THREE UNCERTAINTY PRINCIPLES FOR AN ABELIAN LOCALLY COMPACT GROUP.

TOMASZ PRZEBINDA

Department of Mathematics, University of Oklahoma, Norman, OK 73019, USA

0. Introduction.

The purpose of this article is to direct reader's attention to some recent developments concerning three Uncertainty Principles: the classical Heisenberg-Weyl Uncertainty Principle, The Hirschman-Beckner Uncertainty Principle based on the notion of the entropy, and the Donoho-Stark Uncertainty Principle. We state all three in the first section with some further explanations and proofs in the following sections.

1. The main results.

Let $S(\mathbb{R})$ denote the Schwartz space on $\mathbb{R}$, [H1], and let $f \in S(\mathbb{R})$ be a real valued function. Then, by the Cauchy inequality,

$$\int_{\mathbb{R}} |xf(x)|^2 \, dx \cdot \int_{\mathbb{R}} |f'(x)|^2 \, dx \geq \left(\int_{\mathbb{R}} xf(x)f'(x) \, dx \right)^2.$$

Moreover,

$$\int_{\mathbb{R}} xf(x)f'(x) \, dx = \int_{\mathbb{R}} x \frac{1}{2} \left(f(x)^2 \right)' \, dx = x\frac{1}{2}f(x)^2|_{-\infty}^{\infty} - \int_{\mathbb{R}} \frac{1}{2} f(x)^2 \, dx = -\frac{1}{2} \parallel f \parallel_2^2.$$

Thus

$$(1.1) \qquad \int_{\mathbb{R}} |xf(x)|^2 \, dx \cdot \int_{\mathbb{R}} |f'(x)|^2 \, dx \geq \frac{\parallel f \parallel_2^4}{4}.$$

If the equality holds in (1.1), then there is a real number a such that

$$f'(x) = -2af(x),$$

so that

(1.2)
$$f(x) = \pm e^{-ax^2 - c},$$

where $c \in \mathbb{R}$. Since $\| f \|_2 < \infty$, the number a must be positive. Let us write f as the inverse Fourier transform of the Fourier transform $\hat{f}$ of f:

$$f(x) = \int_{\mathbb{R}} e^{2\pi i x \xi} \hat{f}(\xi) \, d\xi.$$

Then

$$f'(x) = \int_{\mathbb{R}} 2\pi i \xi \, e^{2\pi i x \xi} \hat{f}(\xi) \, d\xi$$

and therefore

$$\int_{\mathbb{R}} |f'(x)|^2 \, dx = \int_{\mathbb{R}} |2\pi i \xi \hat{f}(\xi)|^2 \, d\xi = 4\pi^2 \int_{\mathbb{R}} |\xi \hat{f}(\xi)|^2 \, d\xi.$$

Thus for $\| f \|_2 = 1$,

(1.3)
$$\int_{\mathbb{R}} |x f(x)|^2 \, dx \cdot \int_{\mathbb{R}} |\xi \hat{f}(x)|^2 \, dx \geq \frac{1}{16\pi^2}.$$

A few more easy steps lead to the following theorem of Herman Weyl, (see [Wi]).

Theorem 1.4, (H. Weyl, 1931). *Let $f \in S(\mathbb{R})$, with $\| f \|_2 = 1$. Set*

$$\mu = \int_{\mathbb{R}} x |f(x)|^2 \, dx, \quad \sigma^2 = \int_{\mathbb{R}} (x - \mu)^2 |f(x)|^2 \, dx,$$

$$\tilde{\mu} = \int_{\mathbb{R}} \xi |\hat{f}(\xi)|^2 \, d\xi, \quad \tilde{\sigma}^2 = \int_{\mathbb{R}} (\xi - \tilde{\mu})^2 |\hat{f}(\xi)|^2 \, d\xi.$$

Then

$$\sigma \cdot \tilde{\sigma} \geq \frac{1}{4\pi},$$

and the equality occurs if and only if

$$f(x) = e^{-ax^2 - bx - c},$$

where $a > 0$, and $b, c \in \mathbb{C}$. In other words f is a constant multiple of a translation

$$f(x) \rightarrow f(x + x_0)$$

and a modulation

$$f(x) \rightarrow e^{i y_0 x} f(x)$$

of a Gaussian

$$e^{-ax^2}.$$

The above theorem states that a function and its Fourier transform cannot both be arbitrarily concentrated. In computer applications one deals often with finite cyclic groups, rather than with the real line. In this context there is no obvious or straightforward generalization of the theorem 1.4, because the quantities involved $(\mu, \sigma, ...)$ do not seem to make sense. In order to circumvent this difficulty, Donoho and Stark, [D-S], have introduced a different, elementary, measure of uncertainty, which we explain below.

Let $A = \{0, 1, 2, 3, ..., N-1\}$ be a finite cyclic group of order $|A| = N$. Let $\hat{A}$ denote the dual group of all the characters (i.e. group homomorphisms $\hat{a} : A \to \mathbb{C}^\times$). It is customary, and sometimes convenient, to identify A with $\hat{A}$ be the formula

$$\hat{a}(b) = e^{\frac{2\pi i}{N}ab} \qquad (a, b \in A).$$

For a function $f : A \to \mathbb{C}$ define a Fourier transform of f by

$$\hat{f}(\hat{b}) = \sum_{a \in A} f(a)\hat{b}(-a) \qquad (\hat{b} \in \hat{A}).$$

Theorem 1.5, (Donoho-Stark, 1989, [D-S]). *For any non-zero function $f : A \to \mathbb{C}$,*

$$|supp\, f| \cdot |supp\, \hat{f}| \geq |A|.$$

The equality occurs if and only if f is a constant multiple of a translation

$$f(a) \to f(a + c)$$

and a modulation

$$f(a) \to \hat{b}(a)f(a)$$

of the indicator function of a subgroup of A.

In applications one deals often with multi-dimensional signals, i.e. with functions defined on a finite product of finite cyclic groups. Hence it is natural to ask for a generalization of the theorem 1.5 to this more general context. This has been done by K. Smith.

Theorem 1.6, (K. Smith, 1990, [Sm]). *The above theorem 1.5 holds for any finite Abelian group A (a finite direct product of finite cyclic groups).*

The theorem 1.6 also follows from theorem 1.14 below. In addition, a proof which uses no more than basic concepts from finite dimensional linear algebra over complex numbers and the structure of finite abelian groups is available in [M-Ö-P].

Clearly, the cardinality of the support is not the most precise measure of the concentration of a function. One possible improvement is based on the notion of entropy.

Definition 1.7, (based on Shannon, 1948, [S]). *Let μ be a non-negative measure on a measure space M. Let $\phi : M \to [0, \infty)$ be a probability density function, i.e.*

$$\int_M \phi(x)\, d\mu(x) = 1.$$

The entropy of ϕ is defined as

$$H(\phi) = - \int_M \phi(x)\, log(\phi(x))\, d\mu(x),$$

where the log stands for the natural logarithm, whenever the integral converges.

In his 1957 paper, [Hi], Hirschman has proven the following theorem and stated the following conjecture.

Theorem 1.8. *Let $f \in S(\mathbb{R})$, with $\| f \|_2 = 1$. Then*

$$H(|f|^2) + H(|\hat{f}|^2) \geq 0.$$

Conjecture 1.9. *Let $f \in S(\mathbb{R})$, with $\| f \|_2 = 1$. Then*

(a)
$$H(|f|^2) + H(|\hat{f}|^2) \geq log(\frac{e}{2}).$$

(Notice that since $\frac{e}{2} > 1$, $log(\frac{e}{2}) > 0$.)

(b) *The equality holds in (a) if and only if f is a constant multiple of a translation and a modulation of a Gaussian e^{-ax^2}, $a > 0$.*

As an application of his L^p, L^q estimates for the Fourier transform, Beckner proved part (a) of Hirschman's conjecture.

Theorem 1.10, (Beckner, 1975, [B1], [B2]). *The part (a) of Hirschman's Conjecture is true.*

Theorem 1.11, (Özaydın-Przebinda, 2000, [Ö-P 2]). *The part (b) of Hirschman's Conjecture is true.*

Let A be a finite Abelian group and let α be a Haar measure on A. Thus α is a positive constant multiple of the counting measure on A. Let $\hat{\alpha}$ be the dual Haar measure on the dual group $\hat{A}$, in the sense that for a function $f : A \to \mathbb{C}$, the Fourier transform and the inverse Fourier transform are given by

$$(1.12) \qquad \begin{aligned} \hat{f}(\hat{b}) &= \int_A f(a)\hat{b}(-a) \, d\alpha(a), \\ f(a) &= \int_{\hat{A}} \hat{f}(\hat{b})\hat{b}(a) \, d\hat{\alpha}(\hat{b}). \end{aligned}$$

The following theorem is known. For the idea of a proof, various particular cases and generalizations see [Hi], [M-U], [M] and [D-C-T].

Theorem 1.13. *Let $f \in L^2(A, \alpha)$, with $\| f \|_2 = 1$. Then*

$$H(|f|^2) + H(|\hat{f}|^2) \geq 0.$$

Theorem 1.14, (Özaydın-Przebinda, 2000, [Ö-P 2]). *The equality holds in the above theorem 1.13 if and only if f is a constant multiple of a translation and a modulation of the indicator function of a subgroup of A.*

Corollary 1.15, (DeBrunner-Özaydın-Przebinda, 2001, [P-D-Ö]). *The discretization of the minimizers for the entropy inequality on $\mathbb{R}$ does not give minimizers for the entropy inequality on any finite cyclic group A.*

In other words, it is certain that no discretization of a signal defined on the real line and best concentrated in the time-frequency plane, will lead to a signal defined on a finite cyclic group which is also best concentrated in the corresponding finite time-frequency plane.

In a mathematically natural search for the most general theorem, which would generalize all the cases considered above, we arrive at the notion of a locally compact

Abelian group, (see [H-R], [J-W]). The integration and the notion of the Fourier transform are both well established for such groups.

Let A be a locally compact Abelian group. As was explained to the author by Michael Cowling, a result of Ahern and Jeweet [A-J], together with [H-R, 9.8], imply that A is isomorphic to the direct product of a finite number of copies of $\mathbb{R}$ and an Abelian locally compact group B, which contains an open compact subgroup:

$$(1.16) \qquad\qquad A = \mathbb{R}^n \times B.$$

Let $\hat{A}$ be the Pontryagin dual of A. Then $\hat{A} = R^n \times \hat{B}$, where $\hat{B}$ also contains an open compact subgroup. Let α be a Haar measure on A and let $\hat{\alpha}$ be the Haar measure on $\hat{A}$, dual to α, so that the Fourier transform and the inverse Fourier transform are given by the following formulas (1.12). Let $U(L^2(A, \alpha))$ be the group of unitary (norm preserving) operators on the Hilbert space $L^2(A, \alpha)$, and let $G \subseteq U(L^2(A, \alpha))$ be the group generated by all the translations, all modulations and by the multiplications by complex numbers of absolute value 1. This is the Heisenberg group attached to the Abelian group A.

Theorem 1.17, (Özaydın-Przebinda, 2000, [Ö-P 2]). *For any function $f \in L^2(A, \alpha)$, with $\| f \|_2 = 1$, such that*

$$(*) \qquad\qquad \| f \|_1 < \infty \quad and \quad \| \hat{f} \|_1 < \infty,$$

the following inequality holds

$$(a) \qquad\qquad H(|f|^2) + H(|\hat{f}|^2) \geq n \, log(\frac{e}{2}).$$

The set of functions for which equality occurs in (a) coincides with the union of orbits

$$(b) \qquad\qquad G \cdot f,$$

where $f = g \otimes h$, g is a Gaussian on $\mathbb{R}^n$, and h an appropriate constant multiple of the indicator function of a (open-compact) subgroup of B.

2. The support inequality for a finite Abelian group.

In this section A stands for a finite Abelian group and α for the Haar measure given by $\alpha(a) = \frac{1}{\sqrt{|A|}}$, $a \in A$. Denote the translations and modulations by

$$
\begin{aligned}
T_c &: f(a) \to f(a + c), \\
M_{\hat{b}} &: f(a) \to \hat{b}(a)f(a).
\end{aligned}
\tag{2.1}
$$

Notice that

$$
T_b M_{\hat{b}} T_c M_{\hat{c}} = z M_{\hat{b}\hat{c}} T_{b+c} \qquad (b, c \in A, \ \hat{b}, \hat{c} \in \hat{A}),
\tag{2.2}
$$

where z is a complex number of absolute value 1. Indeed, The action of the operator on the left hand side on a function $f(a)$ may be represented as

$$
f(a) \underset{M_{\hat{c}}}{\to} \hat{c}(a)f(a) \underset{T_c}{\to} \hat{c}(a+c)f(a+c) \underset{M_{\hat{b}}}{\to} \hat{b}(a)\hat{c}(a+c)f(a+c)
$$

$$
\underset{T_b}{\to} \hat{b}(a+b)\hat{c}(a+b+c)f(a+b+c) = (\hat{b}(b)\hat{c}(b+c))(\hat{b}(a)\hat{c}(a))f(a+(b+c)).
$$

In particular the Heisenberg group attached to A may written as

$$
G = \{z M_{\hat{b}} T_c; \ \hat{b} \in \hat{A}, \ c \in A\}.
\tag{2.3}
$$

It is esthetically pleasing, and convenient for applications, to know that an orthonormal basis of a space of functions may be indexed by a group. This is the case for an orthonormal wavelet basis for functions defined on the real line, [D], [Ö-P 1]. The lemma below states that such bases may be chosen from among our minimizers on a finite Abelian group. In contrast, the translations and/or modulations of Gaussians defined on the reals, will not be orthogonal.

Lemma 2.4, (see [D-S], [P-D-Ö] for the cyclic case). *Let B be a subgroup of A and let $\mathbb{I}_B$ denote the indicator function of B. Let $\beta \in \mathbb{C}$ be such tat $\| \beta \mathbb{I}_B \|_2 = 1$. Then, up to constant multiples of absolute value 1, the orbit $G \cdot \beta \mathbb{I}_B$ is an orthonormal basis of the Hilbert space $L^2(A, \alpha)$.*

Proof. For $c \in A$ and $\hat{c} \in \hat{A}$ we have

$$
M_{\hat{c}} T_c \beta \mathbb{I}_B(a) = \hat{c}(a)\beta \mathbb{I}_B(a + c) = \hat{c}(a)\beta \mathbb{I}_{B-c}(a).
$$

Hence, in order to parameterize the orbit, we may choose the $\hat{c}$ modulo $B^\perp (= \{\hat{a} \in \hat{A};\ \hat{a}|_B = 1\})$, and the c modulo B. Thus the cardinality of the orbit, modulo the multiplications by complex numbers of absolute value 1, is equal to

$$|\hat{A}/B^\perp| \cdot |A/B| = |B| \cdot |A/B| = |A| = dim\, L^2(A, \alpha).$$

Let $d \in A$ and let $\hat{d} \in \hat{A}$. Then,

$$\int_A M_{\hat{c}} T_c \beta \mathbb{I}_B(a) \cdot \overline{M_{\hat{d}} T_d \beta \mathbb{I}_B(a)}\, d\alpha(a) = |\beta|^2 \int_A \hat{c}(a) \overline{\hat{d}(a)} \mathbb{I}_{B-c}(a) \mathbb{I}_{B-d}(a)\, d\alpha(a).$$

For this quantity to be non-zero, we must have $B - c = B - d$. Then it is equal to

$$|\beta|^2 \int_{B-c} \hat{c}(a) \overline{\hat{d}(a)}\, d\alpha(a) = |\beta|^2 \int_B \hat{c}(a+c) \overline{\hat{d}(a+c)}\, d\alpha(a)$$

$$= |\beta|^2 \int_B \hat{c}(a) \hat{d}(a)^{-1}\, d\alpha(a) \cdot \hat{c}(c) \hat{d}(c)^{-1},$$

which is non-zero if and only if $\hat{c}\hat{d}^{-1} \in B^\perp$. Then, modulo multiplication by the constant $\hat{c}(c)\hat{d}(c)^{-1}$, the result is

$$|\beta|^2 \int_B d\alpha(a)\, \|\, \beta \mathbb{I}_B\, \|_2^2 = 1.$$

$\square$

Corollary 2.5, (see [P-D-Ö] for the cyclic case). *In terms of 2.4, the orbit* $G \cdot \beta \mathbb{I}_B$ *consists of functions such that*

(a) $$H(|f|^2) = H(|\hat{f}|^2)$$

or equivalently, such that

(b) $$|supp\, f| = |supp\, \hat{f}|,$$

if and only if

(c) $$|B| = \sqrt{|A|}.$$

Proof. Since the entropy and the cardinality of the support for f and for $\hat{f}$ are invariant under the action of the Heisenberg group G, it suffices to consider $f = \beta \mathbb{I}_B$. Then, by a straightforward computation,

$$\hat{f} = \alpha(B)\beta \mathbb{I}_{B^\perp}.$$

Hence, $|supp\,\hat{f}| = |B^{\perp}| = |A|/|B|$, and since $|supp\,f| = |B|$, the equivalence of (b) and (c) follows.

Furthermore

$$H(|\beta\mathbb{I}_B|^2) = -\int_A |\beta|^2\mathbb{I}_B(a)^2\,log(|\beta|^2\mathbb{I}_B(a)^2)\,d\alpha(a) = -|\beta|^2 log(|\beta|^2)\,\alpha(B),$$

and similarly,

$$H(|\beta\alpha(B)\mathbb{I}_{B^{\perp}}|^2) = -|\beta|^2|\alpha(B)|^2 log(|\beta|^2|\alpha(B)|^2)\,\hat{\alpha}(B^{\perp}).$$

Thus (a) is equivalent to

$$log(|\beta|^2) = log(|\beta|^2|\alpha(B)|^2)\alpha(B)\hat{\alpha}(B^{\perp}).$$

Since, by a simple computation,

$$\alpha(B)\hat{\alpha}(B^{\perp}) = 1,$$

we see that (a) is equivalent to $\alpha(B)^2 = 1$. But, by our assumption on α, $\alpha(B) = |B|/\sqrt{|A|}$, so (a) is equivalent to (c). $\square$

For reader's convenience we reproduce here a lemma which is the key to the proof of theorem 1.5, leaving the complete proof of the theorem as an exercise. (See [P-D-Ö] for an exposition.)

Lemma 2.6, [D-S]. *Let $f : A \to \mathbb{C}$ be a non-zero function. Then under the identification $\hat{A} = A$, $\hat{a}(b) = e^{\frac{2\pi i}{|A|}ab}$, the function $\hat{f}$ cannot have more than $|supp\,f|$-consecutive zeros.*

Proof. Let $supp\,f = \{a_1, a_2, a_3, ..., a_m\}$. Let $\omega = e^{-\frac{2\pi i}{|A|}}$. Then

$$\begin{bmatrix} \hat{f}(0) \\ \hat{f}(1) \\ . \\ . \\ . \\ \hat{f}(m-1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & . & . & 1 \\ \omega^{a_1} & \omega^{a_2} & . & . & \omega^{a_m} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \omega^{a_1(m-1)} & \omega^{a_2(m-1)} & . & . & \omega^{a_m(m-1)} \end{bmatrix} \cdot \begin{bmatrix} \hat{f}(a_1) \\ \hat{f}(a_2) \\ . \\ . \\ . \\ \hat{f}(a_m) \end{bmatrix}.$$

By Vandermonde, the above matrix is invertible. Thus the vector on the left hand side is non-zero. Since a translation of $\hat{f}$ is equivalent to a modulation of f, which does not change the support of f, the lemma follows. $\square$

In the remainder of this section we provide an elementary proof of the part of theorem 1.6, which describes the minimizers for the support inequality (which we assume), if A is a finite direct product of groups of order 2: $A = (\mathbb{Z}/2\mathbb{Z})^N$. (See theorem 2.14 below.) This proof is a simplified version of the argumant developed in [M-Ö-P].

Lemma 2.7. *Let B and C be subgroups of A, such that*

$$A = B \oplus C.$$

For a function $f : A \to \mathbb{C}$ let

$$f_c(b) = f(b + c) \qquad (b \in B, \ c \in C).$$

Suppose

$$|supp \, f| \cdot |supp \, \hat{f}| = |A| \quad and \quad f_c \neq 0 \quad for \ all \quad c \in C.$$

Then

(a) $$|supp \, f_c| \cdot |supp \, \hat{f}_c| = |B| \qquad (c \in C),$$

(b) $$|supp \, f_c| = \frac{1}{|C|} |supp \, f| \qquad (c \in C),$$

(c) $$|supp \, \hat{f}_c| = |supp \, \hat{f}| \qquad (c \in C),$$

(d) $$supp \, \hat{f}_c = P(supp \, \hat{f}) \qquad (c \in C),$$

where $P : \hat{A} = \hat{B} \oplus \hat{C} \ni \hat{b} + \hat{c} \to \hat{b} \in \hat{B}$.

Proof. Clearly,

(2.8) $$supp \, f = \bigcup_{c \in C} ((supp \, f_c) + c).$$

Hence,

(2.9) $$|supp \, f| = \sum_{c \in C} |supp \, f_c|.$$

By the Fourier inversion formula on C we have

(2.10) $$\hat{f}_c(\hat{b}) = \frac{1}{\sqrt{|C|}} \sum_{\hat{c} \in \hat{C}} \hat{f}(\hat{b} + \hat{c})\hat{c}(c) \qquad (\hat{b} \in \hat{B}, \ c \in C).$$

Hence,

(2.11) $$supp \, \hat{f}_c \subseteq P(supp \, \hat{f}) \qquad (c \in C),$$

and therefore

$$(2.12) \qquad |supp\ \hat{f}_c| \leq |supp\ \hat{f}| \qquad (c \in C).$$

Suppose $c \in C$ is such that

$$|C| \cdot |supp\ f_c| < |supp\ f|.$$

Then, by (2.12),

$$|C| \cdot |supp\ f_c| \cdot |supp\ \hat{f}_c| < |supp\ f| \cdot |supp\ \hat{f}| = |A|.$$

Hence,

$$|supp\ f_c| \cdot |supp\ \hat{f}_c| < |B|,$$

which contradicts the fact that $f_c \neq 0$. (Here we use the assumption that we have the support inequality for the finite cyclic groups of the form $(\mathbb{Z}/2\mathbb{Z})^N$.) Thus

$$(2.13) \qquad |C| \cdot |supp\ f_c| \geq |supp\ f| \qquad (c \in C).$$

Clearly, (2.9) and (2.13) imply (b). Further (b) and (2.12) imply that for all $c \in C$,

$$|supp\ f_c| \cdot |supp\ \hat{f}_c| = \frac{1}{|C|}|supp\ f| \cdot |supp\ \hat{f}_c| \leq \frac{1}{|C|}|supp\ f| \cdot |supp\ \hat{f}| = |A|/|C| = |B|.$$

Thus (a) follows. Part (c) follows from (a) and (b). Part (d) follows from (2.11) and (c). $\square$

Theorem 2.14. *Let $A = (\mathbb{Z}/2\mathbb{Z})^N$. Suppose $f : A \to \mathbb{C}$ is a minimizer, i.e.*

$$|supp\ f| \cdot |supp\ \hat{f}| = |A|.$$

Then, up to a constant multiple, a modulation and a translation, f is the indicator function of a subgroup of A.

Proof. We proceed via the induction on N. It is easy to check that the statement holds for $N = 2$.

Let $f : A \to \mathbb{C}$ be a minimizer. Suppose there is a proper subgroup $B \subseteq A$, such that $supp\ f \subseteq B$. Then there is a subgroup $C \subseteq A$ such that $A = B \oplus C$. (In order to have some standard linear algebra at the disposal, it might be easier here

to view A a vector space of dimension N over the field $\mathbb{Z}/2\mathbb{Z}$ of two elements.) In these terms

$$f = (f|_B) \otimes \delta,$$

where δ is the Dirac delta at zero on C. Hence

$$\hat{f} = \widehat{(f|_B)} \otimes \hat{\delta},$$

where $\hat{\delta}$ is a constant multiple of the function $\mathbb{I}_C$. We see from the above two equations that $f|_B$ is a minimizer on B. Thus, by the inductive assumption $f|_B$, and hence f has the desired form.

Suppose there is a proper subgroup $B \subseteq A$ and an element $c \in A$ such that $supp\, f \subseteq B + c$. Let $g(a) = f(a + c)$, $a \in A$. Then g is a minimizer on A and $supp\, g \subseteq B$. Hence, by the previous argument, g and hence f has the desired form.

Let $A = B \oplus C$ as in (2.7), with $|C| = 2$. By the previous two cases we may assume that

$$f_c \neq 0 \qquad (c \in C).$$

By (2.7.a), each f_c is a minimizer on B. Therefore, by the inductive assumption, for each $c \in C$, $supp\, \hat{f}_c$ is a coset of a subgroup of $\hat{B}$. By (2.7.d) these cossets do not depend on $c \in C$. Thus there is a subgroup $D \subseteq \hat{B}$ and an element $\hat{b} \in \hat{B}$ such that

$$sup\, \hat{f}_c = D + \hat{b}_0 \qquad (c \in C).$$

Replacing f by an appropriate translation of f we may assume that each $\hat{f}_c$ is a constant on its support. Thus there is a function $h : C \to \mathbb{C}$ such that

$$\hat{f}_c = \mathbb{I}_{D+\hat{b}_0} \otimes h(c) \qquad (c \in C).$$

Therefore

(2.15) $$\hat{f} = \mathbb{I}_{D+\hat{b}_0} \otimes \hat{h}.$$

We see from (2.15) that h is a minimizer for C. But

$$f = \hat{\mathbb{I}}_{D+\hat{b}_0} \otimes h$$

and our assumption ($f_c \neq 0$ for all $c \in C$) implies that $supp\, h = C$. Hence, $|supp\, \hat{h}| = 1$. Thus $supp\, \hat{h} = \{\hat{c}_0\}$, for some $\hat{c}_0 \in \hat{C}$. Therefore, by (2.15),

$$supp\, \hat{f} = (D + \hat{b}_0) + \hat{c}_0 = D + (\hat{b}_0 + \hat{c}_0).$$

Thus $supp\, \hat{f}$ is a coset of a proper subgroup of $\hat{A}$. Since $\hat{f}$ is a minimizer, our previous argument implies that $\hat{f}$, and hence f, has the desired property. $\square$

3. A few words on the notion of the entropy. In this section we recall a few basic facts, which indicate that the notion of the entropy is quite natural and useful. For more details we refer the reader to [C-T] and [S].

Let X be a discrete random variable with the probability distribution function $p(x) = P(X = x)$. Set

(3.1)
$$H(X) = H(p) = -\sum_x p(x)log_2(p(x)).$$

Example 3.2 Suppose X has a uniform distribution over $2^3 = 8$ outcomes. Then

$$H(X) = -\sum_{x=1}^{8} \frac{1}{8}log_2(\frac{1}{8}) = 3.$$

This agrees with the number of bits needed to describe X in binary:

$$c(0) = 000, \ c(1) = 001, \ c(2) = 010, \ c(3) = 011, \ c(4) = 100,$$
$$c(5) = 101, \ c(6) = 110, \ c(7) = 111.$$

Example 3.3 Consider a horse race, with eight horses taking part. Suppose the probability of winning the race is distributed as follows:

$$p(0) = 1/2, \ p(1) = 1/4, \ p(2) = 1/8, \ p(3) = 1/16, \ p(4) = p(5) = p(6) = p(7) = 1/64.$$

Then $H(X) = 2$. We may encode the horses in binary as follows:

$$c(0) = 0, \ c(1) = 10, \ c(2) = 110, \ c(3) = 1110, \ c(4) = 111100,$$
$$c(5) = 111101, \ c(6) = 111110, \ c(7) = 111111.$$

Let $l(c(x))$ be the length of the code word $c(x)$. Then

$$l(c(0)) = 1, \ l(c(1)) = 2, \ l(c(2)) = 3, \ l(c(3)) = 4, \ l(c(5)) = ... = l(c(7)) = 6.$$

Hence, the expected value of $l(c(X))$ is

$$E(l(c(X))) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + 6 \cdot \frac{1}{64} = 2.$$

Thus

$$H(X) = E(l(c(X))).$$

This last equality is not a coincidence, as we shall explain below. For details see [C-T].

A source code for a discrete random variable X is a mapping c from the range of X to

$$\{0,1\} \cup \{0,1\}^2 \cup \{0,1\}^3 \cup \ldots .$$

If a codeword $c(x)$ belongs to $\{0,1\}^k$, let $l(c(x)) = k$ denote the length of $c(x)$. The code is called instantenous if no codeword is a prefix to any other code word. For such codes one can recognize the separate words $c(x_1)$, $c(x_2)$, ... by looking at the string $c(x_1)c(x_2)\ldots$. For instance in our second example, the string 010110111111 is made of the codewords 0, 10, 110, 111111.

Theorem 3.4, [S]**, (see also** [C-T]**).** *For an instantenous code c we have,*

$$E(l(c(X))) \geq H(X).$$

Moreover, there is an instantenous code c such that

$$E(l(c(X))) < H(X) + 1.$$

A theorem of Shannon, which we quote below, gives a characterization of the entropy as a measure of uncertainty of the outcome of an experiment.

Let $P = \{p_1, p_2, p_3, ..., p_n\}$ be a finite probability sequence, i.e. $p_j \geq 0$, $\sum_j p_j = 1$.

Theorem 3.5, (Shannon, 1948, [S]**).** *Suppose H is a function defined on finite probability sequences such that*

(a) $\qquad\qquad H(p_1, p_2, p_3, ..., p_n) \quad$ *is continuous,*

(b) $\qquad H(\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}) \quad$ *is monotonically increasing as a function of n,*

$\qquad\qquad$ *If $Q_j = \{q_{j_1}, q_{j_2}, q_{j_3}, ...\}$, $j = 1, 2, 3, ..., n$, are probability sequences,*

(c) $\qquad$ *then* $\quad H(\bigcup_{j=1}^{n} p_j Q_j) = H(P) + \sum_{j=1}^{n} p_j H(Q_j).$

Then, up to a constant multiple,

$$H(P) = -\sum_{j=1}^{n} p_j log(p_j).$$

Suppose M is a finite set and μ is a positive multiple of the counting measure on M. For a function $f : M \to \mathbb{C}$ and for $1 \leq p < \infty$ we have the L^p norm of f defined by

$$\| f \|_p = \left(\int_M |f(x)|^p \, d\mu(x) \right)^{1/p}.$$

There is a simple and explicit connection between the entropy (see 1.7) and the L^p norms, expressed in the lemma below, which may be verified by a straightforward computation (left to the reader). See [Wi] for more explanation.

Lemma 3.6. *Let* $p = p(t) = \frac{1}{t}$, $0 < t \leq 1$. *Then for any function* $f : M \to \mathbb{C}$,

(a)
$$\frac{d}{dt} log(\| f \|_p) = - \int_M \frac{|f(x)|^p}{\| f \|_p^p} log(\frac{|f(x)|^p}{\| f \|_p^p}) \, d\mu(x).$$

In particular, for f *with* $\| f \|_1 = 1$,

(b)
$$\frac{d}{dt} log(\| f \|_p)|_{p=1} = H(|f|),$$

and for f *with* $\| f \|_2 = 1$,

(b)
$$\frac{d}{dt} log(\| f \|_p)|_{p=2} = H(|f|^2),$$

4. The entropy inequality for a finite Abelian group. In this section we give a proof of theorem 1.14 for a finite Abelian group A. For the general case we refer the reader to [Ö-P 2].

The proof is based on the following four basic theorems.

Hölder's Inequality, [H2].

$$\int_M |\phi(x)\psi(x)| \, dx \leq \| \phi \|_p \| \psi \|_q, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Plancherel's Formula, [H-R].

$$\| \hat{f} \|_2 = \| f \|_2.$$

Riesz-Thorin-Young Inequality, [Z, Chapter 9, (1.11)].

$$\| \hat{f} \|_{1/(1-t)} \leq \| f \|_{1/t} \qquad (\frac{1}{2} \leq t \leq 1).$$

Hopf's Maximum Principle, [H2, Theorem 3.1.6']. *Let $D \subseteq \mathbb{C}$ be an open unit disc, and let $u : D \rightarrow \mathbb{R}$ be a harmonic function which extends to a continuous function on the closure $\overline{D}$ of D, $u : \overline{D} \rightarrow \mathbb{R}$. Suppose z is a point on the boundary of D such that $u(z) \geq u(z')$ for all $z' \in \overline{D}$, and the directional derivative of u at z along the radius which ends at z, is zero. Then*

$$u(z) = u(z') \qquad (z' \in \overline{D}).$$

Let $f : A \rightarrow \mathbb{C}$ be a minimizer. Consider the following function

$$(4.1) \quad F(z) = \int_{\hat{A}} (|f|^{2z} \frac{f}{|f|}) \hat{}(\hat{a}) \, |\hat{f}(\hat{a})|^{2z} \frac{\overline{\hat{f}(\hat{a})}}{|\hat{f}(\hat{a})|} \, d\hat{\alpha}(\hat{a}) \qquad (z \in \mathbb{C}, \; \frac{1}{2} \leq Re(z) \leq 1),$$

where $\frac{f}{|f|} = 0$ outside the support of f, and similarly for $\frac{\hat{f}}{|\hat{f}|}$. A straightforward application of Hölder's inequality and the Riesz-Thorin Theorem shows that for $\frac{1}{2} \leq x \leq 1$, $y \in \mathbb{R}$, $p = \frac{1}{x}$ and q defined by the equation $\frac{1}{p} + \frac{1}{q} = 1$ (with $q = \infty$ if $p = 1$), we have

(4.2)

$$|F(x+iy)| \leq \| \, (|f|^{2x+i2y} \frac{f}{|f|}) \hat{} \, \|_q \cdot \| \, |\hat{f}|^{2x+i2y} \, \|_p \leq \| \, |f|^{2x+i2y} \, \|_p \cdot \| \, |\hat{f}|^{2x+i2y} \, \|_p$$

$$= \| \, f \, \|_2 \cdot \| \, \hat{f} \, \|_2 = 1.$$

The function F is analytic in the open strip (4.1) and continuous in the closed strip. A straightforward calculation shows that

$$F'(z) = \int_{\hat{A}} (|f|^{2z} \frac{f}{|f|} log(|f|^2)) \hat{}(\hat{a}) \, |\hat{f}(\hat{a})|^{2z} \frac{\overline{\hat{f}(\hat{a})}}{|\hat{f}(\hat{a})|} \, d\hat{\alpha}(\hat{a})$$

$$+ \int_{\hat{A}} (|f|^{2z} \frac{f}{|f|}) \hat{}(\hat{a}) \, |\hat{f}(\hat{a})|^{2z} \frac{\overline{\hat{f}(\hat{a})}}{|\hat{f}(\hat{a})|} log(|\hat{f}(\hat{a})|^2) \, d\hat{\alpha}(\hat{a}).$$

Hence, by the Plancherel formula,

$$(4.3) \qquad\qquad F'(\frac{1}{2}) = -H(|f|^2) - H(|\hat{f}|^2).$$

Since f is a minimizer, the right hand side of the equation (4.3) is zero. In particular $Re \, F(z)$ is a real valued harmonic function, on the interior of the disc of radius $\frac{1}{4}$ centered at $z = \frac{3}{4}$, which achieves the maximum at $z = \frac{1}{2}$ and has derivative equal to

zero at this point. Hence the Hopf's Maximum Principle implies that $Re\,F(z) = 1$ on the disc. Hence, $F(z) = 1$ on the disc. In particular,

$$(4.4) \qquad 1 = F(1) = \int_{\hat{A}} (|f|\hat{f})(\hat{a})|\hat{f}(\hat{a})|\overline{\hat{f}(\hat{a})}\,d\hat{\alpha}(\hat{a}).$$

The formula (4.4) may be rewritten as

$$(4.5) \qquad 1 = \int_{\hat{A}} \int_{A} |f(a)|^2|\hat{f}(\hat{a})|^2\hat{a}(-a)\frac{f(a)}{|f(a)|}\frac{\overline{\hat{f}(\hat{a})}}{|\hat{f}(\hat{a})|}\,d\alpha(a)\,d\hat{\alpha}(\hat{a}).$$

Since,

$$1 = \int_{\hat{A}} \int_{A} |f(a)|^2|\hat{f}(\hat{a})|^2\,d\alpha(a)\,d\hat{\alpha}(\hat{a})$$

the equation (4.5) implies that

$$(4.6) \qquad 1 = \hat{a}(-a)\frac{f(a)}{|f(a)|}\frac{\overline{\hat{f}(\hat{a})}}{|\hat{f}(\hat{a})|} \qquad (a \in supp\,f, \hat{a} \in supp\,\hat{f}).$$

Hence,

$$\hat{a}(-a) = \frac{\overline{f(a)}}{|f(a)|}\frac{\hat{f}(\hat{a})}{|\hat{f}(\hat{a})|}$$

Thus for $\hat{a} \in supp\,\hat{f}$

$$\hat{f}(\hat{a}) = \int_{A} f(a)\hat{a}(-a)\,d\alpha(a) = \int_{A} f(a)\frac{\overline{f(a)}}{|f(a)|}\,d\alpha(a)\frac{\hat{f}(\hat{a})}{|\hat{f}(\hat{a})|} = \| f \|_1 \frac{\hat{f}(\hat{a})}{|\hat{f}(\hat{a})|}.$$

Therefore

$$(4.7) \qquad |\hat{f}(\hat{a})| = \| f \|_1 \qquad (\hat{a} \in supp\,\hat{f}),$$

and similarly

$$(4.8) \qquad |f(a)| = \| \hat{f} \|_1 \qquad (a \in supp\,f).$$

The statement (4.7) implies that the function $|\hat{f}|$ is constant on its support. Since $\| \hat{f} \|_2 = 1$, the constant is equal to $\hat{\alpha}(supp\,\hat{f})^{-1/2}$. Hence,

$$H(|\hat{f}|^2) = log(\hat{\alpha}(supp\,\hat{f})).$$

Similarly

$$H(|f|^2) = log(\alpha(supp\,f)).$$

Since f is a minimizer,

$$log(\alpha(supp\, f) \cdot \hat{\alpha}(supp\, \hat{f})) = 2\left(H(|f|^2) + H(|\hat{f}|^2)\right) = 0.$$

Therefore

(4.9) $$\alpha(supp\, f) \cdot \hat{\alpha}(supp\, \hat{f}) = 1.$$

We may assume that $0 \in supp\, \hat{f}$ and $0 \in supp\, f$. Then (4.7) implies

$$\left|\int_A f(a)\, d\alpha(a)\right| = \int_A |f(a)|\, d\alpha(a).$$

Therefore there is $\lambda \in \mathbb{C}$ such that $f = \lambda|f|$. Hence (4.7) may be rewritten as

(4.10) $$\left|\int_A \lambda|f(a)|\hat{a}(-a)\, d\alpha(a)\right| = \int_A |f(a)|\, d\alpha(a) \quad (\hat{a} \in supp\, \hat{f}).$$

Therefore

(4.11) $$supp\, \hat{f} \subseteq (-supp\, f)^\perp,$$

where for a subset $S \subseteq A$, $S^\perp = \{\hat{a} \in \hat{A};\ \hat{a}|_S = 1\}$. Similarly (4.8) implies

(4.12) $$supp\, f \subseteq (supp\, \hat{f})^\perp.$$

By dualizing (4.11) and (4.12) we deduce

(4.13)
$$-supp\, f \subseteq (-supp\, f)^{\perp\perp} \subseteq (supp\, \hat{f})^\perp, \text{ and}$$
$$supp\, \hat{f} \subseteq (supp\, \hat{f})^{\perp\perp} \subseteq (supp\, f)^\perp$$

But, as is well known and easy to check,

(4.14) $$\hat{\alpha}((supp\, \hat{f})^\perp) \cdot \alpha((supp\, \hat{f})^{\perp\perp}) = 1.$$

By combining (4.9), (4.13) and (4.14) we see that the inclusions (4.13) are equalities. In particular $supp\, f$ is a subgroup of A and f is invariant under the translations by this subgroup. Thus f is a constant multiple of the indicator function of a subgroup of A, as claimed.

References

[A-J]   P. Ahern and R. Jewett, *Factorization of locally compact Abelian groups*, Illinois J. Math. **9** (1965), 230–235.

[B1]   W. Beckner, *Inequalities in Fourier Analysis*, Annals of Mathematics **102 Number 6** (1975), 159–182.

[B2]   W. Beckner, *Pitt's Inequality and the Uncertainty Principle*, Proceedings of the AMS **123** (1995), 1897–1905.

[C-T]   T. Cover and J. Thomas, *Elements of information theory*, Elements of information theory, 1991.

[D-C-T]   A. Dembo, T. M. Cover and J. A. Thomas, *Information Theoretic Inequalities*, IEEE Transactions on Information Theory **37** (1991), 1501–21518.

[D-S]   D. L. Donoho and P. B. Stark, *Uncertainty Principles and Signal Recovery*, SIAM Journal of Applied Mathematics **49** (1989), 906–931.

[H-R]   E. Hewitt and K. A. Ross, *Abstract Harmonic Analysis*, Springer Verlag, 1963.

[Hi]   I. I. Hirschman, Jr., *A Note on Entropy*, Amer. J. Math. **79** (1957), 152–156.

[H1]   L. Hörmander, *The Analysis of Linear Partial Differential Operators, I*, Springer Verlag, 1983.

[H2]   L. Hörmander, *Notions of Convexity*, Birkhäuser, 1994.

[M-Ö-P]   E. Matusiak, M. Özaydın and T. Przebinda, *The Donoho - Stark Uncertainty Principle for a Finite Abelian Group*, preprint.

[M]   H. Maassen, *A discrete entropic uncertainty relation*, Quantum Probability and Applications **5** (1988), 263– 266.

[M-U]   H. Maassen and J. Uffink, *Generalized Entropic Uncertainty Relations*, Phys. Rev. Lett. **60** (1988), 1103– 1106.

[Ö-P 1]   M. Özaydın and T. Przebinda, *Platonic Orthonormal Wavelets*, Appeared Applied and Computational Harmonic Analysis **4** (1997), 351–365.

[Ö-P 2]   M. Özaydın and T. Przebinda, *An Entropy-based Uncertainty Principle for a Locally Compact Abelian Group*, preprint available at www.crystal.ou.edu.

[P-D-Ö]   T. Przebinda, V. DeBrunner and M.Özaydın, *The Optimal Transform for the Discrete Hirschman Uncertainty Principle*, IEEE Transactions on Information Theory **47** (2001), 2086–2090.

[S]   C. E. Shannon, *A Mathematical Theory of Communication*, The Bell System Technical Journal **27** (1948), 379–656.

[Sm]   K. T. Smith, *The Uncertainty Principle on Groups*, SIAM Journal of Applied Mathematics **50** (1990), 876–882.

[Wi]   M. Wickerhauser, *Adapted wavelet analysis from theory to software*, Cambridge University Press, 1990.

[Z]   A. Zygmund, *Trigonometric series, second edition, volumes I and II combined*, A K Peters, Ltd., Wellesley, MA, 1994.