The Uncertainty Principle,
A Different Perspective

M. An, J. J. Benedetto, G. Björck, J. Byrnes,
D. Cochran, I. Gertner, W. Moran, D. J. Newman,
G. Ostheimer, M. Ramalho, B. Saffari,
H. S. Shapiro, R. Tolimieri
**Classical Uncertainty Principle**

If $f$ in $L^2(\mathbb{R})$, then both $f$ and $\hat{f}$ can't go to 0 “too quickly” at $\cdot$, i.e., if $f$ is “concentrated” then $\hat{f}$ is “spread out,” and vice versa.

Question: Given a large subset $\mathcal{F}$ of $L^2(\mathbb{R})$, what is the best uniform rate at which all $f$ in $\mathcal{F}$ and all of their Fourier transforms $\hat{f}$ can go to 0 at $\pm$?
Theorem (Kolmogorov)

Suppose \( \sup_\mathcal{F} \| f \|_{L^2} \leq M \). For \( t, \| \| > 0 \) let

\[
A_t = \sup_{f \in \mathcal{F}} \int_{|x| \geq t} |f(x)|^2 \, dx
\]

\[
B_\| = \sup_{f \in \mathcal{F}} \int_{\| \| \geq \|} |\hat{f}(\|)|^2 \, d\|.
\]

If \( \lim_{t \to \infty} A_t = \lim_{\| \to \infty} B_\| = 0 \) then \( \mathcal{F} \) is precompact.

In particular, \( \mathcal{F} \) cannot contain an infinite orthonormal set.
Corollary (H. S. Shapiro)
The following is impossible:
\( F \subseteq L^2(\mathbb{R}) \) an infinite orthonormal set, \( p > \frac{1}{2} \), and \( \square f \subseteq F \)

\[ (*) \quad |f(x)| < \frac{C_1}{(1 + |x|)^p}, \quad |\hat{f}(\square)| < \frac{C_2}{(1 + |\square|)^p} \]

Theorem
\( \square \) an orthonormal basis (ONB) for \( L^2(I) \), where \( I \subseteq \mathbb{R} \) is any finite interval (which, for convenience, we take to be \([0,1]\)), such that

1) \( \square n, x, \square_n(x) \) takes on only \( \pm 1 \)

2) \( \left| \int_0^1 \square_n(x) e^{2\pi i \square x} dx \right| \subseteq \frac{C}{\sqrt{1 + |\square|}} \)

Corollary (a Global Uncertainty Principle)
\( \square \) an ONB \( S \) for \( L^2(\mathbb{R}) \) such that \( (*) \) with \( p = \frac{1}{2} \) is satisfied for all \( f \subseteq S \)
Fundamental Lemma

Let \( \{R_n(z)\} \) be any sequence of upper flat polynomials on the unit circle \(|z| = 1\) with unimodular coefficients. That is,

\[
R_n(z) = \sum_{k=0}^{n-1} c_k(n)z^k \quad \text{with all } |c_k(n)| = 1
\]

and \( \|R_n\| \leq C \sqrt{n} = C\|R_n\|_{L^2} \).

Define, for each \( n \), the piecewise constant function \( y_n(x) \) on \([0, 1]\) to be the coefficients of \( R_n(z) \). That is, \( y_n(x) = c_k(n) \) for \( \frac{k}{n} \leq x < \frac{k+1}{n} \).

Then

\[
\left| \int_0^1 y_n(x)e^{-2i\theta x} \, dx \right| \leq \frac{C}{\sqrt{1 + |\theta|}}.
\]

**Problem:** Choose a collection of such sequences of upper flat polynomials such that all of the resulting \( y_n \)'s are pairwise orthogonal and the set of all such \( y_n \)'s spans \( L^2[0,1] \).
Shapiro Polynomials

\[ P_0(z) = Q_0(z) = 1 \]
\[ P_{n+1}(z) = P_n(z) + z^{2^n} Q_n(z) \]
\[ Q_{n+1}(z) = P_n(z) \quad z^{2^n} Q_n(z) \]

\( P_n \) and \( Q_n \) are polynomials of degree \( 2^n - 1 \) with coefficients \( \pm 1 \).

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_0 )</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>( Q_0 )</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>( P_1 )</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( Q_1 )</td>
<td>1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>( P_2 )</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( Q_2 )</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>
Properties of Shapiro Polynomials

For \(|z|=1|, \quad \left|P_{n+1}(z)\right|^2 + \left|Q_{n+1}(z)\right|^2
= 2 \left(\left|P_n(z)\right|^2 + \left|Q_n(z)\right|^2\right)
= 2^{n+2}
\left|P_n(z)\right| \sqrt{2} \sqrt{2^n} \quad \text{or} \quad \|P_n\| L\sqrt{2}\|P_n\|_L^2

This choice of \(\pm 1's\) gives an excellent bound \(\sqrt{2}\) for the “peak factor” (peak-to-average ratio), thereby spreading the “energy” of these polynomials almost equally around the unit circle.
Three ways of thinking of these sequences:
(a) sequences of ±1’s of length $2^n$
(b) coefficients of polynomials
(c) values of piecewise constant functions on [0,1]

Note: $P_n$ and $Q_n$ are orthogonal in the sense of (c)
Shapiro Sequences are Incomplete

We want:

A collection of functions of type (c) so that the characteristic function of any interval
\[ \left[ \frac{k}{2^n}, \frac{k + 1}{2^n} \right], \quad 0 \leq k < 2^n, \]
can be expressed as a finite linear combination of them.

We have: 2 such functions for each \( n \).

We need: \( 2^n \) such functions for each \( n \).
PONS Sequences

---

...
Mathematical Properties

\[ P(z) = \prod_{k=0}^{2^n-1} \frac{\sqrt{2} \sqrt{2^n}}{2^n-k} \]  
\[ z = e^{it}, \{k\} \text{ a PONS sequence.} \]

\[ |P(z)| \prod \sqrt{2} \sqrt{2^n} \]

\[ P(z)P(\frac{1}{z}) + P(\frac{1}{z})P(\frac{1}{z}) \equiv 2^{n+1} \]

\[ \int_{0}^{1} f(t)e^{2\pi i t} dt \left| \prod \frac{C}{\sqrt[2^n]{1 + \sqrt{t}}} \right|, \prod \mathbb{R}, \]

where \( f(t) = k, \frac{k}{2^n} \leq t < \frac{k+1}{2^n}, 0 \leq k < 2^n. \)
PONS—The Formal Definition

Let \( P_{1,1}(z) = 1 + z \), \( P_{1,2}(z) = 1 - z \).

Given \( P_{n,m}(z) \), \( m = 1, 2, \ldots, 2^n \),

for \( j = 0, 1, 2, \ldots, 2^{n-1} - 1 \) and \( m = 4j + 1 \) define \( P_{n+1,m}(z) \), \( m = 1, 2, \ldots, 2^{n+1} \) by

\[
\begin{align*}
P_{n+1,m} &= P_{n,2j+1} + z^{2^n} P_{n,2j+2} \\
P_{n+1,m+1} &= P_{n,2j+1} - z^{2^n} P_{n,2j+2} \\
P_{n+1,m+2} &= P_{n,2j+2} + z^{2^n} P_{n,2j+1} \\
P_{n+1,m+3} &= -P_{n,2j+2} + z^{2^n} P_{n,2j+1}
\end{align*}
\]
For each $n \geq 1$, each $j = 0, 1, 2, \ldots, 2^{n-1} - 1$, and each $m = 4j + 1$,

$$
\left| P_{n+1,m} \right|^2 + \left| P_{n+1,m+1} \right|^2 \equiv \left| P_{n+1,m+2} \right|^2 + \left| P_{n+1,m+3} \right|^2 \equiv 2^{n+2}
$$

so that all such $P_{n,m}(z)$ have crest factor $\sqrt{2}$.

Now define $Q_{0,0}(z) = 1$, $Q_{0,1}(z) = 1 - z$, and

$$
Q_{n,m}(z) = (1 - z) P_{n,m}(z^2).
$$

All $Q_{n,m}(z)$ have crest factor $\sqrt{2}$, and the piecewise constant functions generated by their coefficients form the required orthogonal basis.
PONS = Walsh+

PONS satisfies all useful Walsh properties \textit{plus}

- PONS polynomials have uniformly low crest factor
- they \textit{are} quadrature mirror filters
- PONS Fourier transforms have the optimal uniform decay rate \( p = \frac{1}{2} \)
- PONS functions satisfy optimal \textit{Global Uncertainty Principle} bounds
- PONS “spreads energy”
32-Coefficient Magnitude Spectra

Magnitude Transfer for PONS/Walsh-32: Poly 3 (dB diff = 9.031)

PONS (blue curve)
Walsh (red curve)
Average Correlation Magnitudes

Average Correlation Magnitudes for Order 64 In-Place PONS Polynomials

Average Correlation Magnitudes for Order 64 Walsh Polynomials
Conjecture

Let $P(z)$ be any PONS polynomial of length $L = 2^n$ and let $m = \lfloor \sqrt{L} \rfloor$. For $0 \leq k < m$ let

$$I_{P,k,m} = \sum_{2^{k/m}}^{2^{(k+1)/m}} \left| P(e^{it}) \right|^2 dt$$

and let $D_{P,m}$ be the “dynamic range” of $I$. That is, $D_{P,m} = \frac{\max_{0 \leq k < m}(I_{P,k,m})}{\min_{0 \leq k < m}(I_{P,k,m})}$.

Then $\lim_{n \to \infty} D_{P,m} = 1$. 

Energy Spreading—One Interpretation
Energy Spreading—Beurling ME Norm

Let $\mu$ be a bounded complex measure on $\mathbb{R}$, $F_\mu$ its Fourier-Stieltjes transform, $F_\mu = \int e^{it\mu} d\mu(t)$.

$x(t)$ a complex-valued function on $\mathbb{R}$, $\|x\|_w$ its Wiener norm, $\|x\|_w = V(\mu)$ if $\mu$ satisfying $F_\mu(t) = x(t)$, otherwise $\|x\|_w = \cdot$.

$\square(t)$ a complex-valued function defined on $E \subseteq \mathbb{R}$, $\|\square(t)\|_{ME} = \inf_{U} \|x\|_w$, where $U = \{x(t) | x(t) = \square(t) \text{ for } t \in E\}$. 
Examples of Wiener Norms

If $\hat{x} \in L'(\mathbb{R})$, then $\|x\|_w = \int_{\mathbb{R}} |\hat{x}(\omega)| d\omega$.

If $x(t) = \sum_{j=1}^{k} a_j e^{i\alpha_j t}$, $a_j \in \mathbb{R}$, then $\|x\|_w = \sum_{j=1}^{k} |a_j|$.

If $\alpha \geq 0$ and $x(t) = F_{\alpha}(t)$, then $\|x\|_w = x(0)$.

So $\|x\|_w = 1$ for $x(t) = e^{-t^2}$, $e^{-|t|}$, $(1 + t^2)^{-1}$,

or $T(t) = \begin{cases} 1 & |t| \leq 1 \\ 0 & |t| > 1 \end{cases}$.

If $y(t) = x(at + b)$ for $a > 0$, $b \in \mathbb{R}$, then $\|y\|_w = \|x\|_w$. 
Energy Spreading Results

A an \( N \times N \) Hadamard matrix, \( x \) a discrete signal of length \( N \), \( S \) a set of such \( x \)'s.

**Trivial bound:** \[ \|Ax\| \leq \sqrt{N}\|x\| \] for all \( x \).

**Claim:** If, for some constant \( M \) not too much larger than 1, \[ \|Ax\| \leq M\|x\| \] for all \( x \in S \), then \( A \) gives good energy spreading on \( S \).
**Theorem:** If $A$ is a PONS matrix, then

$$\|Ax\| \leq \sqrt{2}\|x\|_{ME} \text{ for all } x.$$ 

**Example:** $x = [x' \mid x'']$, $k + l = N$, $A$ PONS

$x' = [a \cos(k \mathcal{W}_1 + b), \ldots, a \cos(k \mathcal{W}_1 + b)]$

$x'' = [c \cos(l \mathcal{W}_2 + d), \ldots, c \cos(l \mathcal{W}_2 + d)]$

Then $\|Ax\| \leq \frac{3\sqrt{6}}{5}\sqrt{a^2 + c^2}$
Related Engineering Work

• Communications
  – Golay
  – Budisin
  – Popovic
  – Boche
  – Stanczak

• Radar
  – Welti
  – Moran
Robustness of PONS-Coded Data in Bursty Channels
PONS Encoding for Burst Error Mitigation

- The PONS transform spreads the energy of time-localized (spatially localized) phenomena approximately evenly among all the coefficients.
- Transmission of PONS coefficients in a medium prone to localized corruption (e.g., burst noise) averages the corruption across all coefficients.
- In the presence of quantization, the effects of modest local distortions can be completely removed.
- When distortion is significant and not localized, PONS coding can be worse than no coding.
- PONS coding at the waveform level can be combined with error detection/correction coding at the source level.
- In two-dimensional applications (e.g., holographic data storage) where the signal mean is non-zero, the absence of a DC term in PONS is an advantage over other Hadamard transforms (e.g., Walsh).
Localized Burst Noise
Reconstructed PONS-Coded Image
Question (Ingrid Daubechies)
1991, Oberwolfach Wavelet Conference

Does there exist a smooth basis whose elements satisfy the optimal Global Uncertainty Principle bounds?

Answer: Yes

Proof: Apply a PONS Transform to an appropriately chosen smooth basis
Converting a Basis to “PONS type”

\[ \left\{ \square_j(t) \right\}_{j=0} \text{ an ONB for } S, \text{ a normed linear space on } [0,1]. \]

For each \( r \geq 0 \), \( P^{(r)} \) is the \( 2^r \times 2^r \) PONS matrix.

Define \( \overline{\square}^{(r)} = \left\langle \square_{2^r1}(t), \square_2(t), \square_{2^r+1}(t), \ldots, \square_{2(2^r1)}(t) \right\rangle^T \)

and \( \overline{W}^{(r)} = \frac{1}{\sqrt{2^r}} P^{(r)} \overline{\square}^{(r)} = \left\langle W_{2^r1}(t), \ldots, W_{2(2^r1)}(t) \right\rangle^T. \)

Then \( \overline{\square}^{(r)} = \frac{1}{\sqrt{2^r}} P^{(r)} \overline{W}^{(r)} \) and \( \left\{ W_j(t) \right\}_{j=0} \) is an ONB for \( S \).
Conversion

\[ W_0(t) = \Box_0(t) \]

\[
\begin{bmatrix}
W_1(t) \\
W_2(t)
\end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix}
1 \\
1
\end{bmatrix} \begin{bmatrix}
\Box_1(t) \\
\Box_2(t)
\end{bmatrix}
\]

\[
\begin{bmatrix}
W_3(t) \\
W_4(t)
\end{bmatrix} = \frac{1}{2} \begin{bmatrix}
1 \\
1
\end{bmatrix} \begin{bmatrix}
\Box_3(t) \\
\Box_4(t)
\end{bmatrix}
\]

\[
\begin{bmatrix}
W_5(t) \\
W_6(t)
\end{bmatrix} = \frac{1}{2} \begin{bmatrix}
1 \\
1
\end{bmatrix} \begin{bmatrix}
\Box_5(t) \\
\Box_6(t)
\end{bmatrix}
\]

Smoothness properties of \( \{W\} \)

□ smoothness properties of \( \{W\} \)

**Problem:**

Choose a smooth basis \( \{\square_n\} \) such that the (necessarily smooth) basis \( \{W_n\} \) satisfies the global uncertainty principle inequalities:

\[
\left| W_j(t) \right| \square C \quad \text{and} \quad \left| \int_0^1 W_j(t) e^{\square i \square t} dt \right| \square \frac{C}{\sqrt{1 + \square}}
\]
A Smooth PONS Basis

Take \( \{ \mathcal{B}_j(t) \} \) \( j = 0 \to \infty \) = \( \{ e^{2\pi i k t} \} \) \( k = -\infty \to \infty \). Let \( \mathcal{B}_0(t) = e_0 \) and, for each \( m \geq 1 \), let

\[
\mathcal{B}_m(t) = e^{2\pi i (j+1) \cdot 2^m} , \quad 2^m \cdot 1 \cdot j + 2^m + 2^m \cdot 1 \cdot j + 2^m + 2^m \cdot 1 \cdot j + 2^m + 2^m \cdot 1 \cdot j + 2^m + 2^m \cdot 1 \cdot j + 2^m .
\]

Thus \( \{ \mathcal{B}_0, \mathcal{B}_1, \mathcal{B}_2, \ldots \} = \{ e_0, e_{-1}, e_1, e_{-2}, e_2, e_{-3}, e_3, e_{-4}, e_4, e_{-5}, e_5, e_{-6}, e_6, e_{-7}, \ldots \} \).
With this ordering, for any $r \geq 0$, and any $k$, $2^r - 1 \leq k \leq 2^{r+1} - 2$,

$$W_k(t) = \frac{1}{\sqrt{2^r}} \prod_{j=1}^{2^r} p_{k2^r+2}^{(r)}(j)e_{2^r+2}^0(t)$$

$$= \frac{1}{\sqrt{2^r}} \prod_{j=1}^{2^r} p_{k2^r+2}^{(r)}(j)e^0_{2^r+1} + \frac{1}{\sqrt{2^r}} \prod_{j=2^r+1}^{2^r} p_{k2^r+2}^{(r)}(j)e_{j}^0$$

$|W_k(t)| \leq 2$ for all $r, k, t$.

Need bound for $\hat{W}_k(t) = \prod_{0}^{1} W_k(t)e^{2i\pi t} dt$. 

Lemma 1

The crest factor of any finite section

\[ Q(z) = \prod_{j=k_1}^{k_2} p(j)z^j \quad (k_1 < k_2 < 2^r) \]

of any PONS polynomial

\[ P(z) = \prod_{j=0}^{2^r-1} p(j)z^j, \]

where \( \{p(j)\}_{j=k_1}^{k_2} \) is the consecutive set of entries of index \( k_1 \) through \( k_2 \) of any row of any \( P^{(r)} \), is less than 5.
Lemma 2

Let \( \{q(j)\}_{j=N}^{2N-1} \) be \( N \) consecutive entries of any \( P^{(r)} \) \( (2^r \geq N) \), and let

\[
\mathcal{Q}(N, \square) = \left(1 \mathcal{Q} e^{2\mathcal{Q}i \square} \right)^{2N \square} \frac{q(j)}{j + \square}.
\]

Then \( \mathcal{Q}(N, \square) \leq C \).
Graphs for $W_7$
Graphs for $W_{17}$

$|W_{17}(t)|$

$|\hat{W}_{17}(\omega)|$
PONS-Related Problems

2. *Correlations*. Construct collections of $K$ sequences, each of length $N$, where the periodic and aperiodic cross- and autocorrelations for shifts up to $\pm M$ (where $M < \sqrt{N}$ certainly, and usually even smaller) are all “small” (except, of course, for the 0 shift in the auto case). All elements of all sequences must have modulus one. Ideally they would be $\pm 1$ (“bipolar sequences”), although “quadri-phase sequences” (entries $\pm 1, \pm i$) are also very interesting. Sequences whose elements are other roots of unity are also interesting, but less so. Typical values for the parameters are $N = 128$ or 256, $N \leq K \leq 8N$, and $M$ about 6.
PONS-Related Problems

3. *Barker*. Let $P$ and $Q$ have coefficients of modulus one and the same length (say $L$, or degree $L - 1$), $|P|^2 + |Q|^2 = 2L$ on $|z| = 1$ and $R = P^2 + zQ^2$. Obviously the $L^1$ norm of $R$ is $\| 2L$. Prove that it is $\| 2L - 1$. From this it would follow that there are no more Barker sequences.


5. *Fourier transform approximation.* Given $f(t)$ in $L^1$ and $L^2$ of $R$, $\square > 0$, and $b > 0$, must there exist a $g(t)$ supported on $[-b, b]$ such that $|\hat{f} - \hat{g}| < \square$ on $R$? If yes, find an algorithm for computing $g(t)$. Note: You can assume $\hat{f}$ has compact support (if that helps).
References