The Optimal Transform for the Discrete Hirschman Uncertainty Principle

Tomasz Przebinda, Victor DeBrunner, *Senior Member, IEEE*, and Murad Özaydın

***Abstract*—We determine all signals giving equality for the discrete Hirschman uncertainty principle. We single out the case where the entropies of the time signal and its Fourier transform are equal. These signals (up to scalar multiples) form an orthonormal basis giving an orthogonal transform that optimally packs a finite-duration discrete-time signal. The transform may be computed via a fast algorithm due to its relationship to the discrete Fourier transform.**

***Index Terms*—Entropy, information measures, orthogonal functions, signal representation theory.**

I. Introduction

In [1], we introduced the weighted average of the entropies of a discrete-time signal and its Fourier transform H_p that measures the concentration of a signal in the sample-frequency phase plane. This was used to show that discretized Gaussian pulses may not be the most compact basis [2], and a lower limit on the compaction in the phase plane was conjectured. We have since discovered that part of this conjectured lower limit was proven in [3] under the moniker of "a discrete Hirschman's uncertainty principle." This principle states that $H_{\frac{1}{2}}$ is at least $\frac{1}{2}\log(N)$, where N is the length of the discrete-time signal. However, that result did not describe the characteristics of the signals that meet the limit, as our conjecture did [1], [4]. We further argued in [5] that this measure indicates two possible "best basis" options:

1) the multitransform (nonorthogonal) option,
2) the orthogonal discrete Hirschman uncertainty principle option.

We have discussed many results in the first option (see [1] for pointers to many references). The second option is the focus of this correspondence. We have found a basis (transform) that is orthogonal and that uniquely minimizes the discrete Hirschman uncertainty principle.

II. Statement of the Main Theorem

Fix a positive integer N. Let A denote the ring $\mathbb{Z}/N\mathbb{Z}$. Thus $A = \{0, 1, 2, \ldots, N-1\}$, with the addition and multiplication modulo N. Often we shall view A as a group with respect to the addition.

The Heisenberg group of degree one, with coefficients in A, is the group $G_1(A)$ of all matrices of the form

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \qquad (x, y, z \in A).$$

We shall identify $G_1(A)$ with the Cartesian product $A \times A \times A$ via the map

$$G_1(A) \ni \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \to (x, y, z) \in A \times A \times A. \tag{1}$$

In terms of (1), the matrix multiplication and the inverse look as follows:

$$(x, y, z)(x', y', z') = (x+x', y+y', z+z'+xy'),$$
$$(x, y, z)^{-1} = (-x, -y, -z+xy)(x, y, z, x', y', z' \in A).$$

Let

$$\chi(a) = \exp(2\pi j a/N) \qquad (a \in A).$$

This is a unitary character of the (additive) group A. For a function $u: A \to \mathbb{C}$ let

$$\|u\|_2 = \left(\sum_{a \in A} |u(a)|^2\right)^{1/2} \tag{2}$$

and let $L^2(A)$ denote the Hilbert space of all such functions, with the norm (2). Let

$$\rho(x, y, z)u(a) = \chi(ay + z)u(a + x)(u \in L^2(A);\ a, x, y, z \in A). \tag{3}$$

It is easy to check that ρ is a group homomorphism from $G_1(A)$ to the group of unitary operators on $L^2(A)$. In other words, ρ is a unitary representation of $G_1(A)$ on the space $L^2(A)$.

Recall the discrete Fourier transform (DFT), defined with respect to the character χ

$$\mathcal{F}u(b) = \hat{u}(b) = |A|^{-1/2} \sum_{a \in A} u(a)\chi(-ab) \qquad (u \in L^2(A),\ b \in A).$$

Here $|A| = N$ is the cardinality of the set A. The inverse Fourier transform is given by

$$u(a) = |A|^{-1/2} \sum_{b \in A} \hat{u}(b)\chi(ab) \qquad (u \in L^2(A),\ a \in A).$$

A straightforward calculation shows that

$$\mathcal{F}\rho(x, y, z)\mathcal{F}^{-1} = \rho(-y, x, z - xy) \qquad (x, y, z \in A). \tag{4}$$

In other words, the Fourier transform normalizes the group $\rho(G_1(A))$. For $u \in L^2(A)$, with $\|u\|_2 = 1$, let

$$H(u) = -\sum_{a \in A} |u(a)|^2 \log\left(|u(a)|^2\right)$$

and let

$$H_p(u) = pH(u) + (1-p)H(\hat{u}) \qquad (0 \le p \le 1).$$

Manuscript received October 5, 1999; revised January 25, 2001. This work was supported in part by the National Science Foundation under Grant DMS-9622610.

T. Przebinda and M. Özaydın are with the Department of Mathematics, The University of Oklahoma, Norman, OK 73019 USA (e-mail: przebin@crystal.ou.edu; mozaydin@ou.edu).

V. DeBrunner is with the School of Electrical and Computer Engineering, The University of Oklahoma, Norman, OK 73019 USA (e-mail: vdebrunn@ou.edu).

Communicated by J. A. O'Sullivan, Associate Editor for Detection and Estimation.

Publisher Item Identifier S 0018-9448(01)04430-3.

It is easy to see that

$$H_p(\rho(h)u) = H_p(u) \qquad (h \in G_1(A),\ 0 \le p \le 1).$$

We would like to consider $u \in L^2(A)$ with $\|u\|_2 = 1$ equivalent to $v = \lambda u$ where $|\lambda| = 1$. As $H(u) = H(v)$ and $H_p(u) = H_p(v)$ for equivalent u and v, H and H_p are defined on the equivalence classes. This set of equivalence classes forms a complex projective space which we will denote by $P(A)$. Note that being orthogonal is well-defined on the equivalence classes, so a subset of $P(A)$ being orthonormal makes sense. There is an induced action of the Heisenberg group $G_1(A)$ on $P(A)$ defined via (3) at the level of representatives for the equivalence classes. Below we will use the same symbol u for an element of $L^2(A)$ with $\|u\|_2 = 1$ and the equivalence class it represents in $P(A)$.

If B is a subset of A, let $\mathbb{1}_B$ denote the indicator function of B. Thus, $\mathbb{1}_B(a) = 1$ if $a \in B$, and $\mathbb{1}_B(a) = 0$ if $a \in A \setminus B$. Here is our main theorem.

Theorem 1 (Main Theorem):

a) If $u \in P(A)$, then $H_{1/2}(u) \ge \frac{1}{2}\log(|A|)$.

b) The set of vectors $u \in P(A)$ and $H_{1/2}(u) = \frac{1}{2}\log(|A|)$ coincides with the union of the orbits

$$\rho(G_1(A))\frac{1}{\sqrt{|B|}}\,\mathbb{1}_B(B\text{—a subgroup of } A). \tag{5}$$

c) Each orbit (5) is an orthonormal basis of $L^2(A)$.

d) The set of vectors $u \in P(A)$ and $H_p(u) = \frac{1}{2}\log(|A|)$ for all $0 \le p \le 1$ is not empty if and only if $|A|$ is a square. In this case, this set coincides with the orbit (5) for the unique subgroup $B \subseteq A$ of cardinality $|B| = \sqrt{|A|}$.

Part a) of the above theorem has been proven by Dembo, Cover and Thomas, [3]. The idea of their proof is based on Hirschman's work, [6]. In fact, those authors name the inequality a) "the discrete Hirschman uncertainty principle." Following this line, we have chosen the title of this correspondence. While unaware of the work in [3], we conjectured a result close to the above theorem in [1]. The conjecture was refined in [4].

The strategy of the proof of part b) is to reduce it to a result of Donoho and Stark [7, Theorem 13]. In order to keep the presentation self-contained, we give proofs for this as well as part a). Part c) suggests a close connection of the functions listed in b) with wavelets, along the lines explored partially in [8].

A generalization of parts a) and b) of the Main Theorem, where the finite cyclic group A is replaced with a compactly generated, locally compact abelian group is available [9]. This includes multidimensional finite (A is a product of finite cyclic groups), continuous ($A = \mathbb{R}^N$), and periodic ($A = (\mathbb{R}/\mathbb{Z})^N$) cases, as well as their products.

III. Proof of the Main Theorem

For a function $u\colon A \to \mathbb{C}$ and a number $0 < p < \infty$ let

$$\|u\|_p = \left(\sum_{a\in A}|u(a)|^p\right)^{1/p}.$$

Also, let

$$\|u\|_\infty = \max\{|u(a)|;\ a \in A\}.$$

A straightforward calculation shows that for a nonzero function $u\colon A \to \mathbb{C}$ and for a number $0 < t < \infty$ the following formulas hold:

$$\frac{d}{dt}\log\|u\|_{1/t} = -\sum_{a\in A}\frac{|u(a)|^{1/t}}{\|u\|_{1/t}^{1/t}}\log\left(\frac{|u(a)|^{1/t}}{\|u\|_{1/t}^{1/t}}\right) \tag{6}$$

and

$$\frac{d^2}{dt^2}\log\|u\|_{1/t} = \frac{1}{t}\sum_{a\in A}\frac{|u(a)|^{1/t}}{\|u\|_{1/t}^{1/t}} \cdot \left(\log\left(\frac{|u(a)|^{1/t}}{\|u\|_{1/t}^{1/t}}\right) - \sum_{b\in A}\frac{|u(b)|^{1/t}}{\|u\|_{1/t}^{1/t}}\log\left(\frac{|u(b)|^{1/t}}{\|u\|_{1/t}^{1/t}}\right)\right)^2.$$

Since the second derivative is nonnegative, the function $\log\|u\|_{1/t}$, $0 < t < \infty$, is convex. Hence, for $u \in L^2(A)$ with $\|u\|_2 = 1$

$$H(u) = \frac{d}{dt}\log\|u\|_{1/t}\bigg|_{t=\frac{1}{2}} \le \lim_{t\to+\infty}\frac{d}{dt}\log\|u\|_{1/t} = \log(|\operatorname{supp} u|) \tag{7}$$

where $|\operatorname{supp} u|$ stands for the cardinality of $\operatorname{supp} u$, the support of u. The inequality (7) is of course well known.

Since $\|\hat{u}\|_2 = \|u\|_2$, and since

$$\|\hat{u}\|_\infty \le |A|^{-1/2}\|u\|_1$$

the Riesz–Thorin theorem, [10, Ch. 12, eq. (1.11)], implies

$$\frac{\|\hat{u}\|_{1/(1-t)}}{\|u\|_{1/t}} \le |A|^{\frac{1}{2}-t} \qquad \left(\frac{1}{2} \le t \le 1,\ u \ne 0\right). \tag{8}$$

By applying negative logarithm to both sides of (8) we obtain the following inequality:

$$\log(\|u\|_{1/t}) - \log(\|\hat{u}\|_{1/(1-t)}) \ge \left(t - \frac{1}{2}\right)\log(|A|) \qquad \left(\frac{1}{2} \le t \le 1\right). \tag{9}$$

As an aside, notice that the left-hand side of (9) is a difference of two convex functions.

We assume from now on that $\|u\|_2 = 1$. Then both sides of (9) are equal to zero for $t = \frac{1}{2}$. Hence, (6) and (7) imply

$$H(u) + H(\hat{u}) \ge \log(|A|). \tag{10}$$

This verifies part a) of the theorem as in [3].

We are interested in functions u for which the equality holds in (10). We are going to use some ideas of Zygmund, [10, Ch. 12, eqs. (1.20)–(1.24)]. For a complex number $z \in \mathbb{C}$ define

$$f(z) = |A|^{-\frac{1}{2}+z}\sum_{b\in A}\mathcal{F}\left(|u|^{2z}\frac{u}{|u|}\right)(b)|\hat{u}(b)|^{2z}\frac{\overline{\hat{u}(b)}}{|\hat{u}(b)|}.$$

Here $\frac{u}{|u|} = 0$ outside the support of u, and, similarly, for $\frac{\hat{u}}{|\hat{u}|}$.

Notice that for $y \in \mathbb{R}$

$$\left|f\left(\tfrac{1}{2}+iy\right)\right| \le \||u|^{1+i2y}\|_2 \cdot \||\hat{u}|^{1+i2y}\|_2 = \|u\|_2 \cdot \|\hat{u}\|_2 = 1 \cdot 1 = 1,$$

and

$$|f(1+iy)| \le |A|^{\frac{1}{2}}\left\|\mathcal{F}\left(|u|^{2+i2y}\frac{u}{|u|}\right)\right\|_\infty \cdot \||\hat{u}|^{2+i2y}\|_1 \le \|u\|_2^2 \cdot \|\hat{u}\|_2^2 = 1.$$

Hence, by the Phragmén and Lindelöf theorem, [10, Ch. 12, eq. (1.1)]

$$|f(z)| \le 1 \qquad \left(\frac{1}{2} \le \operatorname{Re}(z) \le 1\right).$$

A straightforward calculation shows that

$$\begin{aligned}\frac{d}{dz}f(z) &= f'(z)\\ &= f(z)\log(|A|) + |A|^{-\frac{1}{2}+z}\sum_{b\in A}\mathcal{F}\\ &\quad\cdot\left(|u|^{2z}\frac{u}{|u|}\log(|u|^2)\right)(b)|\hat{u}(b)|^{2z}\frac{\overline{\hat{u}(b)}}{|\hat{u}(b)|}\\ &\quad+|A|^{-\frac{1}{2}+z}\sum_{b\in A}\mathcal{F}\left(|u|^{2z}\frac{u}{|u|}\right)(b)|\hat{u}(b)|^{2z}\\ &\quad\cdot\frac{\overline{\hat{u}(b)}}{|\hat{u}(b)|}\log(|\hat{u}(b)|^2).\end{aligned}$$

Hence, by Plancherel's formula

$$f'\left(\frac{1}{2}\right) = \log(|A|) - H(u) - H(\hat{u}).$$

Thus, the equality in (10) is equivalent to $f'(\frac{1}{2}) = 0$. Altogether, we have checked that the function $f(z)$ has the following properties: $f(z)$ is an entire function

$$|f(z)| \le 1, \qquad \text{for } \frac{1}{2} \le \mathrm{Re}(z) \le 1,\ f\left(\frac{1}{2}\right) = 1, \text{ and } f'\left(\frac{1}{2}\right) = 0.$$

In particular, $\mathrm{Re}(f(z))$ is a real-valued harmonic function in the disc of radius $\frac{1}{4}$ centered at $z = \frac{3}{4}$. This harmonic function achieves its maximum at $z = \frac{1}{2}$, and has derivative equal to zero at this point. Hence, the Hopf's Maximum Principle [11, Theorem 3.1.6'], implies that $\mathrm{Re}(f(z))$ is constant on this disc. Hence, by standard properties of entire functions, $f(z) = 1$ for all $z \in \mathbb{C}$. This equation coincides with the formula [10, Ch. 12, eq. (1.24)], which has been obtained there under a slightly stronger assumption, [10, Ch. 12, eq. (1.20)]. In particular, for $z = 1$ we obtain

$$1 = f(1) = |A|^{\frac{1}{2}}\sum_{b\in A}\mathcal{F}\left(|u|u\right)(b)|\hat{u}(b)|\overline{\hat{u}(b)}. \tag{11}$$

Now we follow Zygmund's proof of [10, Ch. 12, eq. (2.18)].

The formula (11) may be rewritten as

$$1 = \sum_{a,\,b\in A}|u(a)|^2|\hat{u}(b)|^2\chi(-ab)\frac{u(a)}{|u(a)|}\frac{\overline{\hat{u}(b)}}{|\hat{u}(b)|}. \tag{12}$$

Since

$$\sum_{a,\,b\in A}|u(a)|^2|\hat{u}(b)|^2 = \|u\|_2^2\cdot\|\hat{u}\|_2^2 = 1$$

(12) implies

$$1 = \chi(-ab)\frac{u(a)}{|u(a)|}\frac{\overline{\hat{u}(b)}}{|\hat{u}(b)|} \qquad (a \in \mathrm{supp}\,u,\ b \in \mathrm{supp}\,\hat{u}).$$

Hence, for $b \in \mathrm{supp}\,\hat{u}$

$$\begin{aligned}\hat{u}(b) &= |A|^{-\frac{1}{2}}\sum_{a\in\mathrm{supp}\ u}u(a)\chi(-ab)\\ &= |A|^{-\frac{1}{2}}\sum_{a\in\mathrm{supp}\ u}u(a)\frac{\overline{u(a)}}{|u(a)|}\frac{\hat{u}(b)}{|\hat{u}(b)|}.\end{aligned}$$

By taking the absolute value of the extreme left and right sides of the above equations, we get

$$|\hat{u}(b)| = |A|^{-\frac{1}{2}}\sum_{a\in\mathrm{supp}\ u}|u(a)| \qquad (b \in \mathrm{supp}\,\hat{u}). \tag{13}$$

Similarly, for $a \in \mathrm{supp}\,u$

$$\begin{aligned}u(a) &= |A|^{-\frac{1}{2}}\sum_{b\in\mathrm{supp}\ \hat{u}}\hat{u}(b)\chi(ab)\\ &= |A|^{-\frac{1}{2}}\sum_{b\in\mathrm{supp}\ \hat{u}}\hat{u}(b)\frac{\overline{\hat{u}(b)}}{|\hat{u}(b)|}\frac{u(a)}{|u(a)|}\end{aligned}$$

and, therefore,

$$|u(a)| = |A|^{-\frac{1}{2}}\sum_{b\in\mathrm{supp}\ \hat{u}}|\hat{u}(b)| \qquad (a \in \mathrm{supp}\,u). \tag{14}$$

The statements (13) and (14) mean that the functions $|u|$ and $|\hat{u}|$ are constant on their support. Since $\|u\|_2 = 1$, it follows that

$$|u(a)| = |\mathrm{supp}\,u|^{-1/2} \text{ and } |\hat{u}(b)| = |\mathrm{supp}\,\hat{u}|^{-1/2} \qquad (a \in \mathrm{supp}\,u,\ b \in \mathrm{supp}\,\hat{u}).$$

Hence,

$$H(u) + H(\hat{u}) = \log(|\mathrm{supp}\,u|) + \log(|\mathrm{supp}\,\hat{u}|).$$

Thus, the equality in (10) implies

$$|\mathrm{supp}\,u|\cdot|\mathrm{supp}\,\hat{u}| = |A|.$$

Thus, part b) of the theorem will follow as soon as we verify the following theorem of Donoho and Stark, [7].

Theorem 2: Let $v \in L^2(A)$. Then the equation

$$|\mathrm{supp}\,v|\cdot|\mathrm{supp}\,\hat{v}| = |A|$$

holds if and only if there is a subgroup $B \subseteq A$, an element $h \in G_1(A)$, and a constant "$const$" such that $v = const\,\rho(h)\mathbb{1}_B$.

Lemma 3 [7]: Let $v \in L^2(A)$ and let $m = |\mathrm{supp}\,u|$. Then $\hat{v}$ cannot have more than m consecutive zeros.

Proof: Since the translations of $\hat{v}$ do not effect the support of v, it shall suffice to show that

$$(\hat{v}(0),\,\hat{v}(1),\,\ldots,\,\hat{v}(m-1)) \neq (0,\,0,\,\ldots,\,0). \tag{15}$$

Let $\mathrm{supp}\,v = \{a_1,\,a_2,\,\ldots,\,a_m\}$. Then

$$\begin{bmatrix}\hat{v}(0)\\ \hat{v}(1)\\ \hat{v}(2)\\ \vdots\\ \hat{v}(m-1)\end{bmatrix} = |A|^{-\frac{1}{2}}\begin{bmatrix}1 & 1 & \cdots & 1\\ \chi(-a_1) & \chi(-a_2) & & \chi(-a_m)\\ \chi(-a_1)^2 & \chi(-a_2)^2 & & \chi(-a_m)^2\\ \vdots & & \ddots & \\ \chi(-a_1)^{m-1} & \chi(-a_2)^{m-1} & & \chi(-a_m)^{m-1}\end{bmatrix}\cdot\begin{bmatrix}v(a_1)\\ v(a_2)\\ v(a_3)\\ \vdots\\ v(a_m)\end{bmatrix}$$

Since, by Vandermonde, the above $m \times m$-matrix is invertible, (15) follows, and we are done.

Next we recall a few facts concerning the Fourier transform. For a subset $S \subseteq A$ let

$$S^{\perp} = \{a \in A; ab = 0 \text{ for all } b \in S\}.$$

It is easy to see that $S^{\perp}$ is a subgroup of A, and that $S^{\perp\perp}$ is the smallest subgroup of A containing S. Furthermore, v is invariant under translations by $(\operatorname{supp} \hat{v})^{\perp}$, i.e.,

$$v(a+b) = v(a) \qquad (a \in A,\ b \in (\operatorname{supp} \hat{v})^{\perp}). \tag{16}$$

Here is a statement dual to (16)

$$\text{for a subgroup } B \subseteq A, \text{ if } v(a+b) = v(a) \text{ for } a \in A \text{ and } b \in B, \text{ then } \operatorname{supp} \hat{v} \subseteq B^{\perp}.$$

An elementary counting argument shows that for any subgroup $B \subseteq A$

$$|A| = |B||B^{\perp}| \tag{17}$$

and

$$\left\| \frac{1}{\sqrt{|B|}} \mathbb{1}_B \right\|_2 = 1$$

$$H\left(\frac{1}{\sqrt{|B|}} I_B \right) = \log(|B|)$$

and

$$\mathcal{F}\left(\frac{1}{\sqrt{|B|}} \mathbb{1}_B \right) = \frac{1}{\sqrt{|B^{\perp}|}} \mathbb{1}_{B^{\perp}}. \tag{18}$$

Proof of Theorem 3.1: Let $B \subseteq A$ be a subgroup and let $h \in G_1(A)$. We know from (4) and (18) that there is $h' \in G_1(A)$ such that

$$\mathcal{F}\left(\rho(h) \frac{1}{\sqrt{|B|}} \mathbb{1}_B \right) = \mathcal{F}\rho(h)\mathcal{F}^{-1}\mathcal{F} \frac{1}{\sqrt{|B|}} \mathbb{1}_B = \rho(h') \frac{1}{\sqrt{|B^{\perp}|}} \mathbb{1}_{B^{\perp}}.$$

Hence, by (17)

$$\left| \operatorname{supp} \mathcal{F}\left(\rho(h) \frac{1}{\sqrt{|B|}} \mathbb{1}_B \right) \right| \cdot \left| \operatorname{supp} \rho(h) \frac{1}{\sqrt{|B|}} \mathbb{1}_B \right| = |B^{\perp}| \cdot |B| = |A|.$$

Conversely, suppose $v \in L^2(A)$ is such that $|\operatorname{supp} v| \cdot |\operatorname{supp} \hat{v}| = |A|$. Then the lemma implies that the elements of $\operatorname{supp} \hat{v}$ are equally spaced. Hence, there is $h \in G_1(A)$ such that $\operatorname{supp} \rho(h) v$ is a subgroup of A. Thus, we may assume that $\operatorname{supp} \hat{v}$ is a subgroup of A. Let B be the unique subgroup of A such that $B^{\perp} = \operatorname{supp} \hat{v}$. Then v is invariant under translations by elements of B, by (16). In particular, $|\operatorname{supp} v|$ is a multiple of $|B|$. But our assumption implies that $|\operatorname{supp} v| = |A|/|B^{\perp}| = |B|$. Hence, v is a translation of $\mathbb{1}_B$.

This completes the proof of part b) of the Main Theorem. Part d) of the Main Theorem is immediate from part b) because the equation

$$H_p(u) = \tfrac{1}{2} \log(|A|) \qquad (0 \le p \le 1)$$

is equivalent to

$$H(u) = H(\hat{u}) = \tfrac{1}{2} \log(|A|)$$

which, for $u = \frac{1}{\sqrt{|B|}} \mathbb{1}_{\mathbb{B}}$, becomes $|B|^2 = |A|$, by (18).

It remains to verify part c) of the Main Theorem. A straightforward argument shows that, under the isomorphism (1), the stabilizer of the complex line $\mathbb{C}\mathbb{1}_B$, in $G_1(A)$, is given by

$$Stab_{G_1(A)}(\mathbb{C}\mathbb{1}_B) = B \times B^{\perp} \times A.$$

This is a normal subgroup of $G_1(A)$, the quotient group $G_1(A)/Stab_{G_1(A)}(\mathbb{C}\mathbb{1}_B)$ is isomorphic to $(A/B) \times (A/B^{\perp})$, via (1), and, by (17), has $(|A|/|B|)(|A|/|B^{\perp}|) = |A|$ elements. Thus, the number of distinct elements in the orbit $\rho(G_1(A)) \frac{1}{\sqrt{|B|}} \mathbb{1}_B$, (5), coincides with the dimension of the space $L^2(A)$.

It remains to check that any two distinct elements of this orbit are orthogonal. Since the representation ρ is unitary, it shall suffice to show that

$$\sum_{a \in A} \rho(x,\ y,\ 0) \cdot \mathbb{1}_B(a)\mathbb{1}_B(a) = 0 (x \in A \setminus B \text{ or } y \in A \setminus B^{\perp}). \tag{19}$$

The left-hand side of (19) is equal to

$$\sum_{a \in B \cap (-x+B)} \chi(ay) = \sum_{a \in B \cap (x+B)} \chi(ay). \tag{20}$$

If $x \in A \setminus B$, then $B \cap (x+B)$ is empty, so the quantity (20) is zero. If $x \in B$, then $B \cap (x+B) = B$, so the quantity (20) is equal to $\mathcal{F}(\mathbb{1}_B)(y) = 0$, by (18). This completes our proof of part c) of the Main Theorem, and thus of the whole theorem.

IV. The Hirschman Optimal Transform

Now that we have seen the theorem that actually defined the discrete Hirschman uncertainty principle optimal transform (HOT), we provide details regarding the transform.

A. The HOT Basis Functions

The basis functions that define the HOT are derived according to part b) of the Main Theorem, and those that are suggested in [7]. Consequently, we use the K-dimensional DFT as the originator signals for our $N = K^2$-dimensional HOT basis. Each of these basis functions must then be shifted and interpolated to produce the sufficient number of orthogonal basis functions that define the HOT. We note that the DFT basis can be extended in a similar manner to produce an $N = KL$-dimensional transform. This basis, however, does not yield a HOT.

To detail this process, consider the three-point DFT defined

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{-i\frac{2\pi}{3}} & e^{-i\frac{4\pi}{3}} \\ 1 & e^{-i\frac{4\pi}{3}} & e^{-i\frac{8\pi}{3}} \end{bmatrix} \begin{bmatrix} x[0] \\ x[2] \\ x[3] \end{bmatrix}.$$

This three-point DFT yields the nine-point HOT shown at the bottom of the next page.

This organization is not unique—the rows can be reordered as desired. This representation would be consistent with the DFT. The Matlab source that implements the general version of the HOT is shown:

```
function H = hot(x);
% This function implements an N = K^2 Hirschman
  optimal transform
% H = hot(x);
% Input: x is a sequence of length N = K^2
% Output: H is the transform sequence
[N, M] = size(x);
K = sqrt(N);
T = zeros(size(N));
W = fft(eye(K));
n = 1 : K : N;
for tr = 0 : K - 1
  T(n + tr, n + tr) = W;
end
T = (1/sqrt(K)) * T;
H = T * x;
```

In this script, `H` is the transform sequence, `T` is the transform, and `x` is the input sequence. The transform is unitary to a scale (just like the DFT), and so the inverse transform can be achieved by taking the conjugate transpose and scaling by $\sqrt{K}$.

B. Fast HOT Computation

Because the HOT is based on periodic shifts of the DFT, the $N = K^2$-point HOT can be accomplished using K separate K-point DFT computations. Because the HOT requires lengths N that are squares of integers, the efficiency of any computational procedure will depend on the exact length N. For $N = 4, 16, 64, 256$, etc., this provides a fast HOT that requires $O(N \log K)$ computations. For other lengths N, the efficiency is less. For instance, in the $N = 9$-point HOT shown above, we can see that the HOT transform coefficients are determined from

$$\begin{bmatrix} H(0) \\ H(3) \\ H(6) \end{bmatrix} = \mathcal{DFT}\left\{ \begin{bmatrix} x[0] \\ x[3] \\ x[6] \end{bmatrix} \right\}$$

and

$$\begin{bmatrix} H(1) \\ H(4) \\ H(7) \end{bmatrix} = \mathcal{DFT}\left\{ \begin{bmatrix} x[1] \\ x[4] \\ x[7] \end{bmatrix} \right\}$$

and, finally, that

$$\begin{bmatrix} H(2) \\ H(5) \\ H(8) \end{bmatrix} = \mathcal{DFT}\left\{ \begin{bmatrix} x[2] \\ x[5] \\ x[8] \end{bmatrix} \right\}.$$

This requires three separate three-point DFT computations. In general, we have the (unitary) transform relationship

$$H(Kr + l) = \frac{1}{\sqrt{K}} \sum_{n=0}^{K-1} x[Kn + l] e^{-j \frac{2\pi}{K} nr}, \qquad 0 \le r, l \le K - 1$$

and its inverse

$$x[Kn + l] = \frac{1}{\sqrt{K}} \sum_{r=0}^{K-1} H(Kr + l) e^{j \frac{2\pi}{K} nr}, \qquad 0 \le n, l \le K - 1.$$

Of course, in practice, the square roots need not be carried out. This is, as is commonly done in the DFT; that is, by moving the one square root out of the analysis relationship and moving it into the synthesis relationship to create the scale $\frac{1}{K}$. The N-point HOT is computationally more efficient than the N-point DFT, and increasingly more efficient as $N \to \infty$. As we have mentioned above, this is somewhat simplistic because the squared integers are not, in general, powers of 2. Consequently, for any length N we should compare specific counts.

Acknowledgment

The authors would like to thank Dr. M. Doroslovacki at George Washington University in Washington, DC, for his comments on [1] that helped lead us to finding some example signals that met the conjectured minimum, and thus ultimately to this proof that defines all such signals that are optimal according to the discrete form of the Hirschman uncertainty principle.

References

[1] V. DeBrunner, M. Özaydın, and T. Przebinda, "Resolution in time-frequency," *IEEE Trans. Signal Processing*, vol. 47, pp. 783–788, Mar. 1999.

[2] V. DeBrunner, M. Özaydın, T. Przebinda, and J. Havlicek, "The optimal solutions to the continuous- and discrete-time versions of the Hirschman uncertainty principle," in *Proc. ICASSP'00*, Istanbul, Turkey, June 5–9, 2000.

[3] A. Dembo, T. M. Cover, and J. A. Thomas, "Information theoretic inequalities," *IEEE Trans Inform. Theory*, vol. 37, pp. 1501–1518, Nov. 1991.

[4] V. DeBrunner, M. Özaydın, and T. Przebinda, "Analysis in a finite time-frequency plane," *IEEE Trans. Signal Processing*, vol. 48, pp. 3586–3587, Dec. 2000.

[5] T. Przebinda, V. E. DeBrunner, and M. Özaydın, "Using a new uncertainty measure to determine optimal bases for signal representations," in *Proc. ICASSP'99*, Phoenix, AZ, Mar. 1999, paper 1575.

[6] I. I. Hirschman, "A Note on entropy," *Amer. J. Math.*, vol. 79, pp. 152–156, 1957.

[7] D. L. Donoho and P. B. Stark, "Uncertainty principles and signal recovery," *SIAM J. Appl. Math.*, vol. 49, pp. 906–931, 1989.

[8] M. Özaydın and T. Przebinda, "Platonic orthonormal wavelets," *Appl. Comput. Harmon. Anal.*, vol. 4, pp. 351–365, 1997.

[9] ——, "An entropy-based uncertainty principle for a locally compact, abelian, compactly generated group," paper, submitted for publication.

[10] A. Zygmund, *Trigonometric Series*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1990, vol. I and II.

[11] L. Hörmander, *Notions of Convexity*. Basel, Switzerland: Birkhäuser, 1994.

$$\begin{bmatrix} H(0) \\ H(1) \\ H(2) \\ H(3) \\ H(4) \\ H(5) \\ H(6) \\ H(7) \\ H(8) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & e^{-i\frac{2\pi}{3}} & 0 & 0 & e^{-i\frac{4\pi}{3}} & 0 & 0 \\ 0 & 1 & 0 & 0 & e^{-i\frac{2\pi}{3}} & 0 & 0 & e^{-i\frac{4\pi}{3}} & 0 \\ 0 & 0 & 1 & 0 & 0 & e^{-i\frac{2\pi}{3}} & 0 & 0 & e^{-i\frac{4\pi}{3}} \\ 1 & 0 & 0 & e^{-i\frac{4\pi}{3}} & 0 & 0 & e^{-i\frac{8\pi}{3}} & 0 & 0 \\ 0 & 1 & 0 & 0 & e^{-i\frac{4\pi}{3}} & 0 & 0 & e^{-i\frac{8\pi}{3}} & 0 \\ 0 & 0 & 1 & 0 & 0 & e^{-i\frac{4\pi}{3}} & 0 & 0 & e^{-i\frac{8\pi}{3}} \end{bmatrix} \begin{bmatrix} x[0] \\ x[1] \\ x[2] \\ x[3] \\ x[4] \\ x[5] \\ x[6] \\ x[7] \\ x[8] \end{bmatrix}.$$