910 \cite{74}, which although a relatively simple basis still finds use to this day. While the term “wavelet” was applied to smooth transient events in a variety of fields throughout the 1970’s \cite{31}, the perfect reconstruction filters introduced in 1976 by Croisier et al. \cite{38} subsequently called \textit{quadrature mirror filters} (QMFs) provided a tantalizing preview of the power of the decimation features of discrete wavelet families to come. In 1982 \cite{87} Morlet developed a modified version of the Gabor transform with superior localization properties \cite{87}, and he and Grossman introduced an analytic transform \cite{72} now known as the Morlet wavelet.

In 1983 Strömberg established the existence of orthonormal spline wavelet bases of Hardy spaces \cite{133,134}, a fact unknown to the small community of wavelet researchers centered around Grossman in Marseille, but in 1986 Yves Meyer stumbled accidentally into construction of Hilbertian basis with optimal localization in time and frequency \cite{106}. The maxflat filters of Ingrid Daubechies published in 1988 \cite{40} provided a set of FIR QMFs with
octave-band characteristics familiar to those accustomed to Fourier techniques as well as with compact support and energy invariance suited to extraction of transient information. In the middle 1980’s Mallat applied the recently developed Laplacian pyramidal algorithm [23] to QMFs [59, 128] and with the assistance of Meyer presented the theory behind multiresolution decomposition and the algorithm for fast wavelet transformations with structure similar to Cooley’s fast Fourier transform [102, 103] in 1989; the Mallat algorithm for multiresolution decomposition underpins an enormous swath of the real-world application of wavelet analysis, and I consider these entwined developments the milestone marking the maturity of first-generation wavelets. In a Ph.D. thesis Feauveau developed biorthogonal wavelets [61] by loosening the requirements on the representation from a basis, i.e. tight Riesz frame, to a Riesz frame with redundant components, and in 1992 along with Cohen and Daubechies published a thorough summary of this new class of transform [32]; these wavelets are now enshrined in the JPEG2000 standard [1, F.4.8].

2.1.4 Second-Generation Wavelets

Once shown relaxing orthnonormality constraints still yielded useful transforms, the flood gates seem to open. In 1993, Lounsbery noted the successive approximation techniques of subdivision algorithms (somewhat analogously to Mallat’s application of the Laplacian pyramid) for arbitrary topologies pioneered in the 1970’s by Catmull [28, 29] and Doo [54, 55], and extended them to wavelets on these general spaces by omitting dilation and translation invariance of the kernel functions [42, 100, 101]; this formulation untethered wavelets from Fourier analysis and actually unknowingly followed in the footsteps of Strömberg’s obscure spline basis [133]. At this same time period, motivated by interpolation and regression work of Breiman, Friedman, Deslauriers, and Dubac [16, 43, 44, 57, 64, 65, 130], Donoho developed adaptive and interpolative wavelets [49, 50, 52] with extensions to the interval [51]; though the polynomial basis functions invoked by Donoho’s work translated and dilated in fashion similar to first-generation wavelets, these techniques were also divorced from Fourier.
Sweldens introduced the *lifting scheme* in 1996 as an algorithm for the construction and implementation of biorthogonal wavelets [136], but proceeded to demonstrate its applicability to the aforementioned non-Fourier techniques and designated the greater class of Fourier-free transformations “second-generation” wavelets [135]. Bearing more than a strong resemblance to the QMF lattice structures pioneered by Vaidyanathan [109, 140, 141], Daubechies and Sweldens showed lifting included cascaded lattices as a special case [41]. Lifting provides a powerful framework for wavelet analysis regardless of topological space or relationship to the Fourier basis, and the technique underpins a wide range of wavelets including transforms with orientation [48], built from IIR filters [151], and on graphs [108].

2.2 Literature Review

2.2.1 Algebraic Wavelets

Abstracting wavelet theory into algebra, specifically group representations, began as early as 1985 with Grossman et al. [70, 71]. Baggett et al abstracted the theorems for wavelet existence [11], and later generalized the notion of multiresolution analysis eliminating the requirement of a scaling function [12]. Through the tail end of the 1990’s and into the early millenium, Klappenecker started developing group and ring theoretic formulations of perfect reconstruction techniques [91] including drawing a distinction between ladder and lattice algorithms [89], deriving perfect reconstruction FIR filters for commutative rings [92], and even shucked time-invariance while maintaining PR [90]; this fascinating work went seemingly unnoticed outside of the optics community.

Brislawn has been doggedly [17–21] chasing a group structure of matrices over Laurent

---

1Sweldens’s dichotomy between the generations has aged so poorly the author considers the distinction largely without utility. The spline wavelets presented as independent of Fourier theory were shown to converge to Shannon wavelets [104, 139] as the degree approached infinity, a transform with poor time localization but excellent frequency properties. Less than a decade later, Adams formulated a technique to counteract the ham-handedness of lifting constrained to the lattice of integers [143]. Lifting-based extension to fields such as quaternions [112] have concrete grounding in the frequency domain, yet bear small resemblance to the filters of old. We include the distinction as a historical curiosity, and subsequently discard it.
polynomials, and discovered remarkable differences between the structural aspects of reversible filters of odd and even lengths. This formulation bears enormous similarity to Borel Group [95, pp.537-540], as well as relation to the Heisenberg group [78,80]. Though currently restricted to orthonormal wavelets through linear lifts, it seems a straightforward extension to nonlinearly conjugate the polyphase lifting matrices to further lift them to integer to integer perfect reconstruction systems as alluded to in [41].

The algebraic structure of wavelets is of vital interest to those concerned with perfect reconstruction, since the inverse transformation is exactly an inverse group action on the space. That much of the aforementioned work seems to be focused upon group characteristic and representation theory, the tools which Püschel used to develop the field of algebraic signal processing [113–115] suggests future discoveries of fundamental properties of (particularly discrete) wavelets will likely arise out of this subfield.

2.2.2 Perfect Reconstruction Wavelets

The vast majority of wavelets are unsuitable to data compression applications requiring perfect bit-fidelity in the reconstructed signal; invariance of measure in $L^2$ may be interpreted geometrically as restriction to a hypersphere and thus transformation can very easily map integers to irrational values. Calderbank et al. addressed briefly addressed transforms for lossless image compression in 1997 [24], offering the S [131], S+P [120], and TS [15, 150] transforms as special cases of lifted integer to integer transforms previously developed [25]; the sequel [26] expands upon the lifting approach and provides expansion factors based upon the precoding technique of Laroia, Tretter, and Farvardin [97] as a less than promising but still viable alternative. Additionally it should be noted that the dream team led by Calderbank in the development of lifting typical wavelets into lossless wavelets [26] was beaten to the punch by Chao et al. [30], which must have been just a bit humiliating.

Apparently unknown to Sweldens at the when he introduced lifting, Bruekers and van der Enden [22] introduced an approach to lossless, reconstructable transformations using
*ladder networks* taking cues from mid-1980’s work of Smith and Barnwell [128] with ties to a tree decomposition developed in the mid-1970’s by Croisier et al. [38]; this development served as a genus independent of lifting for a great deal of perfect reconstruction work from its publication in 1992 through the rest of the decade. One such example is Kofidis et al, who missed the deep connections between lifting and perfect reconstruction networks, but the oversight seems fortunate as they managed to somehow publish the acronyms “PMIS ASS”(*Pairwise Mirror Image-Symmetric Analysis/Synthesis System*) in their 1998 [93] work on lossless transforms via perfect-reconstruction ladder networks. In parallel with development of Daubechies et al. connecting FIR filters, lifting, and perfect reconstruction [26,41], Komatsu and Sezaki introduced their own lossless interpolative transform into the lifting scheme [94] following work in the early 1990’s by Komatsu et al. regarding lossless transform coding [88].

The last decade and a half provided a bountiful harvest of lossless transforms for researchers. The expansion factor techniques eclipsed by lifting found homes on algebraic integer transforms in hardware [58]. Advances in linear algebra on integers clustered around Hao et al. [75,77,84,99,126,144] have provided a suite of robust mathematical tools for lifting arbitrary integer wavelets including multiwavelets [84,127,142]. These advances helped Adams publish a flurry of papers [4,8,143] using a generalized lifting scheme for multiwavelet decompositions reducing the approximation error inevitable when adapting a non-integer transform for perfect reconstruction.

### 2.2.3 Dynamic Range

Chao et al. proposed reliance upon a consistent computational system for signed integer overflow to combat dynamic range expansions [30], and overstating the recklessness of this approach seems impossible. In both the C [85] and C++ [35] standards, signed overflow and underflow are undefined behavior, whereas unsigned integers behave in a modulo fashion [47]. Beyond implementation complications, the security community considers signed integer over-
flow an especially dangerous error \[105,123\]. Srinivasan proposed *modulo transforms* in 2006 and showed that each stage of cascade structure did not increase the overall volume of the codomain of the transform, but noted dynamic range issues remained and required delicate handling possibly by codebooks \[129\]. Senecal et al. preemptively created a codebook transform doing just that in 2004, called the *TLHaar* (Table Lookup Haar) \[124\], and later that year introduced the *PLHaar* (Piecewise Linear Haar) which rotated data clockwise around the unit square \[125\].

### 2.2.4 Infinity-norm Rotation

Yang and Hao \[147\] introduced a rotation transform on \(l^\infty\), the space of absolutely bounded functions (max-norm), and used it to perform both lossy and lossless image compression. Operating on adjacent pairs of data points, they rotate around the *unit square* using an angle they defined on \(\mathbb{R}/\mathbb{Z}^8\), as described in Definition 3.1.1. Conjecturing this was exactly a special case of wavelet lifting \[135\], and the that \(l^\infty\) clearly seemed well-suited to the integer-to-integer transforms \[24, 26\], I managed to prove a far more provocative formula (Equation 3.4) than Yang and Hao’s case-based decomposition in Theorem 3.1.1. Thus far, however, I’ve verified no lifting-type triangular matrix representation, though PLUS \[126\] of Hao et al. seems an option. I’ve verified experimentally the formula in Corollary 3.1.2 which eliminates absolute value and connects the Euclidean angle to Yang and Hao’s angle with Walsh functions. Clearly these rotations may be cascaded and dyadically divided in the same fashion as Givens rotations over \(l^2\) \[132\] Chapter 4.5 \[41\], so the question remains is this a new wavelet formulation or a special case of an older technique.
Chapter 3  Theory

3.1  The Infinity Norm

As stated in A.0.3, for some \( \vec{x} \in \mathbb{F}^n \), \( \| \vec{x} \|_\infty = \max \{ x_1, x_2, \ldots, x_n \} \). Due to having no natural inner product, the typical definition of angle (A.0.11) may not be applied directly to a vector space weighted by \( \| \cdot \|_\infty \). Yang and Hao introduced an angle on the infinity norm in \( \mathbb{R}^n \), generalizing the Euclidean norm to \( p \)-norms.

![Unit Circles for \( L_2, L_4, L_\infty \), rotation by \( \frac{\pi}{6} \) radians on \( L_2 \), rotation by \( 2 - \frac{\sqrt{3}}{2} \) from \( \frac{\sqrt{3}}{2} \) on \( L_\infty \).](image)

Figure 3.1: Unit Circles for \( L_2, L_4, L_\infty \), rotation by \( \frac{\pi}{6} \) radians on \( L_2 \), rotation by \( 2 - \frac{\sqrt{3}}{2} \) from \( \frac{\sqrt{3}}{2} \) on \( L_\infty \).

**Definition 3.1.1** (Yang and Hao \( \theta_p \)). Let \( \vec{x}, \vec{y} \in \mathbb{R}^n \), and \( \vec{u} = \frac{\vec{x}}{\| \vec{x} \|_p}, \vec{v} = \frac{\vec{y}}{\| \vec{y} \|_p} \) designate the corresponding unit vectors, and thus \( \vec{u}, \vec{v} \in S_p \), which in this instance is a compact,
continuous curve on $\mathbb{R}^2$. For arbitrary $p$,

$$\theta_p := \int_{S_p} \vec{u} \cdot d\vec{r}.$$ 

Let $\theta_2^{(u)}, \theta_2^{(v)}$ be the standard Euclidean angles for $\vec{u}$ and $\vec{v}$ respectively, and $f_p : \mathbb{R}/\mathbb{Z}2\pi \rightarrow S_p$ map the typical angle to its corresponding point on the unit sphere. This provides an equivalent form

$$\theta_p = \int_{\theta_2^{(v)}}^{\theta_2^{(u)}} f_p d\theta_2$$

which for $p = 2$ corresponds to the standard Euclidean angle. For $p = \infty$, however, the unit sphere becomes a square,

$$f_\infty (\theta_2) = \begin{cases} 
((-1)^{k+1} \frac{\sin \theta_2}{|\cos \theta_2|})^T & \text{if } \theta_2 \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right] + k\pi \\
\left(\frac{\cos \theta_2}{|\sin \theta_2|}, (-1)^{k+1}\right)^T & \text{if } \theta_2 \in \left[\frac{\pi}{4}, \frac{3\pi}{4}\right] + k\pi 
\end{cases}$$

for $k \in \mathbb{Z}$. This bijects to the unit circle, and since $\frac{\pi}{4} f \mapsto 1$, we may connect the euclidean angle to the infinity angle by the relation

$$\theta_\infty = \frac{2\theta_2}{\pi} \quad (3.3)$$

This definition of angle certainly possesses some attractive qualities, not least of which the elimination of $\pi$ facilitating natural utilization of computationally-friendly approximations.
Theorem 3.1.1. Let $M : \frac{\pi}{\pi^2} \rightarrow [-1, 1]^2$, such that

\[
M(\theta_2) = \frac{(\cos \theta_2 \sin \theta_2)^T}{\left| \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \cos \theta_2 & \sin \theta_2 \\ \sin \theta_2 & \cos \theta_2 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \cos \theta_2 & 1 \\ \sin \theta_2 & \frac{\mathcal{G}^{(\pi)}_1(\theta_2)}{4} \end{pmatrix} \right|}.
\]

(3.4)

Then $M$ is a bijection to the unit square of $l_{\infty}$.

Proof. Without loss of generality, let $\Theta_C = \left( \frac{\pi}{4}, \frac{3\pi}{4} \right) \cup \left( \frac{5\pi}{4}, \frac{7\pi}{4} \right)$. It follows that $\mathcal{G}^{(\pi)}_1(\Theta_C) = \{-1\}$, and thus equation 3.4 becomes

\[
M(\theta_2) = \frac{(\cos \theta_2 \sin \theta_2)^T}{\left| \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \cos \theta_2 & -\sin \theta_2 \\ \sin \theta_2 & \cos \theta_2 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \cos \theta_2 & -1 \\ \sin \theta_2 & \frac{\mathcal{G}^{(\pi)}_1(\theta_2)}{4} \end{pmatrix} \right|} = \frac{(\cos \theta_2 \sin \theta_2)^T}{\left| \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \cos \theta_2 & \cos \theta_2 + \sin \theta_2 \\ \sin \theta_2 & \cos \theta_2 - \sin \theta_2 \end{pmatrix} \right|}.
\]

(3.5)

exactly the value for this region in 3.1. Clearly the orthogonal open sets of angle space correspond to the other case, and since $|\sin \left( k \frac{\pi}{2} + \frac{\pi}{4} \right)| = |\cos \left( k \frac{\pi}{2} + \frac{\pi}{4} \right)| \ \forall k \in \mathbb{Z}$, the corners align regardless the exact definition of the square wave.

Clearly this function is surjective, and since in each case decomposes to a linear rescaling of a vector which by A.0.10 intersects $S^3_{\infty}$ exactly once. \qed

Corollary 3.1.2 (Elimination of Absolute Value). Equation 3.4 is equivalent to

\[
M(\theta_2) = \frac{(\cos \theta_2 \sin \theta_2)^T}{\left| \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \cos \theta_2 & -\sin \theta_2 \\ \sin \theta_2 & \cos \theta_2 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \cos \theta_2 & \frac{\mathcal{G}^{(2\pi)}_1(\theta_2)}{4} \\ \sin \theta_2 & \frac{\mathcal{G}^{(2\pi)}_3(\theta_2)}{4} \end{pmatrix} \right|}.
\]

(3.6)

Proof. For each $\Theta_k = k \frac{\pi}{2} + \left( -\frac{\pi}{4}, \frac{\pi}{4} \right)$, equation 3.6 simplifies to the appropriate value of equation 3.1, and the points $k \frac{\pi}{4}$ are trivially verified by the same reasoning as theorem
Theorem 3.1.3 (Isomorphism Between Euclidean and Infinity Rotation). The function $M$ as defined in equation 3.4 is invertible and thus induces a group structure on $S^2_\infty$ through conjugation of rotations on $S^2_2$. This trivially projects to spheres of arbitrary magnitude.

Proof. In 3.1.1 we showed the bijective properties of $M$, thus eliciting the existence of $M^{-1}$. Let $\text{Ang} : S^2_2 \to \mathbb{R}/\mathbb{Z}2\pi$ map a point on the Euclidean unit circle to its angle, clearly another bijection. Thus the following diagram commutes

$\xymatrix{ S^2_2 \ar[r]_{\text{Ang}} & \mathbb{R}/\mathbb{Z}2\pi \ar[r]^M & S^2_\infty }$

$(M \circ \text{Ang}) R_\theta (M \circ \text{Ang})^{-1}$ corresponds to a rotation of $\frac{2\theta}{\pi}$ on $S^2_\infty$. Since all the operations are magnitude invariant, the operation is isometric on both unit circle and square.

Corollary 3.1.4 (Connecting wavelets on $l^2$ and $l^\infty$). Proof. The lattice decomposition of orthonormal wavelets on $l^2$ take the form of a series of independent two-dimensional Givens rotations [41,132,141], which may be uniquely mapped to infinity rotations by the operation defined in 3.1.3.

Definition 3.1.2 (Integer to Integer Wavelet). A wavelet transform constrained to a discrete lattice, usually designed and implemented in terms of lifting or expansion factors [26], is called a integer to integer wavelet.

Theorem 3.1.5 (Inefficiency of Integer to Integer $l^2$ Wavelets in Bounded Discrete Spaces). Under a hard bit limit, an integer to integer orthogonal or approximately orthogonal $l^2$ wavelet is restricted to at best $\frac{\pi}{4}$ of the available product space of each lifting operation to eliminate any possibility of overflow.

Proof. An integer to integer wavelet lifting step acts as an automorphism on a discrete bounded set, and the constraint of $l^2$-invariance restricts the domain to the values on or within the circle $\|\vec{x}\|_2 \leq N$, where $\vec{x} \in \{-N, -N + 1, \ldots, -1, 0, 1, \ldots, N - 1, N\}^2$. The area
of the direct product of spaces is \((2N + 1)^2\), whereas the area of the circle is \(\pi \left( N + \frac{1}{2} \right)^2 = \frac{\pi}{4} (2N + 1)^2\); hence the quotient is exactly \(\frac{\pi}{4}\) before taking into account the lattice.

**Lemma 3.1.6.** Let \(X = \{-N, -(N - 1), \ldots, -1, 0, 1, \ldots, N - 1, N\}\) and \(X \times X/n = \{\vec{x} \in X^2 : \|\vec{x}\|_\infty = n\}\). \(|X \times X/0| = 1\), and \(|X \times X/n| = 8 |n|\) for \(n \neq 0\).

**Proof.** The 0 case is obvious. For \(n = 1\) the coset has 8 members \{\((\pm 1, 0), (0, \pm 1), (\pm 1, \pm 1)\}\). Assume the coset for \(n\) has \(8n\) members. Each individual entry has an associated entry in the next coset, and the corners each have an additional two, yielding \(8n + 8 = 8(n + 1)\).

**Theorem 3.1.7.** Every integer to integer orthogonal or approximately orthogonal wavelet transform on \(l^2\) embeds into the set of integer to integer transforms on \(l^\infty\), and this embedding eliminates all possibility of overflow for \(\vec{x} \in \{-N, -N + 1, \ldots, -1, 0, 1, \ldots, N - 1, N\}^2\).

**Proof.** We know from 3.1.5 at a radius of \(n\) the isometric cosets of \(l^2\) lie on a circle with circumference \(\pi(2n + 1)\), whereas 3.1.6 showed the \(l^\infty\) cosets are \(8n\) in size. For \(n > 0\),

\[
r(n) := \frac{\pi(2n + 1)}{8n} = \frac{2\pi n + \pi}{8n} = \frac{\pi}{4} + \frac{\pi}{8n},
\]

which for \(r(n + 1) < 1\) and \(r(n) \to \frac{\pi}{4}\) as \(n \to \infty\). At \(n = 1\) the value \(\frac{5}{4} > \frac{3\pi}{8} > 1\), thus in the context of an integer lattice the isometric cosets of both spaces are identically the eight values adjacent to the center point. Applying 3.1.4 with a consistent adherence to the discrete lattice, secure in the knowledge that the square has as much or more values than the circle, we can be certain of the existence of (non-unique) mapping from the circle to the square for every \(n \leq N\). Since square rotation cannot increase the absolute magnitude of any dimension beyond the greatest, overflow is an impossibility.

### 3.2 Additional Techniques

The remainder of this chapter is devoted to techniques not necessarily related to the wavelet transform, but as part of the machinery for the data compression experiments in chapter 4.
Definition 3.2.1 (Piecewise-Linear Haar). The Piecewise-Linear Haar [125] is a special case [147] of square rotation, one in which \( \theta_2 = -1 \) [125], and is denoted \( \text{PLHaar} : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \).

Definition 3.2.2 (S-Transform). The S-Transform [131] is a discretized version of the Haar-transform designed for perfect reconstruction, \( \text{ST} : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \), such that

\[
\begin{pmatrix}
y_0 \\
y_1
\end{pmatrix}
= \text{ST}
\begin{pmatrix}
x_0 \\
x_1
\end{pmatrix}
= \left( \left\lfloor \frac{x_0 + x_1}{2} \right\rfloor , x_1 - x_0 \right)
\tag{3.7}
\]

\[
\begin{pmatrix}
x_0 \\
x_1
\end{pmatrix}
= \text{ST}^{-1}
\begin{pmatrix}
y_0 \\
y_1
\end{pmatrix}
= \left( y_0 - \left\lfloor \frac{1}{2} y_1 \right\rfloor , y_0 + \left\lfloor \frac{y_1 + 1}{2} \right\rfloor \right)
\tag{3.8}
\]

Definition 3.2.3 (Exponential-Golomb Code). Originally called the “Multimode Golomb code” when introduced in 1974 by Bahl and Kobayashi [13] as an alternative to Golomb codes [67] for non-exponential distributions, by 1978 this variable length code [121] was known as the Exponential-Golomb code [138] and finds contemporary use in the H.264/MPEG4 standard [110, 118]. For \( n, k \in \mathbb{N}_0 := \{0, 1, \ldots\} \), where \( n \) is the value to be encoded and \( k \) is a parameter, and \( P_k : \mathbb{N}_0 \rightarrow \mathbb{N}_0 \) determines the prefix of a value for parameter \( k \). Then

\[
P_k(n) = m : \sum_{l=0}^{m-1} 2^{l+k} \leq n < \sum_{l=0}^{m} 2^{l+k}
\]

\[
r_k(n) := n - \sum_{l=0}^{P_k(n)-1} 2^{l+k},
\]

and the code is \( P_k(n) \) prefix bits set 1, followed by a 0, and completed with the value of \( r_k(n) \) as a \( (P_k(n) + k) \)-bit number.
Chapter 4  Experiments

We experiment here with lossless compression of power-quality monitoring signals, both Voltage and Amperage, as a means of demonstrating the utility of infinity norm rotations. The algorithm is quite trivial, and only during the literature review did the author discover the transform tested was introduced by Senecal [125] as the PLHaar transform. The choice of the S-transform as control stemmed from its similarity to the PLHaar—though theorem 3.1.4 demonstrated equivalence to orthonormal $l^2$-wavelets, the method for effectively construction of $l^\infty$-wavelet corresponding to the $l^2$ version remains unclear at this time. These two simple transforms provide an “apples-to-apples” comparison between lifting perfect reconstruction filters and infinity-norm perfect-reconstruction filters.

4.1 Data

The dataset entails about two hundred three phase waveforms each consisting of 16 cycles captured at 512 samples per cycle, i.e. 30,720 Hz. Though ostensibly 16-bit signed samples, the actual sampling was done at 14 bits, but has been rescaled by an unknown algorithm to occupy most of the 16-bit dynamic range. Investigation of the data provided no clear and consistent bijection between the 16-bit and 14-bit scales. Although taken from three-phase systems, the experiments consider each waveform, whether line-to-line Voltage or line current, in isolation.
<table>
<thead>
<tr>
<th>stored bitrate</th>
<th>actual bitrate (estimated)</th>
<th>sample frequency (Hz)</th>
<th>samples-per-cycle</th>
<th>total samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>14</td>
<td>30,720</td>
<td>512</td>
<td>8192</td>
</tr>
</tbody>
</table>

Table 4.1: Dataset Summary

### 4.2 Validation

The “stretching” applied to the data creates a sparse “binning” of the dynamic range, resulting in a sample-space in which no two values are directly adjacent, or for signal $X$ conforming to 4.1

$$
\delta(X) := \sup_{a,b \in X, a \neq b} |a - b| > 1 \quad (4.1)
$$

for every signal considered, and this $\delta(X)$ parameter will be presented for all the results as well as the sample entropy

$$
H(X) := - \sum_{a \in X} p(a | X) \log_2(p(a | X)). \quad (4.2)
$$

Though the techniques applied are for lossless compression, the rescaling of the data induced the S-Transform to exceed the dynamic range in some instances, catastrophically breaking the coding algorithm in the process. Unless otherwise mentioned, the reconstruction has perfect bit fidelity.

### 4.3 Transformed

A signal is completely, recursively decomposed into its dyadic decomposition for each transform, i.e. 13 bands. The 512 values of the lowest (9) frequency subbands were encoded natively, i.e. as 16-bit values. Each subsequent subband was prefixed with 4 bits indicating the $k \in \{1, 2, \ldots, 15\}$ parameter of the Exponential-Golomb code (definition 3.2.3) used to encode the band.
Table 4.2: Compression Format
\[
\begin{array}{|c|c|c|c|c|}
\hline
\{1, \ldots, 512\} & \{513, \ldots, 1024\} & \{1025, \ldots, 2048\} & \{2049, \ldots, 4096\} & \{4097, 8192\} \\
\text{16-bit native} & \text{ExpGol}(k_{10}) & \text{ExpGol}(k_{11}) & \text{ExpGol}(k_{12}) & \text{ExpGol}(k_{13}) \\
\hline
\end{array}
\]

4.4 Results

As mentioned, a large number of S-Transform experiments failed due to dynamic range issues, inadvertently demonstrating the value of the PLHaar.

4.4.1 Signal 1 (Voltage)

In this instance, the PLHaar and ST transforms performed comparably and beat entropy.

<table>
<thead>
<tr>
<th>(\delta(X_1))</th>
<th>(H(X_1))</th>
<th>14-bit PLHaar (%)</th>
<th>PLHaar/(H(X_1))</th>
<th>14-bit S %</th>
<th>S/(H(X_1))</th>
<th>PLHaar/ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>11.76</td>
<td>83.74</td>
<td>0.9971</td>
<td>83.75</td>
<td>0.9973</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

Table 4.3: Signal 1 Results Table
Figure 4.2: Signal 1 Decomposed: green encoded natively, black with Exponential-Golomb, PLHaar left, S-Transform right

Figure 4.3: Signal 1 Native (in red) and several bands of PLHaar (left, black) and S-Transform (right, black)

<table>
<thead>
<tr>
<th>$\delta(X_1)$</th>
<th>$H(X_1)$</th>
<th>14-bit PLHaar (%)</th>
<th>$PLHaar/H(X_1)$</th>
<th>14-bit S %</th>
<th>$S/H(X_1)$</th>
<th>PLHaar/ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>48</td>
<td>10.02</td>
<td>84.01</td>
<td>1.173</td>
<td>84.02</td>
<td>1.175</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

Table 4.4: Signal 1 Results Table
Figure 4.4: Signal 1 Decomposed Green encoded natively, black with Exponential-Golomb, PLHaar left, S-Transform right

<table>
<thead>
<tr>
<th>$\delta(X_1)$</th>
<th>$H(X_1)$</th>
<th>14-bit PLHaar (%)</th>
<th>PLHaar/$H(X_1)$</th>
<th>14-bit S %</th>
<th>S/$H(X_1)$</th>
<th>PLHaar/ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>81</td>
<td>9.260</td>
<td>84.39</td>
<td>1.276</td>
<td>84.56</td>
<td>1.278</td>
<td>0.998</td>
</tr>
</tbody>
</table>

Table 4.5: Signal 3 Results Table

### 4.4.2 Signal 2 (Amps)

This signal had a whopping $\delta$ of 48, and thus neither transform beat entropy, although the percentage from native is still in the 80’s. This is a common issue with current waveforms, because due to their propensity for rapid growth in magnitude, the A/D converters must be configured to have a lot of range to expand.

### 4.4.3 Signal 3 (Amps)

### 4.4.4 Signal 4 (Amps) S-Transform Failure

Notice the S-Transform (left image, figure 4.8) exceeds the dynamic range of 16-bit data.
Figure 4.5: Signal 3 Native (in red) and several bands of PLHaar (left, black) and S-Transform (right, black)

Figure 4.6: Signal 3 Decomposed: green encoded natively, black with Exponential-Golomb, PLHaar left, S-Transform right

<table>
<thead>
<tr>
<th>$\delta(X_1)$</th>
<th>$H(X_1)$</th>
<th>14-bit PLHaar (%)</th>
<th>PLHaar/$H(X_1)$</th>
<th>14-bit S %</th>
<th>S/$H(X_1)$</th>
<th>PLHaar/ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>10.22</td>
<td>85.32</td>
<td>1.168</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Table 4.6: Signal 4 Results Table, * is invalid because S Transform Failed
Figure 4.7: Signal 4 Native (in blue) and several bands of PLHaar (left, black) and S-Transform (right, black)

Figure 4.8: Signal 4 Decomposed: blue encoded natively, black with Exponential-Golomb, PLHaar left, S-Transform right
Chapter 5 Conclusions and Summary

5.1 Further Musings

In this survey [53], the Donoho et al. flippantly dismiss $L^\infty$ and $L^1$ spaces as “not separable” and therefore uninteresting, but they also bring up the bounded mean oscillation ($BMO$) and Hardy ($H_p$) function spaces, in which $BMO$ and $H_1$ are dual analogous to $L^\infty$ and $L^1$ [69, Section 7.2]. For the discrete case this distinction seems immaterial, unless one wished to design their transform in the complete space which is beyond our scope. The unit ball of $BMO$-space is again a cube, but the meaning is quite different [63], though decomposition into a finite series of $L^\infty$ functions is both necessary and sufficient for membership [62]. A maximum operator may be defined bounded for all $BMO$ functions on the extended reals except those identically infinity [3], and according to the preprint [2] a generalized version of $L^\infty$ called the weak $L^\infty$ or $L^\infty_w$ is a proper superset of $BMO$ functions over a doubling measure, which simply demands a uniform upper bound on the ratio between the measure of all equicentered balls radius $r$ and $\frac{r}{2}$. What interests me most is Donoho et al. [53] also indicated that the Haar basis is optimal basis for $BMO$, and Haar is simply an orthonormalized Walsh frame.

Hypercubes have a beautiful geometric connection to direct products of Cantor sets [56], and very recently [39] a large class of Cantor sets permitting doubling measures was established [39]. In the process of developing sufficient and necessary conditions for this property in uniform Cantor sets [146], Wei noted that the sequences characterizing the
cuts generated an infinite series which if bounded, i.e. a member of \( l^1 \), indicated the set as measurable. Hytönen and Kairema demonstrated a finite dyadic decomposition of BMO-space using hypercubes \[82\]. In fact, the locally compact dyadic Cantor group has Walsh functions as its Pontrygin dual \[96\]. Constructions of \( p \)-adic wavelets of integer dilation factors other than 2 were presented by Farkov using Walsh functions for masking purposes \[60\], and it should be noted that 2-adic numbers are the Cantor set under a much different norm than Lang’s \[96\]. Perhaps even more provocatively, dyadic hypercubes, i.e. products of dyadic Cantor sets, are regularly invoked for techniques to approximate multi resolution decompositions \[14, 33, 34, 45, 46, 83\].

Minimax optimization is also regularly viewed in the context of hypercubes \[122, 148, 149\].

5.2 Conjectures and Open Questions

Conjecture 1. It is known that the isometries of hypercube under Hamming distance is a semidirect product of the dimensions with the symmetric permutations its subspaces \[111\], and for dimension \( n \) isomorphic to \( \mathbb{Z}_2 \times S_n \) \[68\]. If there is homomorphism between the PLUS factorization of Hao and She \[77\], can it be used to find a (thus-far unknown) sufficient condition for the existence of the factorization for a given linear operator? There certainly exists a deep connection between wavelets and semidirect products \[10, 20\].

5.3 Conclusion

The PLHaar transform as a special case of infinity rotation works to losslessly compress power quality monitoring signals. Given that this geometry may be applied to more complicated transforms and actually corresponds to traditional wavelets, there should be ample room for research and improvement in this field for this family of techniques.

\(^1\) Also interesting but beyond scope is the measure defined by Lang in \[96\] corresponds exactly with the space of arbitrary precision floating point representations, which suggests a possibility of utilizing these wavelets to simultaneously transform and arithmetically encode a signal.
Reference List


[85] Larry Jones. Wg14 n1539 committee draft iso/iec 9899: 201x, 2010.


Appendices
Mathematical Preliminaries

Definition A.0.1 (Matrix and Vector Notation). Let

- $\mathbb{F}^n$ denote an arbitrary field of dimension $n$
- $M_{m,n}(\mathbb{F})$ designate the set of $m \times n$ matrices over the field $\mathbb{F}$, with $M_n(\mathbb{F}) := M_{n,n}(\mathbb{F})$
- $\vec{x} \in \mathbb{F}^n$ represent a vector, column oriented unless otherwise specified, i.e. $\vec{x} \in M_{n,1}(\mathbb{F})$
- $x_{l,k}$ denote the entry in the (1-indexed) $l$th row and $k$th column of a matrix $x$, with $x_l$ the $l$th entry of a vector $\vec{x}$.
- $A^T$ represent the transpose of $A \in M_{m,n}(\mathbb{F})$, i.e. $B = A^T \in M_{n,m}(\mathbb{F}) : a_{l,k} = b_{k,l}$.
- $C^*$ represent the conjugate transpose of $C \in M_{m,n}(\mathbb{C})$, i.e. $D = C^* \in M_{n,m}(\mathbb{C}) : d_{l,k} = \overline{c_{k,l}}$
- $A \in M_n(\mathbb{F})$ be nonsingular unless specified otherwise
- $GL_n(\mathbb{F})$ denote the general linear group formed by the subset of nonsingular matrices of $M_n(\mathbb{F})$ under typical matrix multiplication
- $SL_n(\mathbb{F}) \subset GL_n(\mathbb{F})$ be the set of matrices $A \in SL_n(\mathbb{F})$ such that $\det A = 1$.
- $I_n \in GL_n(\mathbb{F})$ be the identity matrix, and $\vec{1}_n \in \mathbb{F}^n$ be the vector of all ones.
- $0_n \in M_n(\mathbb{F})$ represent the zero matrix, which is singular, and consequentially $0_n \notin GL_n(\mathbb{F})$. Also $\vec{0}_n \in \mathbb{F}^n$ represents the zero vector.
Definition A.0.2 (Vector Norms). Let $V$ be a vector space over a field $F$, and a function denoted $\| \cdot \| : V \to [0, \infty)$ which $\forall \vec{x}, \vec{y} \in V$ satisfies

1. Nonnegativity: $\| \vec{x} \| \geq 0$
2. Positivity: $0 = \| \vec{x} \| \iff \vec{x} = \vec{0}$
3. Absolute Homogeneity: $\| c\vec{x} \| = |c| \| \vec{x} \| \ \forall c \in F$
4. the Triangle Inequality: $\| \vec{x} + \vec{y} \| \leq \| \vec{x} \| + \| \vec{y} \|$

be called a vector norm.

Definition A.0.3 (Inner Product). Let $V$ be a vector space over $F$, and $\vec{x}, \vec{y}, \vec{z} \in V$. A function designated $\langle \cdot, \cdot \rangle : V^2 \to F$ guaranteeing

1. Nonnegativity: $\langle \vec{x}, \vec{x} \rangle \geq 0$
2. Positivity: $\langle \vec{x}, \vec{x} \rangle = 0 \iff x = 0$
3. Additivity: $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$
4. Homogeneity: $\langle c\vec{x}, \vec{y} \rangle = c \langle \vec{x}, \vec{y} \rangle \ \forall c \in F$
5. the Hermitian Property: $\langle \vec{x}, \vec{y} \rangle = \overline{\langle \vec{y}, \vec{x} \rangle}$

is called an inner product.

Lemma A.0.1 (An Inner Product Induces a Vector Norm). Given any inner product $\langle \cdot, \cdot \rangle$, a norm may be defined as $\| \vec{x} \| : = \sqrt{\langle \vec{x}, \vec{x} \rangle}$ [79] p.262 5.1.7.

Lemma A.0.2 ($p$-norms). Functions $\| \cdot \|_p : \mathbb{C}^n \to [0, \infty)$, defined as $\| \vec{x} \|_p = (\sum_{k=1}^{n} |x_k|^p)^{\frac{1}{p}}$ satisfy the norm axioms for $p \in [1, \infty)$ [79].

Lemma A.0.3 (Infinity Norm). For $x \in \mathbb{C}^n$, $\| \vec{x} \|_{\infty} : = \lim_{p \to \infty} \| \vec{x} \|_p = \max \{|x_1|, |x_2|, \ldots, |x_n|\}$ [79].
Lemma A.0.4 (Inner products and $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$). The Euclidean norm may be defined in terms of the typical inner product, but $\|\cdot\|_1$ and $\|\cdot\|_\infty$ may not, and thus possess no “natural” inner product \[79\]

Definition A.0.4 (Isometry). For a vector space $V$ over a field with $\|\cdot\|$, an isometry a function $f \in \text{End} V$ such that $\|\vec{x}\| = \|f(\vec{x})\|$ $\forall \vec{x} \in V$.

Definition A.0.5 (Group). A group is a pair $(G, \cdot)$, where the binary operator $\cdot : G^2 \to G$ surjects the element pairs of set $G$ onto itself such that $\exists! e \in G : \forall a \in G : e \cdot a = a \cdot e = a$ and $\forall b \in G \exists! b^{-1} : b \cdot b^{-1} = b^{-1} \cdot b = e$.

Definition A.0.6 (Subgroup). A group $H$, designated a subgroup, $H$ of group $(G, \cdot)$ is a set $H \subseteq G : \forall a, b \in Ha \cdot b^{-1} \in H$. $N \triangleleft$.

Definition A.0.7 (Normal Subgroup). A subgroup $N$ of group $G$, i.e. $N \triangleleft G$, is normal if $\forall g \in G : gN = Ng$, or, in English, the cosets $G/N$ are without exception commutative.

Definition A.0.8 (Semidirect Product Group). Let $G$ be a group, $H < G$, $N \triangleleft G$, $N \cup H < G$ and $N \cap H = \{e\}$. The semidirect product $S := N \rtimes H$ [86, p.198].

Remark A.0.5 (Isometry Group). The set of all bijective isometries over a vector field $V$ form a group.

Proof. Let $f, g$ be bijective isometries over $V$ with norm $\|\cdot\|$, and note that $I$ is clearly an isometry. By hypothesis, $\|\vec{x}\| = \|f(\vec{x})\| = \|g \circ g^{-1} \circ f(\vec{x})\| = \|g^{-1} \circ f(\vec{x})\|$.

Definition A.0.9 (Linear Operator). Let $V$ be a vector-space, $x, y \in V$ and $\alpha \in R$. A map $T \in \text{Hom} R$, such that $T$ is homogenous, i.e. $T \alpha x = \alpha Tx$, is called a linear operator.

Definition A.0.10 (Linear Isometry). An linear operator $T \in M_n(F)$ is isometric or an isometry for a vector norm if $\|\cdot\|$ on $F$, $\|T\vec{x}\| = \|\vec{x}\|$ $\forall \vec{x} \in F^n$.

Remark A.0.6 (Every linear operator $T$ has a unique matrix representation [145, p.283 8.15]).
Lemma A.0.7 (Linear Isometry Subgroup). For norm \( \|\cdot\| \) over a vector space \( \mathbb{R}^n \) or \( \mathbb{C}^n \), the set of (nonsingular) matrices isometric on all vectors under standard matrix multiplication forms a group.

Proof. Since \( \|I_n\vec{x}\| = \|\vec{x}\| \) there at bare minimum is a trivial group associated. The null space of any isometric matrix must be limited to \( \vec{0} \) and thus is nonsingular. As a consequence, isometries \( A \) and \( B \) act as changes of basis with full range of the underlying vector space guaranteeing that any \( \vec{x} \) therein

\[
\|\vec{x}\| = \|I_n^2\vec{x}\| = \|B^{-1}BA^{-1}\vec{x}\| = \|B^{-1}A\vec{x}\|.
\]

\[\square\]

Corollary A.0.8 (Isometric Rotation Matrices of Euclidean \( \mathbb{R}^2 \)). The set of matrices, \( R_\theta \subset M_2(\mathbb{R}) \) such that

\[
R_\theta := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \tag{A.1}
\]

is a unitary, isometric under \( \|\cdot\|_2 \), a group under matrix multiplication, \( \{R_\theta : \theta \in \mathbb{R}/\mathbb{Z}2\pi\} \subset SL_2(\mathbb{R}) \), and \( R_\theta^{-1} = R_{-\theta} \).

Proof. \( \det R_\theta = \cos^2 \theta + \sin^2 \theta = 1 \) so clearly \( R_\theta \) is both unitary and nonsingular (and thus in \( SL_2(\mathbb{R}) \)). Also

\[
R_\theta R_\alpha = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \tag{A.2}
\]

\[
= \begin{pmatrix} \cos \theta \cos \alpha - \sin \theta \sin \alpha & -\cos \theta \sin \alpha - \sin \theta \cos \alpha \\ \cos \theta \sin \alpha + \sin \theta \cos \alpha & \cos \theta \cos \alpha - \sin \theta \sin \alpha \end{pmatrix}
\]

\[
= \begin{pmatrix} \cos (\theta + \alpha) & -\sin (\theta + \alpha) \\ \sin (\theta + \alpha) & \cos (\theta + \alpha) \end{pmatrix} = R_{\theta+\alpha}, \tag{A.3}
\]

47
and thus \( R_\theta R_{-\theta} = R_\theta = I_2 \). Let \((x, y)^T \in \mathbb{R}^2\).

\[
\left\| R_\theta \begin{pmatrix} x \\ y \end{pmatrix} \right\|_2 = \sqrt{|x \cos \theta - y \sin \theta|^2 + |x \sin \theta + y \cos \theta|^2}
\]
\[
= \sqrt{\frac{x^2 \cos^2 \theta - 2 |xy \cos \theta \sin \theta| + y^2 \sin^2 \theta + x^2 \sin^2 \theta}{x^2 + y^2}}
\]

\[
= \sqrt{x^2 + y^2}
\]

Definition A.0.11 (Vector Angle). The angle \( \theta \) between two vectors \( \vec{x}, \vec{y} \in \mathbb{C}^n \) endowed with an inner product \( \langle \cdot, \cdot \rangle \) is

\[
\theta := \arccos \left( \frac{|\langle \vec{x}, \vec{y} \rangle|}{\langle \vec{x}, \vec{x} \rangle \langle \vec{y}, \vec{y} \rangle} \right) \tag{A.4}
\]

Remark A.0.9 (\( \text{SL}_n(\mathbb{F}) < \text{GL}_n(\mathbb{F}) \)). Proof. Let \( A, B^{-1} \in \text{SL}_n(\mathbb{F}) \), and \( \det AB^{-1} = \det A \det B^{-1} = 1 \cdot 1 \).

Definition A.0.12 (Square Waves). Let \( \mathcal{S}_\phi^{(T)} : \mathbb{R}/\mathbb{Z}T \to \{\pm 1\} \) for some \( T \in \mathbb{R} \) be a function such that

\[
\mathcal{S}_\phi^{(T)}(t) = \begin{cases} 
1 & \text{if } -\phi \leq t < \frac{T}{2} - \phi \\
-1 & \text{if } \frac{T}{2} - \phi \leq t < T - \phi
\end{cases}
\tag{A.5}
\]

let \( \mathcal{C}_\phi^{(T)}(t) := \mathcal{S}_\phi^{(T)}(t) \), and call these functions Square Sine and Square Cosine respectively.

Definition A.0.13 (Unit Ball and Sphere). For vector space \( V \) with norm \( \| \cdot \| \), the subset \( B := \{ \vec{x} \in V : \| \vec{x} \| \leq 1 \} \) is called the unit ball and \( S := \{ \vec{x} \in V : \| \vec{x} \| = 1 \} \) is called the unit sphere.

Lemma A.0.10. Let \((\mathbb{C}^n, \| \cdot \|)\) be a normed vector space of dimension \( n \in \mathbb{N} \) with unit sphere \( S = \{ \vec{x} \in \mathbb{C}^n : \| \vec{x} \| = 1 \} \). The for some \( \vec{v} \in \mathbb{C}^n \setminus \{0\} \), the curve \( \alpha : (0, \infty) \to \mathbb{C}^n \) defined by \( \alpha(t) = t\vec{v} \) intersects \( S \) at exactly the point \( \alpha \left( \frac{1}{\|\vec{v}\|} \right) = \frac{\vec{v}}{\|\vec{v}\|} \).
Proof. \( \| \alpha \left( \frac{1}{\| \vec{v} \|} \right) \| = \| \frac{\vec{v}}{\| \vec{v} \|} \| = \frac{1}{\| \vec{v} \|} \| \vec{v} \| = 1 \) and thus \( \alpha \left( \frac{\vec{v}}{\| \vec{v} \|} \right) \in S \). Additionally \( \| \alpha(t) \| = \| t \vec{v} \| = |t| \| \vec{v} \| = t \| \vec{v} \| \in C^1 \), and by virtue of strict monotonicity in \( t \) has exactly one solution for \( \| \alpha(t) \| = 1 \). \qed