The Uncertainty Principle: A Mathematical Survey

Gerald B. Folland and Alladi Sitaram

ABSTRACT. *We survey various mathematical aspects of the uncertainty principle, including Heisenberg's inequality and its variants, local uncertainty inequalities, logarithmic uncertainty inequalities, results relating to Wigner distributions, qualitative uncertainty principles, theorems on approximate concentration, and decompositions of phase space.*

Introduction

The uncertainty principle is partly a description of a characteristic feature of quantum mechanical systems, partly a statement about the limitations of one's ability to perform measurements on a system without disturbing it, and partly a meta-theorem in harmonic analysis that can be summed up as follows.

A nonzero function and its Fourier transform cannot both be sharply localized. (0.1)

When translated into the language of quantum mechanics, as we shall do in §2, (0.1) says that the values of a pair of canonically conjugate observables such as position and momentum cannot both be precisely determined in any quantum state. Therefore, it leads to mathematical formulations of the physical ideas first developed in Heisenberg's seminal paper [51] of 1927 and widely promulgated thereafter.

However, the uncertainty principle also has a useful interpretation in classical physics. Namely, if $f(t)$ represents the amplitude of a signal (a sound wave or light wave, perhaps) at time t, the Fourier transform $\widehat{f}$ tells how f is built from sine waves of different frequencies and (0.1) expresses a limitation on the extent to which a signal can be both time-limited and band-limited. This aspect of the uncertainty principle was already expounded by Norbert Wiener in a lecture in Göttingen in 1925. Unfortunately, no written record of this lecture seems to have survived, apart from the nontechnical account in Wiener's autobiography [119, pp. 105–107], so one can only guess at what precise versions of (0.1) it might have contained. Whatever influence this lecture might have had on the physicists in the audience, however, the uncertainty principle did not really sink into the minds of signal analysts until Gabor's fundamental work [40] in 1946. Since then, it has become firmly embedded in the common culture.

On the mathematical side, there were sporadic developments relating to the uncertainty principle in the fifty years after the initial work in the 1920's, followed by a steady stream of results in the last two decades. The purpose of this paper is to give an overview of this work. We shall have nothing to say about the purely physical or epistemological aspects of the uncertainty principle or the applications of the mathematics to particular problems in physics or engineering, and our references to the mathematical physics literature are less than comprehensive. Moreover, the uncertainty principle impinges directly on some other areas of analysis with a large literature, notably, (i) the study of the properties a function implied by restrictions on the support or the decay properties of its Fourier transform, (ii) the construction of orthonormal bases or frames for L^2 whose elements and their Fourier transforms are well localized (wavelets, etc.), and (iii) the body of analytic results relating to signal analysis and communication theory. To do justice to the ramifications of the uncertainty principle in any of these subjects would require a book by itself. Fortunately, such books have already been written—notably Havin and Jöricke [48] and Daubechies [28]—as well as a number of good expository articles such as those by Benedetto [10] and Benedetto, Heil, and Walnut [13] and the collections in Price [92] and Benedetto and Frazier [12]. On these matters, therefore, we shall be brief.

To begin, let us fix some notation and terminology. The reader may wish to proceed to §1 and refer back as necessary. Dym and McKean [32] is a good reference for the relevant background on Fourier analysis.

$\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Z}$ have their usual meanings, and $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. The Lebesgue measure of a set $E \subset \mathbb{R}^n$ is denoted by $|E|$, and the charcateristic function of E is denoted by χ_E. Inner products in any Hilbert space are denoted by $\langle \cdot, \cdot \rangle$.

Suppose μ is a probability measure on $\mathbb{R}$. The *variance* of μ is

$$V(\mu) = \inf_{a \in \mathbb{R}} \int (x-a)^2 \, d\mu(x).$$

If the integral on the right is finite for one value of a, then it is finite for all a, in which case it is a quadratic function of a whose minimum is achieved when a is the *mean* of μ:

$$M(\mu) = \int x \, d\mu(x).$$

Similarly, if μ is a measure on $\mathbb{R}^n$, we say that μ has *finite variance* if $\int |x|^2 \, d\mu(x) < \infty$. In this case we define the mean $M(\mu) \in \mathbb{R}^n$ and the *covariance matrix* $V(\mu) = (V_{jk}(\mu))$ by

$$M(\mu) = \int x \, d\mu(x), \qquad V_{jk}(\mu) = \int y_j y_k \, d\mu(x) \quad (y = x - M(\mu)).$$

If $d\mu(x) = \rho(x)\,dx$ with $\rho \in L^1(\mathbb{R}^n)$, we shall call ρ a *probability density function* and write $M(\rho)$ and $V(\rho)$ instead of $M(\mu)$ and $V(\mu)$.

We shall define the Fourier transform on $(L^1 + L^2)(\mathbb{R}^n)$ by

$$\widehat{f}(\xi) = \mathcal{F}f(\xi) = \int e^{-2\pi i \xi \cdot x} f(x) \, dx.$$

Then the inversion theorem and Parseval formula take the form $\mathcal{F}^{-1}f(y) = \mathcal{F}f(-y)$ and $\|\widehat{f}\|_2 = \|f\|_2$. In particular, if $f \in L^2(\mathbb{R}^n)$ and $\|f\|_2 = 1$, then $|f|^2$ and $|\widehat{f}|^2$ are both probability density functions on $\mathbb{R}^n$.

In this connection the following observation is useful. For $a, b \in \mathbb{R}^n$, let us define

$$f_{a,b}(x) = e^{2\pi i b \cdot x} f(x-a). \tag{0.2}$$

Then

$$(f_{a,b})\widehat{\ }(\xi) = e^{-2\pi i a \cdot (\xi - b)} \widehat{f}(\xi - b) = e^{2\pi i a \cdot b} (\widehat{f})_{b,-a}(\xi).$$

Thus the map $f \to f_{a,b}$ preserves all L^p norms of f and $\widehat{f}$ while shifting the centers of mass of f and $\widehat{f}$ by a and b, respectively.

We shall allow ourselves the following minor abuse of notation, along with other variations on the same theme: $x_j f$ and $\xi_j \widehat{f}$ denote the functions $x \to x_j f(x)$ and $\xi \to \xi_j \widehat{f}(\xi)$. Thus, for example, $(\partial f/\partial x_j)\widehat{\ } = 2\pi i \xi_j \widehat{f}$.

Some of our discussion will pertain to Fourier analysis on groups other than $\mathbb{R}^n$, so we briefly recall the basic notions. (See Folland [38] for more information.) Suppose G is a locally compact group, equipped with a fixed left Haar measure dx. As with $\mathbb{R}^n$, the Haar measure of $E \subset G$ will be denoted by $|E|$. If π is a unitary representation of G (always assumed strongly continuous) on a Hilbert space $\mathcal{H}_\pi$, its *integrated representation* is the representation of the Banach algebra $L^1(G)$ on $\mathcal{H}_\pi$ defined by

$$\pi(f) = \int f(x)\pi(x)\,dx.$$

If G is a Lie group, we also have the *differentiated representation* of its Lie algebra $\mathfrak{g}$. Namely, for $X \in \mathfrak{g}$, $\pi(X)$ is the skew-adjoint operator on $\mathcal{H}_\pi$ that generates the one-parameter group $\pi(\exp\ tX)$ according to Stone's theorem. That is,

$$\pi(X)u = \frac{d}{dt}\pi(\exp\ tX)u\bigg|_{t=0} = \lim_{t\to 0}\frac{\pi(\exp\ tX)u - u}{t}$$

on the domain of all $u \in \mathcal{H}_\pi$ for which the limit exists. This domain includes the space $\mathcal{H}_\pi^\infty$ of C^∞ *vectors* for π, that is, the set of all $u \in \mathcal{H}_\pi$ such that the map $x \to \pi(x)u$ is C^∞ on G. $\mathcal{H}_\pi^\infty$ is dense in $\mathcal{H}_\pi$, $\pi(X)$ maps $\mathcal{H}_\pi^\infty$ into itself for all $X \in \mathfrak{g}$, and the map $X \to \pi(X)|\mathcal{H}_\pi^\infty$ is a homomorphism of Lie algebras. (See, e.g., Knapp [67, pp. 51–57].)

Suppose G is either (a) Abelian, (b) compact, or (c) unimodular, second countable, and type I; we shall call such groups *Plancherel groups*. Let $\widehat{G}$, the *unitary dual* of G, be a set containing exactly one member of each unitary equivalence class of irreducible unitary representations of G. The *Fourier transform* of $f \in L^1(G)$ is the operator-valued function on $\widehat{G}$ defined by

$$\widehat{f}(\pi) = \pi(f).$$

(The convention for defining $\widehat{f}$ in Folland [38] is slightly different.) There is a canonical topology on $\widehat{G}$ and a unique Borel measure $d\pi$ on $\widehat{G}$, the so-called *Plancherel measure*, such that

$$\|f\|_2^2 = \int_{\widehat{G}} \|\widehat{f}(\pi)\|_{\mathrm{HS}}^2\,d\pi,$$

where $\|\cdot\|_{\mathrm{HS}}$ denotes the Hilbert–Schmidt norm. (Implicit in this statement is the fact that $\widehat{f}(\pi)$ is Hilbert–Schmidt for almost every π.) If G is Abelian, $\widehat{G}$ is identified with the group of continuous homomorphisms from G into $\mathbb{T}$, and Plancherel measure is a Haar measure on $\widehat{G}$.

1. Heisenberg's Inequality

When one asks for a precise quantitative formulation of the principle (0.1), the most common response is the following inequality, usually called *Heisenberg's inequality*. This result does not actually appear in Heisenberg's paper [51], which gives an incisive analysis of the physics of the uncertainty principle but contains little mathematical precision. This omission, however, was soon rectified by Kennard [66] and Weyl [118, Appendix 1] (who credits the result to Pauli).

Theorem 1.1.

If $f \in L^2(\mathbb{R})$ and $\|f\|_2 = 1$, then

$$V(|f|^2)V(|\widehat{f}|^2) \geq \frac{1}{16\pi^2}.$$

In other words, for any $f \in L^2(\mathbb{R})$ and any $a, b \in \mathbb{R}$,

$$\int (x-a)^2|f(x)|^2\,dx \int (\xi-b)^2|\widehat{f}(\xi)|^2\,d\xi \geq \frac{\|f\|_2^4}{16\pi^2}. \tag{1.2}$$

Equality holds in (1.2) if and only if $f(x) = Ce^{2\pi i bx}e^{-\gamma(x-a)^2}$ for some $C \in \mathbb{C}$ and $\gamma > 0$.

Proof. By using the transformation (0.2) we may assume that $a = b = 0$, and clearly we may also assume that the integrals in (1.2) are finite. Since $(f')\widehat{\ }(\xi) = 2\pi i\xi\widehat{f}(\xi)$, the finiteness of $\int|\xi\widehat{f}|^2$ implies that f is absolutely continuous and $f' \in L^2$. The derivative of $|f|^2 = f\overline{f}$ is $2\,\mathrm{Re}\,f\overline{f'}$, so if $-\infty < c < d < \infty$, integration by parts yields

$$2\,\mathrm{Re}\int_c^d xf(x)\overline{f'(x)}\,dx = x|f(x)|^2\Big|_c^d - \int_c^d |f(x)|^2\,dx.$$

Since f, xf, and f' are all in L^2, the integrals in this equality approach finite limits as $c \to -\infty$ or $d \to \infty$ and hence so do $c|f(c)|^2$ and $d|f(d)|^2$. The latter limits must be zero, for otherwise $|f(x)|^2$ would be comparable to x^{-1} for large x and f would not be in L^2. Therefore,

$$\int_{-\infty}^{\infty}|f(x)|^2\,dx = -2\,\mathrm{Re}\int_{-\infty}^{\infty}xf(x)\overline{f'(x)}\,dx. \tag{1.3}$$

Inequality (1.2) now follows from the Schwarz inequality and the Plancherel formula:

$$\|f\|_2^4 \leq 4\int x^2|f(x)|^2\,dx\int|f'(x)|^2\,dx = 16\pi^2\int x^2|f(x)|^2\,dx\int\xi^2|\widehat{f}(\xi)|^2\,d\xi.$$

Equality holds here if and only if f' is a real multiple of xf, say $f'(x) = -2\gamma xf(x)$ with $\gamma \in \mathbb{R}$. This implies that $f(x) = Ce^{-\gamma x^2}$, and of course γ must be positive for f to be in L^2. □

An analogous result (Corollary 2.6) holds for functions on $\mathbb{R}^n$, but the proof is technically harder because the square-integrability of the distribution derivatives $\partial f/\partial x_j$ does not guarantee the continuity of f. One must first work under the assumption that f is smooth and rapidly decaying at infinity and then apply an approximation argument. The (somewhat lengthy) details can be found in Benedetto [10, Appendix A]. We shall present this argument in an abstract setting in §2. (Alternatively, one can reduce to the one-dimensional case by invoking the Stone–von Neumann theorem; see Folland [37, §1.5].)

2. The Uncertainty Inequality in Hilbert Space

Heisenberg's inequality (1.2) is an instance of a more general inequality concerning selfadjoint operators on a Hilbert space, which also has an interpretation in terms of quantum observables. Although our focus is on functions and their Fourier transforms, we shall take a little time to discuss this general situation. See Folland [37] for more information.

The states of a quantum mechanical system are represented by unit vectors in an appropriate Hilbert space $\mathcal{H}$, and the observable quantities of the system are represented by selfadjoint operators on $\mathcal{H}$. The way this works is as follows. If A is a selfadjoint operator, by the spectral theorem there is a projection-valued measure P on $\mathbb{R}$ such that $A = \int\lambda\,dP(\lambda)$. If u is a unit vector, the map $\mu_u(E) = \langle P(E)u,\,u\rangle$ is a probability measure on $\mathbb{R}$ that represents the distribution of the observable A in the state u. The mean and variance of this measure are given by

$$M(\mu_u) = \int\lambda\,\langle dP(\lambda)u,u\rangle = \langle Au,u\rangle,$$

$$V(\mu_u) = \int(\lambda - M(\mu_u))^2\,\langle dP(\lambda)u,u\rangle = \|(A - M(\mu_u))u\|^2.$$

$M(\mu_u)$ represents the expected value of A in the state u, while $V(\mu_u)$ is a measure of the uncertainty of A in the state u. In this context, the general uncertainty principle says that there is a positive lower bound for the product of the uncertainties of two observables A and B in a state u whenever $\langle ABu, u\rangle \neq \langle BAu, u\rangle$.

To make this more precise, suppose A and B are densely defined operators on $\mathcal{H}$, with domains $D(A)$ and $D(B)$. Then the domain of the product AB is

$$D(AB) = \{u \in D(B) : Bu \in D(A)\},$$

and likewise for $D(BA)$. The commutator $[A, B]$ is defined as

$$[A, B] = AB - BA \quad \text{on} \quad D([A, B]) = D(AB) \cap D(BA).$$

Note that $D([A, B]) \subset D(A) \cap D(B)$.

Proposition 2.1.

If A and B are selfadjoint operators and $\alpha, \beta \in \mathbb{C}$,

$$\|(A-\alpha)u\| \, \|(B-\beta)u\| \geq \tfrac{1}{2}|\langle [A, B]u, u\rangle| \quad \textit{for all } u \in D([A, B]). \tag{2.2}$$

Proof. Since subtracting multiples of the identity operator from A and B does not affect $[A, B]$, we may assume that $\alpha = \beta = 0$. If $u \in D([A, B])$,

$$|\langle [A, B]u, u\rangle| = |\langle Bu, Au\rangle - \langle Au, Bu\rangle| = 2|\operatorname{Im}\langle Au, Bu\rangle| \leq 2\|Au\| \, \|Bu\|. \quad \square$$

The triviality of this proof should arouse one's suspicions, and indeed there is less to Proposition 2.1 than meets the eye. In the first place, $D([A, B])$ need not be dense in $\mathcal{H}$; it can even be $\{0\}$. This rarely happens in practice, but a more subtle difficulty is lurking in the shadows. The operator $[A, B]$ is usually not closed. If we denote its closure (the operator whose graph is the closure of the graph of $[A, B]$ in $\mathcal{H} \times \mathcal{H}$) by C, that is, $C = \overline{[A, B]}$, we would expect to have

$$\|Au\| \, \|Bu\| \geq \tfrac{1}{2}|\langle Cu, u\rangle| \quad \text{for all } u \in D(A) \cap D(B) \cap D(C). \tag{2.3}$$

But this is generally *false*. For example, take $\mathcal{H} = L^2([0, 1])$; $Af = if'$ on the domain of all absolutely continuous f on $[0, 1]$ such that $f' \in L^2$ and $f(0) = f(1)$; and $Bf(x) = xf(x)$ ($D(B) = \mathcal{H}$). Then $[A, B] = iI$ on the domain of all absolutely continuous f such that $f' \in L^2$ and $f(0) = f(1) = 0$. Since this domain is dense in $\mathcal{H}$ and $[A, B]$ is bounded, C is simply iI on $\mathcal{H}$. But if u is the constant function 1, we have $Au = 0$ while $|\langle Cu, u\rangle| = 1$, in violation of (2.3).

Of course, (2.3) follows immediately from (2.2) if for any $u \in D(A) \cap D(B) \cap D(C)$ there is a sequence $\{u_k\}$ in $D([A, B])$ such that $u_k \to u$, $Au_k \to Au$, $Bu_k \to Bu$, and $Cu_k \to Cu$; the trouble with the above example is that this condition does not hold. The following theorem, a slight extension of a result of Kraus [69] (but with a new proof), describes an important situation in which it does.

Theorem 2.4.

Let G be a connected Lie group with Lie algebra $\mathfrak{g}$, and let π be a unitary representation of G on $\mathcal{H}_\pi$. Suppose that $X, Y \in \mathfrak{g}$ and that the linear span $\mathfrak{I}$ of X, Y, and $[X, Y]$ is an ideal in $\mathfrak{g}$. Then (2.3) holds with $A = \pi(X)$, $B = \pi(Y)$, and $C = \pi([X, Y])$.

Remark. Let $\mathcal{H}_\pi^\infty$ be the space of C^∞ vectors for π. By results of Nelson [87], $C|\mathcal{H}_\pi^\infty$ is essentially skew-adjoint. Since $[A, B]$ is skew-Hermitian and $[A, B]|\mathcal{H}_\pi^\infty = C|\mathcal{H}_\pi^\infty$, it follows that $C = \overline{[A, B]}$.

To prove Theorem 2.4, choose a sequence $\{\phi_k\} \subset C_c^\infty(G)$ such that $\int \phi_k = 1$ for all k, $\operatorname{supp}(\phi_k) \to \{1\}$ as $k \to \infty$, and $\sup_k \int |\phi_k| = M < \infty$. It is a classic result of Gårding (see Knapp [67, p. 56]) that if $u \in \mathcal{H}_\pi$, then $\pi(\phi_k)u \in \mathcal{H}_\pi^\infty$ and $\pi(\phi_k)u \to u$ as $k \to \infty$. We shall show that if $u \in D(\pi(Z))$ for all $Z \in \mathfrak{I}$, then $\pi(Z)\pi(\phi_k)u \to \pi(Z)u$ for all such Z. Since $\mathcal{H}_\pi^\infty \subset D([\pi(X), \pi(Y)])$, the theorem then follows from the remarks preceding the statement.

Suppose then that $u \in D(\pi(Z))$ for all $Z \in \mathfrak{I}$. We first claim that $\pi(x)u \in D(\pi(Z))$ and $\pi(Z)\pi(x)u = \pi(x)\pi(\mathrm{Ad}(x^{-1})Z)u$ for all $Z \in \mathfrak{I}$ and $x \in G$. Indeed, since G is generated by $\exp \mathfrak{g}$ and $\mathrm{Ad}(\exp W) = \exp(\mathrm{ad}\, W)$ where $(\mathrm{ad}\, W)(Z) = [W, Z] \in \mathfrak{I}$, we have $u \in D(\pi(\mathrm{Ad}(x)Z))$ for all $x \in G$ and $Z \in \mathfrak{I}$. But then

$$\pi(Z)\pi(x)u = \frac{d}{dt}\pi((\exp\, tZ)x)u\bigg|_{t=0} = \pi(x)\frac{d}{dt}\pi(x^{-1}(\exp\, tZ)x)u\bigg|_{t=0} = \pi(x)\pi(\mathrm{Ad}(x^{-1})Z)u.$$

Next, integrating both sides of this equation against $\phi_k(x)$, we obtain

$$\begin{aligned}\pi(Z)\pi(\phi_k)u &= \int \phi_k(x)\pi(x)\pi(\mathrm{Ad}(x^{-1})Z)u\,dx \\ &= \pi(\phi_k)\pi(Z)u + \int \phi_k(x)\pi(x)\big[\pi(\mathrm{Ad}(x^{-1})Z)u - \pi(Z)u\big]\,dx.\end{aligned}$$

Now $\pi(\phi_k)\pi(Z)u \to \pi(Z)u$ as $k \to \infty$, and the second term on the right is bounded by

$$M \sup_{x \in \mathrm{supp}\,\phi_k} \|\pi(\mathrm{Ad}(x^{-1})Z)u - \pi(Z)u\|.$$

Since $x \to \mathrm{Ad}(x^{-1})Z$ is continuous from G to $\mathfrak{I}$ and since $Z \to \pi(Z)u$ is linear and hence continuous from $\mathfrak{I}$ to the finite-dimensional space $\pi(\mathfrak{I})u$, this supremum tends to zero as $k \to \infty$, and we are done. □

As far as we know, it is an open question whether the hypothesis in Theorem 2.4—that X, Y, and $[X, Y]$ span an ideal—is necessary. In stating the theorem, one could perfectly well assume that $\mathfrak{g} = \mathfrak{I}$ (the case considered by Kraus [69]), but the statement as given is natural for the proof and also for the most important application, the n-dimensional generalization of Theorem 1.1.

To wit, on $L^2(\mathbb{R}^n)$ we consider the selfadjoint operators P_j and Q_j $(1 \le j \le n)$ corresponding in quantum mechanics to the components of momentum and position (with Planck's constant h taken to be 1). They are defined by

$$P_j f = \mathcal{F}^{-1}[\xi_j \widehat{f}(\xi)] = \frac{1}{2\pi i}\frac{\partial f}{\partial x_j}, \qquad Q_j f(x) = x_j f(x)$$

on the domains of all $f \in L^2$ such that $\xi_j \widehat{f} \in L^2$ or $x_j f \in L^2$, respectively. $2\pi i P_j$ and $2\pi i Q_j$ are the infinitesimal generators of the one-parameter unitary groups

$$U_j(t)f(x) = f(x + te_j), \qquad V_j(t)f(x) = e^{2\pi i t x_j} f(x),$$

where e_j is the jth standard basis vector for $\mathbb{R}^n$. These groups fit together to make a unitary representation σ of the $(2n + 1)$-dimensional *Heisenberg group* H_n, which is the group whose underlying set is $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ and whose group law is given by

$$(p, q, z)(p', q', z') = \big(p + p',\ q + q',\ z + z' + \tfrac{1}{2}(p \cdot q' - q \cdot p')\big).$$

Its Lie algebra $\mathfrak{h}_n$ has the same underlying set, with Lie product

$$\big[(p, q, z),\ (p', q', z')\big] = (0,\ 0,\ p \cdot q' - q \cdot p'),$$

and the exponential map $\exp : \mathfrak{h}_n \to H_n$ is the identity map in these coordinates. The representation in question is

$$\sigma(p, q, z)f(x) = e^{2\pi i z + 2\pi i q \cdot x + \pi i p \cdot q} f(x + p). \tag{2.5}$$

Clearly $U_j(t) = \pi(te_j, 0, 0)$ and $V_j(t) = \pi(0, te_j, 0)$; thus, if we set

$$X_j = (e_j, 0, 0), \qquad Y_j = (0, e_j, 0), \qquad Z = (0, 0, 1),$$

we have

$$\sigma(X_j) = 2\pi i P_j, \qquad \sigma(Y_j) = 2\pi i Q_j, \qquad \sigma([X_j, Y_j]) = \sigma(Z) = 2\pi i I.$$

For each j, the span of X_j, Y_j, and Z is an ideal in $\mathfrak{h}_n$, so Theorem 2.4 applies to give the following result.

Corollary 2.6.

If $f \in L^2(\mathbb{R}^n)$, $a, b \in \mathbb{R}^n$, and $1 \le j \le n$,

$$\int (x_j - a_j)^2 |f(x)|^2 \, dx \int (\xi_j - b_j)^2 |\widehat{f}(\xi)|^2 \, d\xi \ge \frac{\|f\|_2^4}{16\pi^2}.$$

The case $n = 1$ of Corollary 2.6 is of course Theorem 1.1. The basic integration-by-parts argument in the proof of Theorem 1.1 finds its general expression in Proposition 2.1, and the approximation argument needed to finish the proof in dimensions $n > 1$ is embodied in Theorem 2.4. If one works out the proof of Theorem 2.4 for the particular case considered here, one finds that the space of C^∞ vectors for the representation σ in (2.5) is simply the Schwartz class $\mathcal{S}(\mathbb{R}^n)$ and the approximation procedure in the proof of Theorem 2.4 amounts to approximating L^2 functions by Schwartz functions in the obvious way, that is convolving with a smooth bump function and then multiplying by a smooth cutoff function. (More precisely, this is the result if one takes the functions ϕ_k in the proof to be of the form $\phi_k(p, q, z) = e^{-\pi i p \cdot q} \phi_k^1(p) \phi_k^2(q) \phi_k^3(z)$.)

In general, two selfadjoint operators A and B on a Hilbert space $\mathcal{H}$ such that $\overline{[A, B]} = (2\pi i)^{-1} I$ are said in quantum mechanics to be *canonically conjugate*. In this case the prescription $\pi(X) = 2\pi i A$, $\pi(Y) = 2\pi i B$, $\pi(Z) = 2\pi i I$ defines a representation of the Lie algebra $\mathfrak{h}_1$. Theorem 2.4 implies that one has an uncertainty inequality

$$\|(A - \alpha)u\| \, \|(B - \beta)u\| \ge \frac{\|u\|^2}{4\pi} \qquad (u \in \mathcal{H}) \tag{2.7}$$

(with the understanding that the left side is infinite if $u \notin D(A)$ or $u \notin D(B)$) *provided that* the representation π of $\mathfrak{h}_1$ exponentiates to a unitary representation of H_1. But the example following Proposition 2.1 shows that this hypothesis is not a mere formality.

Taking the square root of both sides in Corollary 2.6, summing over j, and using the Schwarz inequality for vectors in $\mathbb{C}^n$, we arrive at the following n-dimensional form of Heisenberg's inequality. We shall obtain an improved version of it in §5.

Corollary 2.8.

If $f \in L^2(\mathbb{R}^n)$ and $a, b \in \mathbb{R}^n$,

$$\int |x - a|^2 |f(x)|^2 \, dx \int |\xi - b|^2 |\widehat{f}(\xi)|^2 \, d\xi \ge \frac{n^2}{16\pi^2} \|f\|_2^4. \tag{2.9}$$

When specialized to radial functions, Corollary 2.8 is equivalent to an inequality concerning Hankel transforms of integer or half-integer order. The generalization to Hankel transforms of arbitrary positive order has been established by Bowie [21].

Gesztesy and Pittner [42] give further conditions under which (2.3) is valid, with examples and counterexamples. Chistyakov [24] has some generalizations of Proposition 2.1 to n-tuples of operators. Ishigaki [59] discusses relationships between uncertainty inequalities of the type (2.7) and other conditions on the operators A and B. Lahti and Maczynski [71] examine the role of uncertainty inequalities in general quantum logics. Kempf [65] derives an uncertainty inequality for operators satisfying a quantum-group analogue of the canonical commutation relation. Helffer and Nourrigat [52] prove a Heisenberg-type inequality for systems of pseudodifferential operators satisfying a generalized form of the canonical commutation relations. Spera [109] discusses uncertainty inequalities in the context of geometric quantization of Kähler manifolds and shows that certain analogues of Gaussian wave packets provide the extremal functions.

3. Variations on Heisenberg's Inequality

In this section we return to the question of giving precise formulations of the principle (0.1) by considering some generalizations and modifications of Theorem 1.1. For this discussion we shall avail ourselves of the transformation (0.2) at the outset to remove the constants a and b from Theorem 1.1 and write the conclusion in the form

$$\|xf\|_2\|\xi\widehat{f}\|_2 \geq \frac{\|f\|_2^2}{4\pi}. \tag{3.1}$$

One obvious way to extend (3.1) is to replace L^2 norms by L^p norms or the factors of x and ξ by other powers of x and ξ. For example, we can obtain the generalization

$$\|xf\|_p\|\xi\widehat{f}\|_p \geq \frac{\|f\|_2^2}{4\pi} \qquad (1 \leq p \leq 2)$$

by starting with (1.3) and applying first Hölder's inequality,

$$\|f\|_2^2 \leq 2\|xf\|_p\|f'\|_{p'},$$

and then the Hausdorff–Young inequality together with the Fourier inversion theorem,

$$\|f'\|_{p'} \leq \|(f')^\wedge\|_p = 2\pi\|\xi\widehat{f}\|_p.$$

More generally, one can consider inequalities of the form

$$\big\||x|^a f\big\|_p^\gamma \big\||\xi|^b\widehat{f}\big\|_q^{1-\gamma} \geq K\|f\|_2 \qquad (f \in L^2(\mathbb{R})), \tag{3.2}$$

where $a, b \in (0, \infty)$, $p, q \in [1, \infty]$, and $\gamma \in (0, 1)$. Two observations are crucial to the understanding of (3.2). First, invariance under dilations imposes a restriction on the parameters a, b, p, q, γ; and second, under this restriction, (3.2) is equivalent to an analogous "additive" inequality. To be precise, we have the following lemma.

Lemma 3.3.

A necessary condition for the validity of (3.2) *is that*

$$\gamma\left(a + \frac{1}{p} - \frac{1}{2}\right) = (1-\gamma)\left(b + \frac{1}{q} - \frac{1}{2}\right). \tag{3.4}$$

Moreover, if (3.4) *is satisfied,* (3.2) *is equivalent to*

$$\gamma\big\||x|^a f\big\|_p + (1-\gamma)\big\||\xi|^b\widehat{f}\big\|_q \geq K\|f\|_2 \qquad (f \in L^2(\mathbb{R})). \tag{3.5}$$

Proof. Let $f_c(x) = f(cx)$ $(c > 0)$. If we substitute f_c for f in (3.2), we obtain

$$c^{-\gamma(a+(1/p))+(1-\gamma)(b+(1/q)-1)}\big\||x|^a f\big\|_p^\gamma\big\||\xi|^b\widehat{f}\big\|_q^{1-\gamma} \geq Kc^{-1/2}\|f\|_2.$$

If this is to be true for all c, the exponents of c on the left and right must be equal, that is, (3.4) must hold.

Inequality (3.2) implies (3.5) because of the elementary inequality $s^\gamma t^{1-\gamma} \leq \gamma s + (1-\gamma)t$ $(s, t \geq 0,\ \gamma \in (0, 1))$. On the other hand, if we substitute f_c for f in (3.5) and multiply through by $c^{1/2}$, we obtain

$$\gamma c^{-a-(1/p)+(1/2)}\big\||x|^a f\big\|_p + (1-\gamma)c^{b+(1/q)-(1/2)}\big\||\xi|^b\widehat{f}\big\|_q \geq K\|f\|_2.$$

If (3.5) is valid, then this inequality holds for all $c > 0$, and under condition (3.4) it is easily verified that the minimum value of the left side (as a function of c) is nothing but the left side of (3.2). □

It therefore suffices to study inequalities of the form (3.5), and if one is not worried about the sharpness of the constant K, one can dispense with the factors of γ and $(1-\gamma)$ on the left. Here there is no restriction of the form (3.4), and the definitive result has been obtained by Cowling and Price [26].

Theorem 3.6.

Suppose $p, q \in [1, \infty]$ and $a, b > 0$. There is a constant K such that

$$\left\| |x|^a f \right\|_p + \left\| |\xi|^b \widehat{f} \right\|_q \geq K \|f\|_2 \tag{3.7}$$

for all tempered functions f such that $\widehat{f}$ is also a function, if and only if

$$a > \frac{1}{2} - \frac{1}{p} \qquad \text{and} \qquad b > \frac{1}{2} - \frac{1}{q}. \tag{3.8}$$

Consequently, (3.2) *is valid (with perhaps a different constant K) if and only if* (3.4) *and* (3.8) *both hold.*

The proof of this theorem in [26] involves a fair amount of work, but it requires only standard real-variable machinery together with the fact that $\int_{-\delta}^{\delta} |f(x)|^2\, dx \leq C_\delta < 1$ when $\|f\|_2 = 1$ and $\widehat{f}$ is supported in a fixed bounded set, which we shall prove (Theorem 8.4). The case $p = q = 2$, $a = b > 0$ of (3.2) was first obtained via different methods by Hirschman [54].

Cowling and Price [26] also prove generalizations of (3.7) in which $|x|^a$ and $|\xi|^b$ are replaced by more general weight functions. Generalizations of (3.2) of the same sort can be found in Benedetto [10], [11] and Heinig and Smith [50].

De Bruijn [29] observed that a sharpened form of Theorem 1.1 can be derived using the Hermite functions

$$h_k(x) = \frac{2^{1/4}}{\sqrt{k!}} \left(\frac{-1}{2\sqrt{\pi}} \right)^k e^{\pi x^2} \frac{d^k}{dx^k} (e^{-2\pi x^2}).$$

(The normalizations here are a bit different from the usual ones.) It is well known (see Folland [37, §1.7]) that (i) $\{h_k\}_0^\infty$ is an orthonormal basis for $L^2(\mathbb{R})$, (ii) $\widehat{h}_k = i^{-k} h_k$, and (iii) $2\sqrt{\pi}\, x h_k(x) = \sqrt{k+1}\, h_{k+1}(x) + \sqrt{k}\, h_{k-1}(x)$. Given $f \in L^2(\mathbb{R})$, if one expands f in a Hermite series according to (i) and then uses (ii) and (iii) to obtain the corresponding expansions of xf and $\xi \widehat{f}$, one easily arrives at the identity

$$\|xf\|_2^2 + \|\xi \widehat{f}\|_2^2 = \frac{1}{2\pi} \sum_0^\infty (2k+1) |\langle f, h_k \rangle|^2. \tag{3.9}$$

Since $\|f\|_2^2 = \sum_0^\infty |\langle f, h_k \rangle|^2$, this implies that

$$\|xf\|_2^2 + \|\xi \widehat{f}\|_2^2 \geq \frac{\|f\|_2^2}{2\pi},$$

with equality if and only if $f(x) = ch_0(x) = c' e^{-\pi x^2}$. This in turn implies Theorem 1.1 by the dilation argument used to derive (3.2) from (3.5). Another sharpened form of Theorem 1.1 using Hermite functions can be found in Mustard [83].

The identity (3.9) also yields an improvement on Theorem 1.1 for odd functions. Namely, if f is odd, then $\langle f, h_k \rangle = 0$ for k even, so

$$\|xf\|_2^2 + \|\xi \widehat{f}\|_2^2 \geq \frac{3\|f\|_2^2}{2\pi} \qquad (f(-x) = -f(x)).$$

with equality if and only if $f(x) = ch_1(x) = c'xe^{-\pi x^2}$. As above, this implies that

$$\|xf\|_2\|\xi \widehat{f}\|_2 \geq \frac{3\|f\|_2^2}{4\pi} \qquad (f(-x) = -f(x)),$$

with equality if and only if $f = cxe^{-dx^2}$. (Note that if f is odd, then $|f|^2$ and $|\widehat{f}|^2$ are centered at 0, so the left side of (1.2) is minimized at $a = b = 0$.) Skoog [102] has used this result to derive an improved uncertainty inequality for functions vanishing on a half-line.

The Hermite function h_k is the eigenfunction with eigenvalue $2\pi(2k+1)$ of the Hermite operator $-(d/dx)^2 + 4\pi^2x^2$. Thus the extremal functions for Heisenberg's inequality (1.2) are the ground states (eigenfunctions of lowest eigenvalues) for the operators $-(d/dx)^2 + cx^2$ $(c > 0)$. These are the quantum Hamiltonians for particles moving in a potential well $V(x) = cx^2$, and it is reasonable to expect the ground states for $-(d/dx)^2 + V(x)$ also to have a rather small uncertainty product for many other potentials V. Some precise results along these lines have been obtained by Kahane, Lévy-Leblond, and Sjöstrand [61]. Moreover, Borchi and Marsaglia [19] have observed that one can find functions f supported in a finite interval $[-a, a]$ (and with $\|f\|_2 = 1$) for which $V(|f|^2)V(|\widehat{f}|^2)$ is arbitrarily close to $1/16\pi^2$ by taking f to be the ground state for a Hermite operator $-(d/dx)^2 + cx^2$ on $[-a, a]$ subject to the boundary condition $f(-a) = f(a) = 0$.

Theorem 1.1 is somewhat unsatisfactory from the point of view of signal analysis, for the following reason. Suppose f represents the amplitude of a signal; for convenience we assume $\|f\|_2 = 1$. f must be real-valued, which means that $\widehat{f}(-\xi) = \overline{\widehat{f}(\xi)}$, and in particular $|\widehat{f}|^2$ is even. Thus, to say that a signal is localized in frequency can only mean that $|\widehat{f}|^2$ has a peak at some point ξ_0 and an equal one at $-\xi_0$. But if ξ_0 is large, the variance $V(|\widehat{f}|^2)$ will be large even if the peaks are narrow, so Heisenberg's inequality provides little information. One way around this difficulty is to use the local uncertainty inequalities that we shall discuss in the next section. Another one, suggested by Gabor [40], is to replace f by the "complex signal" $f + iHf$ (H being the Hilbert transform), whose Fourier transform is $2\widehat{f}\chi_{(0,\infty)}$. A third one is to to consider the "one-sided variance"

$$V^+(|\widehat{f}|^2) = \inf_{b>0} \int_0^\infty (\xi - b)^2 |\widehat{f}(\xi)|^2 \, d\xi$$

instead of $V(|\widehat{f}|^2)$. Hilberg and Rothe [53] have shown that for real f with $\|f\|_2 = 1$, the product $V(|f|^2)V^+(|\widehat{f}|^2)$ has a positive lower bound, which is the smallest eigenvalue of a certain Sturm–Liouville problem, and the extremal functions are the Fourier transforms of the corresponding eigenfunctions. See also Kay and Silverman [64] for the earlier history of this problem.

Uffink and Hilgevoord [115, 116] have developed a different version of the uncertainty principle. Given two fixed numbers $\alpha, \beta \in (0, 1)$, for a function $f \in L^2(\mathbb{R})$ with $\|f\|_2 = 1$ they define the *mean width* $W(f)$ and the *mean peak width* $w(f)$ of f to be, respectively, the smallest W and the smallest w such that

$$\max_c \int_{c-W/2}^{c+W/2} |f(x)|^2 \, dx = \alpha, \qquad \left| \int f(x-w)\overline{f(x)} \, dx \right| = \beta.$$

They then derive inequalities relating $W(f)$ to $w(\widehat{f})$ and argue that these inequalities capture the physics of the uncertainty principle more effectively than Heisenberg's inequality.

Garofalo and Lanconelli [41], Thangavelu [114], and Sitaram, Sundari, and Thangavelu [101] have obtained three related but inequivalent analogues of Heisenberg's inequality for functions on the Heisenberg group H_n. Price and Sitaram [97] and Hogan [56] have obtained inequalities of the same sort for functions on symmetric spaces of noncompact type and locally compact Abelian groups, respectively, and Thangavelu [114] has some related results for Hermite and Laguerre expansions. Nahmod [84] has derived an uncertainty inequality in a very general setting that relates to the spectral geometry of elliptic and subelliptic operators.

4. Local Uncertainty Inequalities

Heisenberg's inequality says that if f is highly localized, then $\widehat{f}$ cannot be concentrated near a single point, but it does not preclude $\widehat{f}$ from being concentrated in a small neighborhood of two or more widely separated points. In fact, the latter phenomenon cannot occur either, and it is the object of local uncertainty inequalities to make this precise.

The first such inequalities were obtained by Faris [34], and they were subsequently sharpened and generalized by Price [91, 93]. The principal results in the setting of L^2 norms are summarized in the following theorem. As in §3, we implicitly use (0.2) to reduce to the case where f and $\widehat{f}$ are centered at the origin.

Theorem 4.1.

i. *If $0 < \alpha < \frac{1}{2}n$, there is a constant K_α such that for all $f \in L^2(\mathbb{R}^n)$ and all measurable $E \subset \mathbb{R}^n$,*

$$\int_E |\widehat{f}|^2 \le K_\alpha |E|^{2\alpha/n} \big\| |x|^\alpha f \big\|_2^2.$$

ii. *If $\alpha > \frac{1}{2}n$, there is a constant K_α such that for all $f \in L^2(\mathbb{R}^n)$ and all measurable $E \subset \mathbb{R}^n$,*

$$\int_E |\widehat{f}|^2 \le K_\alpha |E| \, \|f\|_2^{2-(n/\alpha)} \big\| |x|^\alpha f \big\|_2^{n/\alpha}.$$

Part **i** is proved in Price [91] and Price and Sitaram [97], and part **ii** is proved in Price [93]. For both parts, the case $\alpha = 1$ is due to Faris [34]; related results are in Benedetto [10]. Price [91] also contains a generalization involving the L^p norm of $|x|^\alpha f$ rather than the L^2 norm. As in the preceding section, the relations among the exponents in these inequalities are forced by homogeneity considerations. As discussed in [93], examples show that the restriction $\alpha > \frac{1}{2}n$ is necessary for part **ii**; hence part **i** also fails for $\alpha = \frac{1}{2}n$, and it is a simple exercise to see that **i** cannot hold for $\alpha > \frac{1}{2}n$. The constants K_α can be described quite explicitly, but we shall not do so here.

Let us indicate the proof of part **i**. Let χ_r denote the characteristic function of $\{x : |x| < r\}$ and $\chi_r' = 1 - \chi_r$. Then for any $r > 0$ we can write

$$\begin{aligned}\left(\int_E |\widehat{f}|^2\right)^{1/2} &= \|\widehat{f}\chi_E\|_2 \le \|(f\chi_r)\widehat{\ }\chi_E\|_2 + \|(f\chi_r')\chi_E\|_2 \\ &\le |E|^{1/2}\|(f\chi_r)\widehat{\ }\|_\infty + \|f\chi_r'\|_2.\end{aligned}$$

Now

$$\|(f\chi_r)\widehat{\ }\|_\infty \le \|f\chi_r\|_1 \le \big\| |x|^{-\alpha}\chi_r \big\|_2 \big\| |x|^\alpha f \big\|_2 \le C_\alpha r^{(n/2)-\alpha} \big\| |x|^\alpha f \big\|_2$$

and

$$\|f\chi_r'\|_2 \le \big\| |x|^{-\alpha}\chi_r' \big\|_\infty \big\| |x|^\alpha f \big\|_2 \le r^{-\alpha} \big\| |x|^\alpha f \big\|_2,$$

so

$$\left(\int_E |\widehat{f}|^2\right)^{1/2} \le (C_\alpha |E|^{1/2} r^{(n/2)-\alpha} + r^{-\alpha}) \big\| |x|^\alpha f \big\|_2.$$

The desired result is obtained by choosing r so as to minimize the quantity on the right.

Each of the inequalities in Theorem 4.1, for any fixed value of α, implies a corresponding global uncertainty inequality of the type (3.2). For example, if $0 < \alpha < \frac{1}{2}n$, we have

$$\|f\|_2^2 = \|\widehat{f}\|_2^2 = \int_{|\xi|<r} |\widehat{f}|^2 + \int_{|\xi|\geq r} |\widehat{f}|^2$$
$$\leq K'_\alpha r^{2\alpha} \big\| |x|^\alpha f \big\|_2^2 + r^{-2\alpha} \big\| |\xi|^\alpha \widehat{f} \big\|_2^2.$$

Choosing r so as to minimize the expression on the right, we obtain

$$\|f\|_2^2 \leq K''_\alpha \big\| |x|^\alpha f \big\|_2 \big\| |\xi|^\alpha \widehat{f} \big\|_2.$$

A similar argument yields the same result when $\alpha > \frac{1}{2}n$. Thus, local uncertainty inequalities are qualitatively stronger than the global ones of §3. It should be noted, however, that in the case $\alpha = 1$ the constant K''_1 thus obtained is not the optimal constant $4\pi/n$ of Corollary 2.8, even if one uses the best constant K_α in Theorem 4.1.

The form of the inequalities in Theorem 4.1 adapts itself readily to other Plancherel groups. Indeed, the analogue of part i for such a group should be

$$\int_E \|\widehat{f}(\pi)\|_{\mathrm{HS}}^2 \, d\pi \leq K_\theta |E|^{2\theta} \|w^\theta f\|_2^2 \qquad (0 < \theta < \tfrac{1}{2}), \tag{4.2}$$

where $|E|$ is the Plancherel measure of E and w is a weight function on G related to the distance to the group identity or perhaps the distance to a suitable "thin" subset of G. Results of this sort have been obtained in the following situations:

1. (Price and Racki [94]) G is the n-torus $\mathbb{T}^n$ and $w(x) = |x|^n$, $|x|$ being the Euclidean distance from x to the identity. There is also a generalization with $\|w^\theta f\|_2$ replaced by $\|w^\theta f\|_p$.
2. (Price and Sitaram [96]) G is a compact metric group and $w(x)$ is the measure of the smallest ball about the identity containing x. Here the $|E|$ on the right, however, is not Plancherel measure but a somewhat larger measure. An analogue for functions on compact Riemannian manifolds, relative to the spectral decomposition of the Laplacian, is also given.
3. (Price and Sitaram [97]) G is either a noncompact semisimple Lie group or a Euclidean motion group, and $w(x)$ is the measure of the set of points whose distance (in a suitable sense) to the maximal compact subgroup K of G is at most that of x. Here, however, the authors establish (4.2) only for K-finite functions; this restriction is necessary to obtain a bound for $\|\widehat{f}(\pi)\|_{\mathrm{HS}}$ in terms of $\|f\|_1$.
4. (Price and Sitaram [97]) G is the Heisenberg group H_n described in §2 and $w(p, q, z) = |z|$. Another version of (4.2) for H_n, with $w(p, q, z) = ((|p|^2 + |q|^2)^2 + |z|^2)^{n+1}$ but involving a more refined description of the Fourier transform on H_n, appears in [101].
5. (Hogan [56]) G is a locally compact Abelian group.

In all of these cases, the basic idea of the proof is similar to that given above, but the complete argument involves results from the representation theory of the group in question.

We return to $\mathbb{R}^n$. One consequence of Theorem 4.2 is that if $\|f\|_2 = 1$ and E is the *complement* of a set of small measure, a sufficiently small upper bound on $\int_E |\widehat{f}|^2$ will force a positive lower bound on the variance $V(|f|^2)$. Strichartz [110] has shown that a similar result holds for much "thinner" sets E provided that they are "evenly distributed." For example, if E is the union of a collection of evenly spaced concentric spheres or a collection of evenly spaced parallel hyperplanes and σ denotes surface measure on E, a sufficiently small bound on $\int_E |\widehat{f}|^2 \, d\sigma$ will imply a positive lower bound on $V(|f|^2)$. (It is no restriction to assume that $V(|f|^2) < \infty$, which means that $\widehat{f}$ belongs to the Sobolev space L_1^2; this is enough to guarantee that the restriction of $\widehat{f}$ to a codimension-one submanifold is well defined.) Strichartz [110] also has a similar result for functions f on the unit sphere $S_n \subset \mathbb{R}^n$: If $f = \sum f_k$ is the expansion of f in spherical harmonics and E is an "evenly spaced" subset of $\mathbb{Z}^+$, a small upper bound on $\sum_E \|f_k\|_2^2$ will imply a positive lower bound on $\int_{S_n} \sin^2 d(x)|f(x)|^2 \, d\sigma(x)$ where $d(x)$ is the distance from x to a fixed $x_0 \in S_n$. An analogue of Strichartz's theorem for functions on real hyperbolic n-space has been obtained by Sun [112].

If the function f is supported in a bounded set, one easily obtains bounds on $\widehat{f}$ and its derivatives that limit the concentration of $\widehat{f}$ in any small set and may provide lower bounds for the concentration of $\widehat{f}$ in sufficiently large sets. For example, one has the following simple local uncertainty inequality:

$$\int_E |\widehat{f}|^2 \leq |E| \, \|\widehat{f}\|_\infty^2 \leq |E| \, \|f\|_1^2 \leq |E| \, |\{x : f(x) \neq 0\}| \, \|f\|_2^2. \tag{4.3}$$

A local uncertainty inequality in this spirit, but applying in some cases to sets E of infinite Plancherel measure, has been obtained for the spherical Fourier transform on certain noncompact symmetric spaces by Shahshahani [98].

We close by quoting an interesting theorem of Logvinenko and Sereda [79] and Kacnel'son [60] (see also Havin and Jöricke [48]), obtained by studying L^p norms on spaces of entire functions and then applying the Paley–Wiener theorem. If $E \subset \mathbb{R}^n$ and $1 \leq p < \infty$, the following conditions are equivalent: (i) for every bounded $B \subset \mathbb{R}^n$ there exists $c > 0$ such that $\int_E |\widehat{f}|^p \geq c\|\widehat{f}\|_p^p$ for all f supported in B, and (ii) there exist $\gamma > 0$ and a cube $K \subset \mathbb{R}^n$ such that $|E \cap (K + x)| \geq \gamma$ for all $x \in \mathbb{R}^n$.

5. Logarithmic Uncertainty Inequalities

Suppose ρ is a probability density function on $\mathbb{R}^n$. Following Shannon [99], we define the *entropy* of ρ to be

$$E(\rho) = -\int \rho(x) \log \rho(x) \, dx.$$

This notion of entropy is related but not identical to the more familiar entropy $-\sum p_i \log p_i$ (also due to Shannon [99]) of a probability distribution on a discrete sample space. Unlike the latter, $E(\rho)$ can have any value in $[-\infty, \infty]$, and it can also be undefined (i.e., of the form $\infty - \infty$). Clearly any sharp peaks in ρ will tend to make $E(\rho)$ negative, whereas a slowly decaying tail will tend to make $E(\rho)$ positive; hence $E(\rho)$ is a measure of how localized ρ is. (See Bialynicki-Birula [17] for a discussion of the significance of entropy in quantum mechanics.) $E(\rho)$ is related to the covariance matrix $V(\rho)$ as follows.

Theorem 5.1.

If ρ is a probability density function on $\mathbb{R}^n$ with finite variance, then $E(\rho)$ is well defined and

$$E(\rho) \leq \tfrac{1}{2} \log\left[(2\pi e)^n \det V(\rho)\right]. \tag{5.2}$$

This theorem is due to Shannon [99], who argued by proposing to maximize $E(\rho)$ among all ρ with a given variance. He solved a calculus of variations problem to find that the critical points for E are the Gaussians, computed $E(\rho)$ for ρ Gaussian, and claimed (5.2) as a result. That the critical points actually give the global maximum can be established by using the concavity of the functional E—a point Shannon omitted to mention. Rather than give the details, we shall present an elegant proof that was communicated to us by W. Beckner.

By composing ρ with a translation and a rotation, which does not affect the quantities in (5.2), we may assume that $M(\rho) = 0$ and that the covariance matrix $V_{jk}(\rho) = \int x_j x_k \rho(x) \, dx$ is diagonal. Moreover, if $\rho(x)$ is replaced by

$$c_1 \cdots c_n \rho(c_1 x_1, \ldots, c_n x_n) \qquad (c_1, \ldots, c_n > 0),$$

then both sides of (5.2) decrease by the amount $\sum \log c_j$; so by taking $c_j = \sqrt{V_{jj}(\rho)}$ we may even assume that $V(\rho) = I$. Let

$$\phi(x) = (2\pi)^{n/2} e^{|x|^2/2} \rho(x), \qquad d\gamma(x) = (2\pi)^{-n/2} e^{-|x|^2/2}\, dx$$

so that $\int \phi\, d\gamma = \int \rho\, dx = 1$. Since γ is a probability measure and $t \log t$ is a convex function of t, Jensen's inequality gives

$$\begin{aligned} 0 = \left[\int \phi\, d\gamma\right] \log \left[\int \phi\, d\gamma\right] &\le \int \phi \log \phi\, d\gamma \\ &= \int \rho(x) \left[\frac{n}{2} \log 2\pi + \frac{1}{2}|x|^2 + \log \rho(x)\right] dx \\ &= \frac{n}{2} \log 2\pi + \frac{1}{2} \sum V_{jj}(\rho) - E(\rho). \end{aligned}$$

Since $V_{jj}(\rho) = 1 = \det V(\rho)$, (5.2) follows.

The fundamental uncertainty inequality in terms of entropy is the following.

Theorem 5.3.

If $f \in L^2(\mathbb{R}^n)$ and $\|f\|_2 = 1$, we have

$$E(|f|^2) + E(|\widehat{f}|^2) \ge n(1 - \log 2)$$

whenever the left side is well defined.

Hirschman [54] conjectured Theorem 5.3 but was able to prove only the weaker inequality

$$E(|f|^2) + E(|\widehat{f}|^2) \ge 0. \tag{5.4}$$

(Leipnik [76] independently discovered Theorem 5.3, but his argument contains the same sort of gap as Shannon's proof of Theorem 5.1, and concavity is no help here.) Hirschman's proof of (5.4) consists of combining the Hausdorff–Young inequality $\|\widehat{f}\|_q \le \|f\|_p$ $(1 \le p \le 2,\ p^{-1} + q^{-1} = 1)$ with the following trivial but useful lemma.

Lemma 5.5.

Suppose $\phi(t) \le \psi(t)$ for $a \le t \le b$ and $\phi(a) = \psi(a)$. If ϕ and ψ are differentiable at a, then $\phi'(a) \le \psi'(a)$.

If one writes the Hausdorff–Young inequality as

$$\int |\widehat{f}|^q \le \left[\int |f|^{q/(q-1)}\right]^{q-1} \qquad (q \ge 2)$$

and applies Lemma 5.5 to the expressions on the left and right as functions of q (with $a = 2$), assuming f is such that all the integrals in question are finite for q near 2, one immediately obtains (5.4). (For the straightforward limiting argument to remove the restriction on f, we refer to Hirschman [54].) As observed by both Beckner [6] and Białynicki-Birula and Mycielski [18], Theorem 5.3 follows by applying the same argument to the sharp Hausdorff–Young inequality of Beckner [6],

$$\|\widehat{f}\|_q \le p^{n/2p} q^{-n/2q} \|f\|_p \qquad (1 \le p \le 2,\ p^{-1} + q^{-1} = 1). \tag{5.6}$$

If one combines Theorems 5.1 and 5.3, one immediately obtains the following corollary.

Corollary 5.7.

If $f \in L^2(\mathbb{R}^n)$ and $\|f\|_2 = 1$,

$$\det V(|f|^2) \det V(|\widehat{f}|^2) \ge (16\pi^2)^{-n}.$$

Corollary 5.7 is a strengthening of Corollary 2.8, for if ρ is any probability density function on $\mathbb{R}^n$,

$$[\det V(\rho)]^{1/n} \leq \frac{1}{n}\int |x - M(\rho)|^2 \rho(x)\,dx.$$

(This is just the inequality of arithmetic and geometric means applied to the eigenvalues of the matrix $V(\rho)$.) Hence Theorem 5.3 can be regarded as a sharp form of Heisenberg's inequality.

The preceding results are discussed in Heinig and Smith [50], which also contains a version of Theorem 5.3 with weighted norms.

Since the proof of Heisenberg's inequality is of an elementary nature, whereas Beckner's inequality (5.6) is a deep theorem, one may wonder whether we have used heavy machinery merely to obtain a mild improvement on Heisenberg's inequality or whether Theorem 5.3 is really a more powerful result. In fact, the latter alternative is the case. As Beckner [7] has shown, Theorem 5.3 yields a short proof of a remarkable improvement on Gross's logarithmic Sobolev inequality, itself a deep theorem closely related to (5.6) and Nelson's hypercontractivity theorem. (See Gross [44, 45] and Beckner [6].) Indeed, let

$$d\gamma(x) = (2\pi)^{-n/2}e^{-|x|^2/2}\,dx \qquad (x \in \mathbb{R}^n)$$

and define the unitary map $T : L^2(\gamma) \to L^2(\mathbb{R}^n)$ by

$$Tg(x) = 2^{n/4}e^{-\pi|x|^2}g(2\pi^{1/2}x).$$

If we apply Theorem 5.3 to $f = Tg$, use the facts that

$$\int |\xi|^2|\widehat{f}(\xi)|^2\,d\xi = \frac{1}{4\pi^2}\int |\nabla f(x)|^2\,dx$$

and

$$|\nabla f(x)|^2 = 2^{2+(n/2)}\pi\Big[\,|\nabla g(2\pi^{1/2}x)|^2 + \pi|x|^2|g(2\pi^{1/2}x)|^2$$
$$-2\pi^{1/2}\operatorname{Re}[g(2\pi^{1/2}x)x\cdot\overline{\nabla g(2\pi^{1/2}x)}]\Big]e^{-2\pi|x|^2},$$

and integrate the cross term in this last expression by parts, we obtain the following theorem.

Theorem 5.8.

Suppose $\int |g|^2\,d\gamma = 1$, *and let* $\widetilde{g} = T^{-1}\mathcal{F}Tg$. *Then*

$$\int |g|^2\log|g|\,d\gamma + \int |\widetilde{g}|^2\log|\widetilde{g}|\,d\gamma \leq \int |\nabla g|^2\,d\gamma. \tag{5.9}$$

Gross's inequality is (5.9) with the term involving $\widetilde{g}$ omitted, which follows from (5.9) since both terms on the left are nonnegative. (To see this, use Jensen's inequality as in the proof of Theorem 5.1 with $\phi = |g|^2$ or $|\widetilde{g}|^2$.)

Beckner [7] has recently proved another logarithmic uncertainty inequality:

$$\int |f(x)|^2\log|x-a|\,dx + \int |\widehat{f}(\xi)|^2\log|\xi-b|\,d\xi \geq [\psi(\tfrac{1}{4}n) - \log\pi]\int |f(x)|^2\,dx \tag{5.10}$$

for all $f \in L^2(\mathbb{R}^n)$ for which the quantity on the left is defined, where ψ is the logarithmic derivative of the gamma function. Beckner first establishes the inequality

$$\int |\widehat{f}(\xi)|^2|\xi|^{-\alpha}\,d\xi \leq \frac{\pi^\alpha\Gamma(\frac{1}{4}(n-\alpha))}{\Gamma(\frac{1}{4}(n+\alpha))}\int |f(x)|^2|x|^\alpha\,dx \qquad (0 \leq \alpha < n)$$

and then applies Lemma 5.5 to get (5.10) with $a = b = 0$; the general case follows by using (0.2). Like Theorem 5.3, (5.10) is related to logarithmic Sobolev inequalities and implies Heisenberg's inequality.

6. Wigner Distributions and Ambiguity Functions

In this section we discuss some uncertainty relations for functions on $\mathbb{R}^n$ that are expressed in terms of certain functions on $x\xi$-space ("phase space"). For a more detailed explanation of these ideas, including several calculations that are elided here, we refer the reader to Folland [37, §§1.4 and 1.8].

To begin with, we consider the matrix elements of the phase-space translations

$$\sigma(p, q, 0)f(x) = e^{2\pi i q\cdot x + \pi i p\cdot q} f(x + p).$$

(Here σ is the representation of H_n given by (2.5). Since $\sigma(p, q, z) = e^{2\pi i z}\sigma(p, q, 0)$, no essential information is lost by restricting to $z = 0$.) That is, for $f, g \in L^2(\mathbb{R}^n)$, we define

$$A(f, g)(p, q) = \langle \sigma(p, q, 0)f, g\rangle = \int e^{2\pi i q\cdot y} f(y + \tfrac{1}{2}p)\overline{g(y - \tfrac{1}{2}p)}\, dy. \tag{6.1}$$

$A(f, g)$ is called the *Fourier–Wigner transform* of f and g in Folland [37]; in the radar engineering literature it is known as the *cross ambiguity function* of f and g, and $A(f, f)$ is the *ambiguity function* of f. Also, $A(f, g)$ differs only by a factor of $e^{\pi i p\cdot q}$ and the substitution $q \to -q$ from the *windowed Fourier transform* or *short-time Fourier transform*

$$\mathcal{F}_g f(p, q) = \int e^{-2\pi i q\cdot y} g(y - p) f(y)\, dy.$$

The Fourier transform of $A(f, g)$ is the *Wigner transform* of f and g, namely,

$$\begin{aligned} W(f, g)(\xi, x) &= \iint e^{-2\pi i(\xi\cdot p + x\cdot q)} A(f, g)(p, q)\, dp\, dq \\ &= \int e^{-2\pi i \xi\cdot p} f(x + \tfrac{1}{2}p)\overline{g(x - \tfrac{1}{2}p)}\, dp. \end{aligned}$$

(The second equality follows from the Fourier inversion theorem.) Clearly $A(f, g)$ and $W(f, g)$ are related not only by the Fourier transform but by the more elementary identity

$$W(f, g)(\xi, x) = 2^n A(f, \widetilde{g})(2x, -2\xi) \qquad [\widetilde{g}(x) = g(-x)]. \tag{6.2}$$

$W(f, f)$ is called the *Wigner distribution* of f and has the following quantum interpretation. Suppose $\|f\|_2 = 1$, so f represents a quantum state. We would like to speak of the joint distribution ρ of momentum P and position Q in the state f. Such a thing does not exist because the uncertainty principle forbids the simultaneous determination of momentum and position, but if it did, its inverse Fourier transform $\check{\rho}(p, q) = \iint e^{2\pi i(p\cdot\xi + q\cdot x)}\rho(\xi, x)\, d\xi\, dx$ ought to be the expected value of the observable $\exp 2\pi i(p\cdot P + q\cdot Q)$ in the state f. But if we interpret P and Q as the operators $(2\pi i)^{-1}\partial/\partial x$ and x as in §2, $\exp 2\pi i(p\cdot P + q\cdot Q)$ is nothing but $\sigma(p, q, 0)$, so the desired expected value is $A(f, f)(p, q)$. Hence ρ ought to be $W(f, f)$.

This almost works! In general, $W(f, f)$ is not a probability distribution function because it can assume negative values, but it is not hard to verify that it has the right marginal distributions for position and momentum in the state f:

$$\int W(f, f)(\xi, x)\, d\xi = |f(x)|^2, \qquad \int W(f, f)(\xi, x)\, dx = |\widehat{f}(\xi)|^2. \tag{6.3}$$

Therefore, $W(f, f)$ can be considered a "phase-space portrait" of the function f. More classically, if $n = 1$ and f is interpreted as the amplitude of a signal, $W(f, f)$ is the "time-frequency portrait" of f; de Bruijn [29] calls it the "musical score" of f.

In this setting, the uncertainty principle says that $W(f, f)$ *cannot be too sharply localized*. Indeed, by (6.3), Corollary 2.8, and the inequality $\alpha^2 + \beta^2 \geq 2\alpha\beta$,

$$\begin{aligned}\iint (|x-a|^2 + |\xi - b|^2) W(f, f)(\xi, x)\, d\xi\, dx \\ = \int |x-a|^2 |f(x)|^2\, dx + \int |\xi - b|^2 |\widehat{f}(\xi)|^2\, d\xi \geq \frac{\|f\|_2^2}{2\pi},\end{aligned} \tag{6.4}$$

which is the analogue of Heisenberg's inequality for Wigner distributions.

Additional inequalities of this type can be found in de Bruijn [29]. Another intriguing result of de Bruijn [29] (see also Folland [37, §1.8]) is the following. As we have stated above, $W(f, f)$ can have negative values. However, let us set $\gamma_\epsilon(x) = \epsilon^{-n} e^{-\pi(x/\epsilon)^2}$, a Gaussian of total mass 1 whose peak has width roughly ϵ. Then

$$W_{\epsilon,\delta}(f, f)(\xi, x) = \iint W(f, f)(\xi - \eta, x - y)\gamma_\epsilon(\eta)\gamma_\delta(y)\, d\eta\, dy \geq 0 \tag{6.5}$$

for all $f \in L^2$ if and only if $\epsilon\delta \geq 2$. Inequality (6.5) guarantees that $W_{\epsilon,\delta}(f, f)$ is a genuine probability distribution function. Intuitively, it is the joint distribution of "momentum to within an error ϵ" and "position to within an error δ," and the uncertainty principle is the fact that this makes sense precisely when $\epsilon\delta \geq 2$. Further results along these lines can be found in Ali and Prugovecki [1] and Busch [22].

Other uncertainty inequalities for $W(f, f)$, or more generally $W(f, g)$, can be obtained by estimating its L^p norm. By (6.2), this is equivalent to estimating the L^p norm of $A(f, g)$, for which the formulas turn out to be a little simpler. First, it is obvious from (6.1) and the Schwarz inequality that

$$\|A(f, g)\|_\infty \leq \|f\|_2 \|g\|_2. \tag{6.6}$$

It is less obvious, but still easy to verify, that

$$\|A(f, g)\|_2 = \|f\|_2 \|g\|_2. \tag{6.7}$$

Thus $\int_E |A(f, g)|^2 \leq \|A(f, g)\|_\infty^2 |E| \leq \|A(f, g)\|_2^2 |E|$, so the mass of $|A(f, g)|^2$ cannot be concentrated in any set of small measure.

If we normalize f and g so that $\|f\|_2 \|g\|_2 = 1$, we have $|A(f, g)| \leq 1$ by (6.6) and (6.7), so $\int |A(f, g)|^p$ is a decreasing function of p. The rate of decrease is less rapid when $A(f, g)$ is more concentrated; the extreme case (not actually achieved) would be when $|A(f, g)| = \chi_E$ with $|E| = 1$. Hence the uncertainty principle can be embodied in a lower bound for the rate of decrease of $\int |A(f, g)|^p$ as p increases. In fact, Lieb [78] has shown that

$$\begin{aligned}\int |A(f, g)|^p \leq (2/p)^n \|f\|_2^p \|g\|_2^p \quad \text{if } p \geq 2, \\ \int |A(f, g)|^p \geq (2/p)^n \|f\|_2^p \|g\|_2^p \quad \text{if } p \leq 2.\end{aligned} \tag{6.8}$$

(Lieb [78] deals explicitly only with the case $n = 1$, but the passage to higher dimensions is straightforward. By (6.2), the same estimates hold for $\int |W(f, g)|^p$ with an extra factor of $2^{n(p-2)}$ on the right.) Lieb [78] also has generalizations of these estimates in which the L^2 norms of f and g are replaced by L^q norms for other values of q.

By applying Lemma 5.5 to (6.8), one obtains the entropy inequality

$$-\int |A(f, g)|^2 \log |A(f, g)|^2 \geq 1 \quad \text{whenever } \|f\|_2 \|g\|_2 = 1. \tag{6.9}$$

In a somewhat different form (more general in one respect and less so in another), this had been conjectured by Wehrl and proved by Lieb in an earlier paper [77], where one can also find a discussion of its physical interpretation. As was first shown by Grabowski [43], one can deduce Heisenberg's inequality from (6.9). The simplest way is to translate (6.9) into an inequality for $W(f, g)$ via (6.2), set $g = f$, and apply Theorem 5.1 to obtain (6.4).

In quantum statistical mechanics one considers not only "pure states" defined by unit vectors $f \in L^2$ but also "mixed states" defined by positive trace-class operators T of trace 1, that is, operators of the form $T = \sum c_j P_{f_j}$ where $\{f_j\}$ is an orthonormal sequence, P_{f_j} is the orthogonal projection onto $\mathbb{C}f_j$, $c_j > 0$, and $\sum c_j = 1$. (The pure states are those for which the sequence $\{f_j\}$ has only a single term f.) The Wigner distribution of such an operator T is defined to be $W(T) = \sum c_j W(f_j, f_j)$. This again resembles a probability distribution function except that it can have negative values. Assuming that its second moments are finite, one can consider its mean and covariance

$$M = \int zW(T)(z)\,dz, \qquad V_{jk} = \int \zeta_j \zeta_k W(T)(z)\,dz \qquad (z = (\xi, x),\ \zeta = z - M).$$

Narcowich [85] has made an interesting study of the uncertainty principle in terms of the matrix V. He characterizes those real positive definite matrices V that are covariances of Wigner distributions, gives a symplectically invariant formulation of Heisenberg's inequality for the state T in terms of invariants of V, and interprets it in terms of the geometry of the quadratic form defined by V.

7. Qualitative Uncertainty Principles

By a "qualitative uncertainty principle" we mean a theorem that, without giving quantitative estimates for f and $\widehat{f}$, says f and $\widehat{f}$ cannot both be too localized unless $f = 0$. Here "too localized" can be taken in several senses, of which we shall focus on two: restrictions on the sets where f and $\widehat{f}$ are nonzero, and bounds on the rate of decay of f and $\widehat{f}$ at infinity.

Our first group of results concerns the sets

$$\Sigma(f) = \{x : f(x) \neq 0\} \qquad \text{and} \qquad \Sigma(\widehat{f}) = \{\xi : \widehat{f}(\xi) \neq 0\}.$$

The first simple result, valid on any locally compact Abelian group, is that

$$0 \neq f \in L^2 \implies |\Sigma(f)|\,|\Sigma(\widehat{f})| \geq 1. \tag{7.1}$$

This follows immediately from (4.3) by taking $E = \Sigma(\widehat{f})$. Equation (7.1) was first derived by Matolcsi and Szücs [80]; it has been generalized to commutative hypergroups by Kumar [70]. On $\mathbb{R}^n$, however, something much stronger is true.

Theorem 7.2.

If $f \in L^1(\mathbb{R}^n)$ and $|\Sigma(f)|\,|\Sigma(\widehat{f})| < \infty$, then $f = 0$.

(Note that if $f \in L^p$ $(p > 1)$ and $|\Sigma(f)| < \infty$, then $f \in L^1$ and that if $f \in L^1$ and $|\Sigma(\widehat{f})| < \infty$, then $f \in L^p$ for all $p > 1$; hence the theorem applies equally to L^p functions.) This theorem is due to Benedicks [16], whose elegant proof, first circulated as a preprint in 1974 but not formally published for another decade, we reproduce below. It relies on the following form of the Poisson summation formula, the proof of which is an amusing exercise (or see Benedetto, Heil, and Walnut [13]).

Lemma 7.3.

If $f \in L^1(\mathbb{R}^n)$, the series $\phi(x) = \sum_{k\in\mathbb{Z}^n} f(x+k)$ converges in $L^1(\mathbb{T}^n)$, and the Fourier series of ϕ is $\sum_{k\in\mathbb{Z}^n} \widehat{f}(k)e^{2\pi i k\cdot x}$.

To prove Theorem 7.2, we may assume that $|\Sigma(f)| < 1$ by composing f with a dilation. We have

$$\int_{[0,1]^n} \sum_{k\in\mathbb{Z}^n} \chi_{\Sigma(\widehat{f})}(\xi + k)\, d\xi = \int_{\mathbb{R}^n} \chi_{\Sigma(\widehat{f})}(\xi)\, d\xi = \left|\Sigma(\widehat{f})\right| < \infty,$$

$$\int_{[0,1]^n} \sum_{k\in\mathbb{Z}^n} \chi_{\Sigma(f)}(x + k)\, dx = \int_{\mathbb{R}^n} \chi_{\Sigma(f)}(x)\, dx = |\Sigma(f)| < 1.$$

These inequalities imply, respectively, that

i. There exists $E \subset [0,1]^n$ with $|E| = 1$ such that $\sum \chi_{\Sigma(\widehat{f})}(a+k) < \infty$ for $a \in E$, and hence $\widehat{f}(a+k) \neq 0$ for only finitely many k if $a \in E$;

ii. There exists $F \subset [0,1]^n$ with $|F| > 0$ such that $\sum \chi_{\Sigma(f)}(x+k) = 0$ for $x \in F$, and hence $f(x+k) = 0$ for all k if $x \in F$.

Given $a \in E$, let

$$\phi_a(x) = \sum_{k\in\mathbb{Z}^n} f(x+k)e^{-2\pi i a(x+k)}.$$

By Lemma 7.3, $\phi_a \in L^1(\mathbb{T}^n)$ and the Fourier series of ϕ_a is $\sum \widehat{f}(a+k)e^{2\pi i k\cdot x}$. Since $a \in E$, ϕ_a is a trigonometric polynomial. In particular, ϕ_a is analytic, so either $\phi_a = 0$ or $\{x : \phi_a(x) = 0\}$ intersects every line in a discrete set and hence $\phi_a \neq 0$ a.e. On the other hand, $|\phi_a(x)| \leq \sum |f(x+k)| = 0$ for $x \in F$. We conclude that $\phi_a = 0$ for all $a \in E$, whence $\widehat{f}(a+k) = 0$ for all $a \in E$ and $k \in \mathbb{Z}$. In other words, $\widehat{f} = 0$ a.e., so $f = 0$.

Amrein and Berthier [2] have given a different proof of Theorem 7.2, and their methods also yield the following complementary result: If E and F are sets of finite measure in $\mathbb{R}^n$, the space of all $f \in L^2 \cap L^\infty$ such that $f = 0$ on E and $\widehat{f} = 0$ on F is infinite-dimensional. See also Busch [22].

There is a large literature on the existence or nonexistence of functions f on $\mathbb{R}$ or $\mathbb{T}$ subject to various restrictions on $\Sigma(f)$ and $\Sigma(\widehat{f})$. We refer the reader to Benedicks [15], Havin and Jöricke [48], and Benedetto [9] for a fuller discussion. Here we shall just mention a few results related to Theorem 7.2.

1. If $0 \neq f \in L^1(\mathbb{R})$ and $\Sigma(f)$ is bounded, then $\widehat{f}$ is the restriction of an entire function on $\mathbb{C}$, so $\mathbb{R} \setminus \Sigma(\widehat{f})$ is a countable discrete set. Moreover, by the Whittaker–Shannon sampling theorem (see Dym and McKean [32, p. 129]; Benedetto [11]; or Benedetto, Heil, and Walnut [13]), $\mathbb{R} \setminus \Sigma(\widehat{f})$ cannot contain any complete arithmetic progression $\xi_0 + b\mathbb{Z}$ with $b \leq [\mathrm{diam}(\Sigma(f))]^{-1}$.
2. (Kargaev [62] and Kargaev and Volberg [63]) There exists a set $E \subset \mathbb{R}$ of positive finite measure, such that $\widehat{\chi}_E$ vanishes on an interval, and a function $f \in L^1(\mathbb{R})$, such that $|\Sigma(f)| < \infty$ and $|\mathbb{R} \setminus \Sigma(\widehat{f})| = \infty$.
3. Melenk and Zimmerman [81] have recently given an explicit elementary construction of an infinite-dimensional family of functions $f \in L^1 \cap L^\infty \cap C^\infty$ on $\mathbb{R}$ such that both f and $\widehat{f}$ vanish on sets of the form $[-a, a] + 8a\mathbb{Z}$, where a can be specified independently for f and $\widehat{f}$.

It should be noted that Theorem 7.2 does not extend to distributions. Indeed, in the language of distributions Lemma 7.3 says that the periodic delta-function $\sum_{k\in\mathbb{Z}^n} \delta(x-k)$ is its own Fourier transform, and its support $\mathbb{Z}^n$ has measure zero.

It is natural to conjecture the following variant of Theorem 7.2 relating to Wigner distributions: *If* $|\Sigma(W(f,f))| < \infty$, *then* $f = 0$. As far as we know, this is an open question, but the following partial results are available. First, from (6.3) and Theorem 7.2 it is clear that $\Sigma(W(f,f))$ cannot be bounded unless $f = 0$. Second, if f is either even or odd, then $W(f,f)$ is its own Fourier transform up to a linear change of variable by (6.2); hence in these cases the conjecture follows by applying Theorem 7.2 to $W(f, f)$. (This is an unpublished remark of D. Mustard.)

Theorem 7.2 can be generalized to many other locally compact groups. A little experimentation with examples suggests the following formulation of a qualitative uncertainty principle for a Plancherel group G. For $E \subset G$ and $F \subset \widehat{G}$ let $|E|$ and $|F|$ denote the Haar measure of E and the Plancherel measure of F, respectively.

$$\text{Suppose } f \in L^1(G). \text{ If } |\Sigma(f)| < |G| \text{ and } |\Sigma(\widehat{f})| < |\widehat{G}|, \text{ then } f = 0. \tag{7.4}$$

If $G = \mathbb{T}^n$, (7.4) simply says that a nonzero trigonometric polynomial cannot vanish on a set of positive measure, a fact that we have already noted in the course of proving Theorem 7.2. The same reasoning shows that (7.4) is valid when G is any connected compact Lie group. The following additional results are known.

1. (Hogan [57]) Suppose G is infinite and compact. Then (7.4) holds if and only if G is connected. **Corollary:** *If G is a discrete Abelian group, (7.4) holds precisely when $\widehat{G}$ is connected or, equivalently, when G is torsion-free.*
2. (Hogan [55], [57]) If G is Abelian, noncompact, and nondiscrete, (7.4) holds precisely when the identity component of G is noncompact. Hogan [57] also has an extension of this result that applies to certain non-Abelian groups, and Kumar [70] has generalized it to certain commutative hypergroups (but see Voit [117, Remark 2.4] for comments on Kumar's hypotheses).
3. (Price and Sitaram [95] and Sitaram, Sundari, and Thangavelu [101]) Assertion (7.4) is valid for the Heisenberg group H_n, where it can actually be strengthened in several ways.
4. (Cowling, Price, and Sitaram [27]) Assertion (7.4) is valid when G is a connected, noncompact, semisimple Lie group with finite center, provided the condition $|\Sigma(f)| < |G| \,(= \infty)$ is replaced by $|K\Sigma(f)K| < \infty$, K being the maximal compact subgroup of G.
5. (Echterhoff, Kaniuth, and Kumar [33]) If G has a noncompact, nondiscrete normal subgroup H such that (7.4) holds for H and G/H is compact, then (7.4) holds for G. In particular, (7.4) holds for the group of rigid motions of $\mathbb{R}^n$ and for $\mathbb{R}^n \times K$ where K is compact. (See Price and Sitaram [95] for some variants of the latter results.) [33] also contains a number of other related theorems. Note, however, that the uncertainty principle considered throughout [33] is not (7.4) but the assertion that if $|\Sigma(f)|\,|\Sigma(\widehat{f})| < \infty$, then $f = 0$; this excludes compact groups and discrete Abelian groups from consideration.
6. (Meshulam [82]) Suppose G is a finite group. If f is a function on G let $|\Sigma(f)|$ denote the cardinality of $\Sigma(f)$ and $R(f)$ the rank of the convolution operator $g \to f * g$. Then $|\Sigma(f)|R(f) \geq |G|$ unless $f = 0$. If $f(1) = 1$, $|\Sigma(f)|R(f) = |G|$ if and only if $H = \Sigma(f)$ is a subgroup of G and $f|H$ is a one-dimensional character of H. Note that if G is Abelian, then $R(f)$ is the cardinality of $\Sigma(\widehat{f})$; the result in this case is due to Donoho and Stark [31].

It should be emphasized that (7.4) is false for many disconnected groups. The following result, while not of maximum generality, covers most of the interesting cases.

Theorem 7.5.
If G has a normal compact open subgroup H, not equal to G or $\{1\}$, such that G/H is either Abelian or finite, then (7.4) *is not valid.*

Proof. (See Folland [38] for the necessary background.) G/H is discrete, so $(G/H)\widehat{\ }$ is either a compact Abelian group or a finite set, and it sits inside $\widehat{G}$ as the set of irreducible representations of G that are trivial on H. If $\pi \in \widehat{G}$, the Schur orthogonality relations easily imply that $\widehat{\chi}_H(\pi) = \int_H \pi(x)\,dx = |H|P$, where P is the orthogonal projection onto the space

$$\mathcal{H}_\pi^H = \{v \in \mathcal{H}_\pi : \pi(h)v = v \text{ for all } h \in H\}.$$

But since $\pi(h)\pi(x) = \pi(x)\pi(x^{-1}hx)$ and H is normal, $\mathcal{H}_\pi^H$ is invariant under π. Since π is irreducible, $\mathcal{H}_\pi^H$ is either $\mathcal{H}_\pi$ (which means that $\pi \in (G/H)\widehat{}$) or $\{0\}$.

Thus, $\Sigma(\widehat{\chi}_H) = (G/H)\widehat{}$. Since $H \neq \{1\}$, we have $|\widehat{G} \setminus (G/H)\widehat{}| > 0$, and $|(G/H)\widehat{}| < \infty$, so $|\Sigma(\widehat{\chi}_H)| < |\widehat{G}|$. On the other hand, since $H \neq G$, it follows that $|\Sigma(\chi_H)| = |H| < |G|$. □

We now turn to the results concerning the decay of f and $\widehat{f}$ at infinity. The prototype of these is the following theorem of Hardy [49].

Theorem 7.6.

For $a, b > 0$ let $E(a, b)$ be the space of all measurable functions f on $\mathbb{R}$ such that

$$|f(x)| \leq ce^{-a\pi x^2} \text{ and } |\widehat{f}(\xi)| \leq ce^{-b\pi\xi^2} \quad \text{for some } c > 0.$$

If $ab < 1$, then $\dim E(a, b) = \infty$; if $ab = 1$, then $E(a, b) = \mathbb{C}e^{-a\pi x^2}$; and if $ab > 1$, then $E(a, b) = \{0\}$.

We give a sketch of the proof; Dym and McKean [32, §3.2] is a good reference for the details. The rescaling $f(x) \to f(\lambda x)$ maps $E(a, b)$ onto $E(\lambda^2 a, \lambda^{-2} b)$, so we may assume $a = b$. First, if $a < 1$, the Hermite functions discussed in §3 all belong to $E(a, a)$. Next, if $f \in E(1, 1)$, the condition $|f(x)| \leq ce^{-\pi x^2}$ easily implies that $\widehat{f}$ extends to an entire function on $\mathbb{C}$ and satisfies $|\widehat{f}(z)| \leq c'e^{\pi|z|^2}$. Since also $|\widehat{f}(\xi)| \leq ce^{-\pi\xi^2}$ for ξ real, a Phragmén–Lindelöf argument allows one to conclude that $\widehat{f}(z) = Ce^{-\pi z^2}$ for some C and hence $f(x) = Ce^{-\pi x^2}$. Finally, if $a > 1$, then $E(a, a) \subset E(1, 1)$ and $e^{-\pi x^2} \notin E(a, a)$, so $E(a, a) = \{0\}$.

Cowling and Price [25] have obtained the following L^p complement for Theorem 7.6: Suppose $p, q \in [1, \infty]$ and $\min(p, q) < \infty$. If $\|e^{a\pi x^2} f\|_p + \|e^{b\pi\xi^2}\widehat{f}\|_q < \infty$ with $ab \geq 1$, then $f = 0$. (Again, the Hermite functions show the necessity of the condition $ab \geq 1$.)

The case $ab > 1$ of Theorem 7.6 and its L^p version is an easy corollary of the following elegant result of Beurling, whose proof, in the same spirit as that of Theorem 7.6, has been published by Hörmander [58]: For $f \in L^1(\mathbb{R})$,

$$\iint |f(x)\widehat{f}(\xi)|e^{2\pi|x\xi|}\,dx\,d\xi < \infty \qquad \Longrightarrow \qquad f = 0. \tag{7.7}$$

Sitaram, Sundari, and Thangavelu [101] have derived analogs of Theorem 7.6, using Theorem 7.6 itself as a tool, for $\mathbb{R}^n$ and the Heisenberg group H_n. For $\mathbb{R}^n$ the result is identical to Theorem 7.6 with x^2 and ξ^2 replaced by $|x|^2$ and $|\xi|^2$, and the proof consists of using the Radon transform to reduce to the one-dimensional case. For H_n, the result is as follows: Suppose f is a function on H_n such that $|f(p, q, z)| \leq g(p, q)e^{-a\pi z^2}$ and $\|\widehat{f}(\sigma_\lambda)\|_{HS} \leq Ce^{-b\pi\lambda^2}$, where $g \in (L^1 \cap L^2)(\mathbb{R}^{2n})$ and $\sigma_\lambda(p, q, z) = \sigma(p, \lambda q, \lambda z)$ with σ given by (2.5); then $f = 0$ provided $ab > 1$. Also, Pati et al. [88] have obtained an analogue of Theorem 7.6 for Hermite expansions on $\mathbb{R}^n$—namely, if f and its Hermite coefficients both decay very rapidly at infinity, then $f = 0$.

The crucial fact that allows the use of complex analysis to prove Theorem 7.6 is that the characters $e_\xi(x) = e^{2\pi i\xi x}$ of $\mathbb{R}$ can be analytically continued in ξ to give (nonunitary) characters with exponential growth in x. A similar phenomenon happens for irreducible representations of many noncompact non-Abelian groups G. More precisely, one may have families of unitary representations of G indexed by $\mathbb{R}^n$ that can be analytically continued to get a family of (nonunitary) representations indexed by $\mathbb{C}^n$ whose matrix elements satisfy certain growth estimates on G. If these representations suffice for the Plancherel formula, one can hope to obtain an analogue of Theorem 7.6. This has been done by Sundari [113] when G is the group of rigid motions of $\mathbb{R}^n$ and by Sitaram and Sundari

[100] when G is a connected semisimple Lie group with finite center and either (i) G has only one conjugacy class of Cartan subgroups or (ii) attention is restricted to right K-invariant functions.

The rapid decay of $\widehat{f}$ at infinity imposes restrictions not only on the decay of f at infinity but also on the local decay of f near a point. For example, if f is a function on $\mathbb{T}$ and $|\widehat{f}(k)| \leq ce^{-\epsilon|k|}$, then f is analytic and so cannot have a zero of infinite order. More sophisticated theorems of this sort for functions on $\mathbb{R}$ or $\mathbb{T}$ can be found in Havin and Jöricke [48]. Pati et al. [88] have derived a sort of hybrid of this result and Theorem 7.6 for eigenfunction expansions of elliptic operators with analytic coefficients on compact Riemannian manifolds.

Havin and Jöricke [48] contains many additional results concerning local or global decay conditions on f and $\widehat{f}$. We shall mention only one, a neat theorem of Nazarov [86] whose flavor is similar to (7.7) but which implies Theorem 7.2 (for $n = 1$) rather than Theorem 7.6: There is a constant $c > 0$ such that for all $A, B \subset \mathbb{R}$ of finite measure and all $f \in L^2(\mathbb{R})$,

$$\|f\|_2^2 \leq ce^{c|A||B|}\left[\int_{\mathbb{R}\setminus A}|f|^2 + \int_{\mathbb{R}\setminus B}|\widehat{f}|^2\right].$$

8. Theorems on Approximate Concentration

Despite their mathematical solidity, the results of the preceding section—with the exception of the simple-minded (7.1)—have little to say about physical phenomena because they are unstable under the small errors that inevitably arise in the correspondence between theory and experiment. After all, the world is full of signals that are synthesized from a finite band of frequencies and last for a finite length of time, no matter what Theorem 7.2 says, and one would like a mathematical theory that says something useful about such signals. Thus, we wish to consider functions f on $\mathbb{R}^n$ and sets $A, B \subset \mathbb{R}^n$ such that f and $\widehat{f}$ are "negligibly small" on the complements of A and B, respectively; and we ask what sort of sets A and B allow functions with this behavior and what sort of functions they are. The uncertainty principle will be expressed as a restriction on the sizes of A and B.

Here is a very simple result of this sort, due to Williams [120]. Suppose f and $\widehat{f}$ are both in L^1; then

$$\int_A |f| \leq \|f\|_\infty |A| \leq \|\widehat{f}\|_1 |A|.$$

Multiplying this inequality by an analogous one with f and $\widehat{f}$ interchanged, we obtain

$$\frac{\int_A |f|}{\int |f|}\frac{\int_B |\widehat{f}|}{\int |\widehat{f}|} \leq |A|\,|B|. \tag{8.1}$$

Note that this gives another proof of (7.1).

More interesting, however, are the results relating to L^2 norms. If G is any locally compact Abelian group, $f \in L^2(G)$, $A \subset G$, and $\epsilon > 0$, we shall say that f is *ϵ-concentrated on A* if $\int_{G\setminus A}|f|^2 \leq \epsilon^2 \int_G |f|^2$, and we wish to know what can be said about f, $A \subset G$, and $B \subset \widehat{G}$ if f is ϵ-concentrated on A and $\widehat{f}$ is δ-concentrated on B. This problem, for $G = \mathbb{R}^n$, was first discussed in a lecture by Fuchs [39] at the 1954 International Congress. Landau, Pollak, and Slepian then made a detailed study of the case where $G = \mathbb{R}$ and A and B are intervals; we shall discuss their work below. However, the simplest and most general results, valid on any locally compact Abelian group, are more recent, and we shall begin by discussing them.

Almost everyone who has worked on this problem has relied on the interplay between the orthogonal projections P_A and Q_B on $L^2(G)$, defined for $A \subset G$ and $B \subset \widehat{G}$ by

$$P_A f = f\chi_A, \qquad (Q_B f)\widehat{\ } = \widehat{f}\chi_B.$$

The basic facts are summarized in the following theorem.

Theorem 8.2.

Suppose $A \subset G$ and $B \subset \widehat{G}$ have finite measure.

a. *$P_A Q_B$ is bounded from $L^p(G)$ to $L^q(G)$ for $1 \le p \le 2$ and $q \ge 1$, and $\|P_A Q_B f\|_q \le |A|^{1/q}|B|^{1/p}\|f\|_p$.*

b. *$P_A Q_B$ is a Hilbert–Schmidt operator on $L^2(G)$, and $\|P_A Q_B\|_{\mathrm{HS}} = |A|^{1/2}|B|^{1/2}$.*

c. *If there is a nonzero $f \in L^2(G)$ such that f is ϵ-concentrated on A and $\widehat{f}$ is δ-concentrated on B, then $1 - \epsilon - \delta \le \|P_A Q_B\|$, where $\|P_A Q_B\|$ is the norm of $P_A Q_B$ as an operator on L^2.*

Proof. Since $Q_B f = f * \mathcal{F}^{-1}\chi_B$, $P_A Q_B$ is an integral operator:

$$P_A Q_B f(x) = \int K(x,y) f(y)\,dy, \qquad K(x,y) = \chi_A(x)\widehat{\chi}_B(x-y).$$

Hence, by the Hölder and Hausdorff–Young inequalities, if $1 \le p \le 2$ and $p' = p/(p-1)$,

$$|P_A Q_B f(x)| \le \chi_A(x)\|\widehat{\chi}_B\|_{p'}\|f\|_p \le \chi_A(x)\|\chi_B\|_p\|f\|_p = \chi_A(x)|B|^{1/p}\|f\|_p.$$

Part **a** follows by taking the L^q norm of both sides. Moreover,

$$\|P_A Q_B\|_{\mathrm{HS}}^2 = \iint |K(x,y)|^2\,dx\,dy = \int |\chi_B|^2 \int |\chi_A|^2 = |B||A|,$$

which proves part **b**. Finally, suppose f is ϵ-concentrated on A and $\widehat{f}$ is δ-concentrated on B, and $\|f\|_2 = 1$. Since $\|P_A(f - Q_B f)\|_2 \le \|f - Q_B f\|_2 = \|(1-\chi_B)\widehat{f}\|_2 \le \delta$, we have

$$1 - \epsilon - \delta \le \|f\|_2 - \|f - P_A f\|_2 - \|P_A(f - Q_B f)\|_2 \le \|P_A Q_B f\|_2 \le \|P_A Q_B\|,$$

which proves part **c**. □

Corollary 8.3.

If $f \ne 0$ is ϵ-concentrated on A and $\widehat{f}$ is δ-concentrated on B, then $|A|\,|B| \ge (1 - \epsilon - \delta)^2$.

Proof. Combine parts **a** and **c** of the theorem. □

Theorem 8.2 and Corollary 8.3 were proved by Donoho and Stark [31] for $G = \mathbb{R}$ or $G = \mathbb{Z}/n\mathbb{Z}$ and generalized by Smith [108] to arbitary locally compact Abelian groups. These papers also contain an analog of Theorem 8.2**c** (in a slightly weakened form) for L^p norms ($1 \le p \le 2$) and a discussion of the sharpness (or lack thereof) of the estimate in Theorem 8.2**a**. In addition, Donoho and Stark [31] give some interesting applications to problems in signal analysis.

Wolf [121] has extended Theorem 8.2 to Gelfand pairs (that is, to K-biinvariant functions on a locally compact group G, where K is a compact subgroup of G such that convolution of K-biinvariant functions is commutative), and Voit [117] has extended it even further to commutative hypergroups. Also, de Jeu [30] has proved a version of the L^2 part of Theorem 8.2 that concerns integral operators on abstract measure spaces possessing some of the features of the Fourier transform, and Koppinen [68] has obtained results analogous to Theorem 8.2 in the setting of Hopf algebras.

The quantity $\|P_A Q_B\|$ that intervenes decisively in Theorem 8.2 has an interesting geometric interpretation: it is the cosine of the angle between the ranges of P_A and Q_B. Indeed, we have

$$\begin{aligned}\|P_A Q_B\| &= \sup\bigl\{|\langle P_A Q_B f, g\rangle| : \|f\|_2 = \|g\|_2 = 1\bigr\} \\ &= \sup\bigl\{|\langle Q_B f, P_A g\rangle| : \|f\|_2 = \|g\|_2 = 1\bigr\} \\ &= \sup\bigl\{\mathrm{Re}\langle u, v\rangle : \|u\|_2 = \|v\|_2 = 1,\ Q_B u = u,\ P_A v = v\bigr\},\end{aligned}$$

and $\mathrm{Re}\langle u, v\rangle$ is the cosine of the angle between the unit vectors u and v.

A deeper analysis of the approximate concentration problem can be achieved by studying the operator

$$(P_A Q_B)^* P_A Q_B = Q_B P_A Q_B.$$

This was done by Landau, Pollak, Slepian, and Widom in a remarkable series of papers [73–75, 103–105, 107], parts of which we now describe briefly. Expositions of this work, with references to related papers, can also be found in Landau [72] and Slepian [106]. In what follows we shall assume that $G = \widehat{G} = \mathbb{R}^n$ to simplify the discussion, although some of the results are actually more general, and A and B will always denote sets of positive finite measure.

$Q_B P_A Q_B$ has the advantage of being selfadjoint and positive, and by Theorem 8.2b it is compact, in fact trace-class. Hence it has an orthonormal eigenbasis, and the nonzero eigenvalues are positive, of finite multiplicity, and accumulate only at 0. Let $\{\lambda_k\}_1^\infty$ be the nonzero eigenvalues, listed with multiplicity in decreasing order ($\lambda_1 \geq \lambda_2 \geq \cdots$), and let $\{\psi_k\}_1^\infty$ be a corresponding orthonormal set of eigenfunctions. We then have

$$\lambda_1 = \|Q_B P_A Q_B\| = \|P_A Q_B\|^2,$$

so $\sqrt{\lambda_1}$ is the cosine of the angle betwen the ranges of P_A and Q_B. The crucial fact is the following.

Theorem 8.4.
$\lambda_1 < 1$.

Proof. Clearly $\lambda_1 = \|Q_B P_A Q_B\| \leq 1$. If $\lambda_1 = 1$, there exists $f \neq 0$ such that $Q_B P_A Q_B f = f$. Thus f is in the range of Q_B, and it is also in the range of P_A because $\|P_A g\|_2 < \|g\|_2$ unless $g \in \text{range}(P_A)$. But this is impossible by Theorem 7.2. □

Suppose now that $\|f\|_2 = 1$, f is ϵ-concentrated on A, and $\widehat{f}$ is δ-concentrated on B. The angle between f and $P_A f$ is

$$\arccos \frac{\langle f, P_A f\rangle}{\|P_A f\|_2} = \arccos \|P_A f\|_2 \leq \arccos \sqrt{1-\epsilon^2},$$

and likewise the angle between f and $P_B f$ is at most $\arccos\sqrt{1-\delta^2}$. The angle between $P_A f$ and $P_B f$ is, on the one hand, at most the sum of these two angles, and on the other, at least $\arccos\sqrt{\lambda_1}$. Thus,

$$\arccos\sqrt{1-\epsilon^2} + \arccos\sqrt{1-\delta^2} \geq \arccos\sqrt{\lambda_1}.$$

In fact, by taking suitable linear combinations of the eigenfunctions ψ_k, one can construct examples where f and $\widehat{f}$ have any desired concentrations on A and B subject to this restriction, and one arrives at the following theorem.

Theorem 8.5.
Suppose $0 \leq \alpha, \beta \leq 1$ *and* $(\alpha, \beta) \neq (1, 0)$ *or* $(0, 1)$. *There is a function* $f \in L^2(\mathbb{R}^n)$ *with* $\|f\|_2 = 1$, $\|P_A f\|_2 = \alpha$, *and* $\|Q_B f\|_2 = \beta$ *if and only if*

$$\arccos\alpha + \arccos\beta \geq \arccos\sqrt{\lambda_1} = \arccos\|P_A Q_B\|. \tag{8.6}$$

The full proof can be found in Landau and Pollak [73] or Dym and McKean [32, §2.9]; these authors state the result for $G = \mathbb{R}$ and A and B intervals, but the arguments are quite general. (If A is bounded, the pair $(\alpha, \beta) = (1, 0)$ is not admissible, for if $f = P_A f$, then $\widehat{f}$ is analytic and so cannot vanish on B; likewise if B is bounded, then $(\alpha, \beta) = (0, 1)$ is not admissible. But Kargaev's example [62, 63] (see §7) shows that the boundedness assumption is necessary.) Another version of the uncertainty inequality (8.6) has been proved by Benedetto [8]; the Logvinenko–Sereda–Kacnel'son theorem quoted at the end of §4 is also of interest here.

The eigenfunctions ψ_k have a number of interesting and pleasing properties. For example, assuming A and B are bounded, $\{\psi_k\}_1^\infty$, $\{\lambda_k^{-1/2} P_A \psi_k\}_1^\infty$, and $\{\widehat{\psi}_k\}_1^\infty$ are (respectively) orthonormal bases for the range of Q_B, the range of P_A, and the range of P_B consisting of eigenfunctions for $Q_B P_A Q_B$, $P_A Q_B P_A$, and $P_B Q_{-A} P_B$. (The reader is invited to work out the rather easy proofs of these facts.) For our purposes, however, the crucial thing is that $\widehat{\psi}_k$ is 0-concentrated on B because $\psi_k = \lambda_k^{-1} Q_B P_A Q_B \psi_k \in \text{range}(Q_B)$ and ψ_k is $(1 - \sqrt{\lambda_k})$-concentrated on A because

$$\|P_A \psi_k\|_2^2 = \|P_A Q_B \psi_k\|_2^2 = \langle Q_B P_A Q_B \psi_k, \psi_k \rangle = \lambda_k \|\psi_k\|_2^2.$$

It follows that one can obtain functions f such that f and $\widehat{f}$ are well concentrated on A and B by taking linear combinations of the ψ_k for which λ_k is close to 1; therefore, the situation calls for an analysis of the eigenvalues λ_k.

The finest results in this direction are those for the case $n = 1$, $A = (-\frac{1}{2}T, \frac{1}{2}T)$ (the factor of $\frac{1}{2}$ is traditional), and $B = (-\Omega, \Omega)$. By rescaling one can reduce to the case $\Omega = 1$, and it follows that the eigenvalues λ_k depend only on the product ΩT. We cite two major theorems; others can be found in Landau and Pollak [74].

First, it was conjectured by Slepian and proved by Landau and Widom [75] that if $N(\Omega T, \alpha)$ denotes the number of eigenvalues λ_k that exceed α $(0 < \alpha < 1)$, then

$$N(\Omega T, \alpha) = 2\Omega T + \left[\frac{1}{\pi^2} \log \frac{1-\alpha}{\alpha}\right] \log \Omega T + o(\log \Omega T).$$

Thus if $\Omega T \gg 1$, λ_k is very close to 1 for $k \ll 2\Omega T$ and very close to zero for $k \gg 2\Omega T$, and the transition from large to small takes place over an interval of length $O(\log \Omega T)$.

Second, if $\|f\|_2 = 1$, f is ϵ-concentrated on $(-\frac{1}{2}T, \frac{1}{2}T)$, and $\widehat{f}$ is 0-concentrated on $(-\Omega, \Omega)$, then

$$\left\| f - \sum_1^{[2\Omega T]} \langle f, \psi_k \rangle \psi_k \right\|_2^2 \leq 12\epsilon^2.$$

Moreover, for any $\eta > 0$ there exists $C > 0$ such that

$$\left\| f - \sum_1^{[2\Omega T]+C \log \Omega T} \langle f, \psi_k \rangle \psi_k \right\|_2^2 \leq (1+\eta)\epsilon^2.$$

Similar results hold if $\widehat{f}$ is merely δ-concentrated on $(-\Omega, \Omega)$; the proof can be found in Landau and Pollak [74].

These results give substance and precision to the folk wisdom that there are about $2\Omega T$ degrees of freedom in a signal of duration T constructed from frequencies of magnitude $\leq \Omega$. Another variation on the same theme can be found in Slepian [104].

The eigenfunctions ψ_k for the case $A = (-\frac{1}{2}T, \frac{1}{2}T)$ and $B = (-\Omega, \Omega)$ are well-known special functions. By the change of variable $r = 2x/T$ we can assume that $T = 2$, and in this case it turns out that $Q_B P_A Q_B$ commutes with the differential operator

$$L_\Omega = (1 - r^2)\frac{d^2}{dr^2} - 2r\frac{d}{dr} - \Omega^2 r^2. \tag{8.7}$$

The eigenfunctions ψ_k are therefore also eigenfunctions of L_Ω, and the corresponding eigenvalues μ_k are singled out as the only values of μ for which the equation $L_\Omega u = \mu u$ has a solution that is continuous at both $x = 1$ and $x = -1$. Since the operator L arises from the Laplacian in $\mathbb{R}^3$ by separation of variables in ellipsoidal coordinates, the functions ψ_k have been saddled with the ungainly name of "prolate spheroidal wave functions," and they have been studied rather extensively. The papers cited above contain more details and references; here we shall just mention one recent result. Let $B^p = \{f \in L^p(\mathbb{R}) : \operatorname{supp} \widehat{f} \subset [-\Omega, \Omega]\}$. As noted above, $\{\psi_k\}$ is an orthonormal basis for B^2; Barceló and Córdoba [5] have proved that it is a basis for B^p if and only if $\frac{4}{3} < p < 4$.

Results similar to the one-dimensional theory have been obtained by Slepian [103] for the case where A and B are balls centered at the origin in $\mathbb{R}^n$. When A is the ball of radius 1 and B is the ball of radius Ω, the eigenfunctions ψ_k are products of spherical harmonics of degree k with functions of $r = |x|$ that are eigenfunctions of $L_\Omega + (\frac{1}{4} - (k + \frac{1}{2}n - 1)^2)r^{-2}$, where L_Ω is given by (8.7).

Analogues of the Landau–Pollak–Slepian theory have been developed in several other settings involving Fourier-type expansions: for the groups $G = \mathbb{Z}/n\mathbb{Z}$ by Pearl [89] and Grünbaum [46]; for the Walsh–Paley group by Pearl [89], for $G = \mathbb{T}$ and $G = \mathbb{Z}$ by Slepian [105], for orthogonal polynomial expansions by Perlstadt [90], and for some situations involving non-Abelian groups (rotation groups, spheres, and hyperbolic spaces) by Grünbaum, Longhi, and Perlstadt [47]. See also Landau [72] and Slepian [106] for references to other related work.

9. Minimal Rectangles in Phase Space

Suppose $f \in L^2(\mathbb{R})$. If f is concentrated (in some sense) on an interval I and $\widehat{f}$ is concentrated on an interval J, we shall think of f as "occupying" the rectangle $I \times J$ in (x, ξ)-space, or phase space. (One can interpret "concentration" as in §8, or with $\|f\|_2 = 1$ one can take $I = [M - \sigma, M + \sigma]$ where $M = M(|f|^2)$ and $\sigma = \sqrt{V(|f|^2)}$ and similarly for J; other variants of this idea are also possible.) The results of the preceding sections give several ways of making precise the vague assertion that in this case, $|I \times J| = |I|\,|J|$ must be at least on the order of magnitude of unity. Likewise, if f is a function on $\mathbb{R}^n$ that is concentrated on a rectangular box $I = \prod_1^n I_k$, where each I_k is an interval in $\mathbb{R}$, and $\widehat{f}$ is concentrated on another such box $J = \prod_1^n J_k$, Heisenberg's inequality in the form of Corollary 2.6 indicates (roughly) that $|I_k|\,|J_k| \geq 1$ for all k. These considerations suggest the following heuristic form of the uncertainty principle: *The smallest significant regions in phase space are sets of the form*

$$\prod_1^n I_k \times \prod_1^n J_k, \qquad |I_k|\,|J_k| = 1 \text{ for all } k.$$

We shall call such a set a *minimal rectangle*. This section is devoted to a brief discussion of some interesting phenomena that can be understood in terms of this principle.

First, phase space is the stage for *microlocal analysis*, a body of techniques developed in the past thirty years for studying local behavior of partial differential equations and generalizations thereof. In this regard, Fefferman and Phong [35, 36] have proved a number of deep theorems concerning boundedness, positivity, and eigenvalue estimates for differential and pseudodifferential operators that are based on the following principle: The size of an open set S in phase space should be measured not by its volume but by the maximum number of minimal rectangles, or images of such under canonical transformations, that can be fitted inside S without overlapping. For example, suppose $L = \sigma(x, D)$ is a selfadjoint differential operator with symbol $\sigma(x, \xi) = \sum_{|\alpha| \leq k} a_\alpha(x)\xi^\alpha$ that is real and bounded below. A classical rule of thumb says that the number of eigenvalues of L less than some constant C is roughly equal to the volume of $S_C = \{(x, \xi) : \sigma(x, \xi) < C\}$; but one obtains better estimates for the eigenvalues by counting minimal rectangles inside S_C. We refer the reader to the introductions of [35] and [36] for more details.

The other matter we wish to discuss is the problem of constructing interesting bases for L^2 (preferably, but not necessarily, orthonormal) whose elements and their Fourier transforms are well localized. We shall restrict attention to the dimension $n = 1$. The idea is the following: Suppose $\{I_k \times J_k\}$ is a tiling of the phase plane by minimal rectangles; we would like to find a basis $\{\phi_k\}$ for $L^2(\mathbb{R})$ such that ϕ_k occupies the box $I_k \times J_k$ in the sense described above.

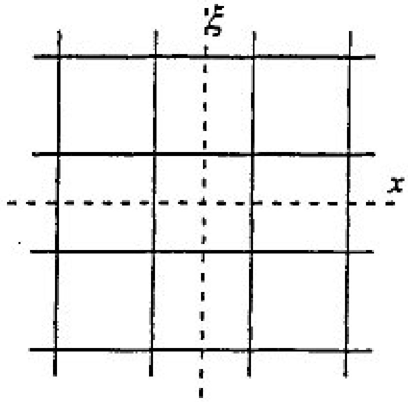

FIGURE 1.

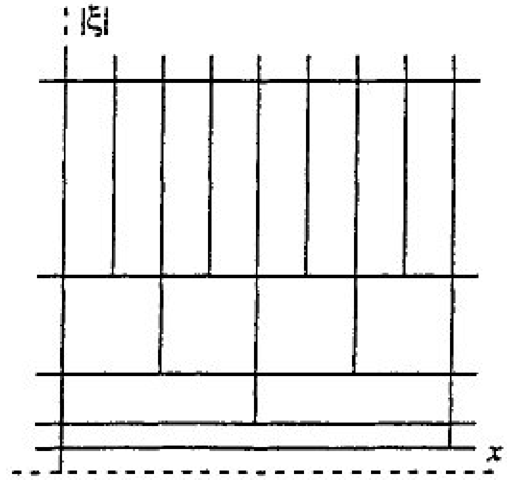

FIGURE 2.

The simplest such tiling is the set of all squares with vertices in the lattice $(\mathbb{Z}+\frac{1}{2})^2$, as shown in Figure 1, and the simplest way to try to find a basis associated to this tiling is to take a function ϕ that occupies the center square $[-\frac{1}{2},\ \frac{1}{2}]^2$ and translate it:

$$\phi_{jk}(x) = e^{2\pi ikx}\phi(x-j). \tag{9.1}$$

ϕ_{jk} thus occupies the square $[j-\frac{1}{2},\ j+\frac{1}{2}]\times[k-\frac{1}{2},\ k+\frac{1}{2}]$. This idea was proposed by Gabor [40], who argued that with $\phi(x)=e^{-\pi x^2}$, $\{\phi_{jk} : j, k\in\mathbb{Z}\}$ should be a good basis for L^2. Unfortunately, this turns out to be false. For this ϕ, $\{\phi_{jk}\}$ does span L^2, but it fails to be a *frame*; that is, it fails to satisfy an estimate of the form

$$C^{-1}\|f\|_2 \le \sum |\langle f, \phi_{jk}\rangle|^2 \le C\|f\|_2 \quad \text{for all } f\in L^2. \tag{9.2}$$

The trouble is not just an unfortunate choice of ϕ; in fact, if $\|\phi\|_2=1$ and $\{\phi_{jk} : j, k\in\mathbb{Z}\}$ satisfies (9.2), then $V(|\phi|^2)V(|\widehat{\phi}|^2)=\infty$, so ϕ cannot really occupy any finite box. This is an extended form of the Balian–Low theorem; see Benedetto, Heil, and Walnut [13]; Benedetto and Walnut [14]; or Daubechies [28, p. 108]. Another version of this result, that $\{\phi_{jk}\}$ cannot satisfy (9.2) if ϕ is continuous and $|\phi(x)|\le C(1+|x|)^{-1-\epsilon}$, is implicit in the arguments in [37, §3.4].

Another interesting tiling is shown in Figure 2, where the strip $2^j\le\xi\le 2^{j+1}$ is cut up into rectangles of width 2^{-j}. Here, if ψ occupies the box $[0, 1]\times[1, 2]$, one can manufacture functions to occupy all the other boxes by translating and dilating ψ:

$$\psi^{jk}(x) = 2^{j/2}\psi(2^jx-k).$$

In this situation it is indeed possible to find ψ's that are quite well localized in both x and ξ for which $\{\psi^{jk} : j, k\in\mathbb{Z}\}$ is an orthonormal basis for L^2. These are the *wavelets* that have received much attention in recent years; we refer the reader to Daubechies [28] and Strichartz [111] for accounts of their construction.

We have, however, cheated a little bit, as the vertical axis in Figure 2 denotes $|\xi|$ rather than ξ. It is, in fact, characteristic of wavelets that $\widehat{\psi}$ has two peaks, one in the region $\xi>0$ and one in the region $\xi<0$, so that ψ actually occupies the two rectangles $[0, 1]\times[1, 2]$ and $[0, 1]\times[-2, -1]$. (The fact that the region associated to ψ is twice as big as a minimal rectangle should not cause concern. The uneasy reader may compensate for it by compressing the $|\xi|$-axis in Figure 2 by a factor of 2, as Figure 2 is only a heuristic guide anyhow.)

It turns out that if one is willing to apply the same kind of fudging to Figure 1 by replacing the exponential in (9.1) by a sine or cosine, thereby replacing the frequency peak at $\xi = k$ by two peaks at $\xi = \pm\frac{1}{2}k$, one can get around the Balian–Low obstruction. More precisely, there exist orthonormal bases for L^2 of the form

$$\phi(x-j)\tau[(k+\tfrac{1}{2})\pi(x-j)] \qquad (j \in \mathbb{Z},\ k \in \mathbb{Z}^+),$$

where ϕ is a smooth approximation to the characteristic function of $[-\frac{1}{2}, \frac{1}{2}]$ and τ denotes either sine or cosine. The construction of these bases, discovered by Coifman and Meyer and by Malvar, is delightfully elementary and generalizes to produce orthonormal bases of L^2 associated to many other tilings of the phase plane by minimal rectangles—but always with the "two-peak" phenomenon. In fact, one can even construct wavelets this way. We refer the reader to Auscher, Weiss, and Wickerhauser [4] and Auscher [3] for a detailed treatment and a discussion of related constructions.

The papers of Bourgain [20] and Byrnes [23] give other constructions of orthonormal bases of $L^2(\mathbb{R})$ whose elements and their Fourier transforms satisfy uniform uncertainty estimates.

References

[1] Ali, S. T. and Prugovecki, E. (1977). Systems of imprimitivity and representations of quantum mechanics in fuzzy phase space. *J. Math. Phys.* **18**, 219–228. MR 58 #8964

[2] Amrein, W. O. and Berthier, A.M. (1977). On support properties of L^p functions and their Fourier transforms. *J. Funct. Anal.* **24**, 258–267. MR 57 #1013

[3] Auscher, P. (1994). Remarks on the local Fourier bases. *Wavelets: Mathematics and Applications* (J. J. Benedetto and M. W. Frazier, eds.). CRC Press, Boca Raton, FL. 203–218. MR 94i:42039

[4] Auscher, P., Weiss, G., and Wickerhauser, M. V. (1992). Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets. *Wavelets: A Tutorial in Theory and Applications* (C. K. Chui, ed.). Academic Press, Boston, 237–256. MR 93e:42042

[5] Barceló, J. A. and Córdoba, A. J. (1989). Band-limited functions: L^p convergence. *Trans. Amer. Math. Soc.* **313**, 655–669. MR 90g:42017

[6] Beckner, W. (1975). Inequalities in Fourier analysis. *Ann. of Math.* **102**, 159–182. MR 52 #6317

[7] ———. (1995). Pitt's inequality and the uncertainty principle. *Proc. Amer. Math. Soc.* **123**, 1897–1905. MR 95g:42021

[8] Benedetto, J. J. (1985). An inequality associated to the uncertainty principle. *Rend. Circ. Mat. Palermo* (2) **34**, 407–421. MR 87g:26024

[9] ———. (1985). Fourier uniqueness criteria and spectrum estimation theorems. *Fourier Techniques and Applications* (J. F. Price, ed.). Plenum, New York, 149–170.

[10] ———. (1990). Uncertainty principle inequalities and spectrum estimation. *Recent Advances in Fourier Analysis and Its Applications* (J. S. Byrnes and J. L. Byrnes, eds.). Kluwer Acad. Publ., Dordrecht. MR 91i:94010

[11] ———. (1994). Frame decompositions, sampling, and uncertainty principle inequalities. *Wavelets: Mathematics and Applications* (J. J. Benedetto and M. W. Frazier, eds.). CRC Press, Boca Raton, FL, 247–304. MR 94i:94005

[12] Benedetto, J. J. and Frazier, M. W. (eds.). (1994). *Wavelets: Mathematics and Applications*. CRC Press, Boca Raton, FL.

[13] Benedetto, J. J., Heil, C., and Walnut, D. (1992). Uncertainty principles for time-frequency operators. *Continuous and Discrete Fourier Transforms, Extension Problems, and Wiener-Hopf Equations* (I. Gohberg, ed.). Birkhäuser, Basel, 1–25. MR 94a:42041

[14] Benedetto, J. J. and Walnut, D. F. (1994). Gabor frames for L^2 and related spaces. *Wavelets: Mathematics and Applications* (J. J. Benedetto and M. W. Frazier, eds.). CRC Press, Boca Raton, FL, 97–162. MR 94i:42040

[15] Benedicks, M. (1984). The support of functions and distributions with a spectral gap. *Math. Scand.* **55**, 255–309. MR 86f:43006

[16] ———. (1985). On Fourier transforms of functions supported on sets of finite Lebesgue measure. *J. Math. Anal. Appl.* **106**, 180–183. MR 86j:42017

[17] Bialynicki-Birula, I. (1985). Entropic uncertainty relations in quantum mechanics. *Quantum Probability and Applications III* (L. Accardi and W. von Waldenfels, eds.). *Lecture Notes in Math.* **1136**. Springer, Berlin, 90–103. MR 87b:81007

[18] Bialynicki-Birula, I. and Mycielski, J. (1975). Uncertainty relations for information entropy in wave mechanics. *Comm. Math. Phys.* **44**, 129–132. MR 52 #7403

[19] Borchi, E. and Marsaglia, S. (1985). A note on the uncertainty product of bandlimited functions. *Signal Process.* **9**, 277–282. MR 87a:94007

[20] Bourgain, J. (1988). A remark on the uncertainty principle for Hilbertian bases. *J. Funct. Anal.* **79**, 136–143. MR 89f:81025

[21] Bowie, P. C. (1971). Uncertainty inequalities for Hankel transforms. *SIAM J. Math. Anal.* **2**, 601–606. MR 46 #4113

[22] Busch, P. (1984). On joint lower bounds of position and momentum observables in quantum mechanics. *J. Math. Phys.* **25**, 1794–1797. MR 85h:81004

[23] Byrnes, J. S. (1994). Quadrature mirror filters, low crest factor arrays, functions achieving optimal uncertainty principle bounds, and complete orthonormal sequences—a unified approach. *Appl. Comput. Harmonic Anal.* **1**, 261–266. MR 95i:94004

[24] Chistyakov, A. L. (1976). On uncertainty relations for vector-valued operators. *Teoret. Mat. Fiz.* **27**, 130–134; English transl., *Theor. Math. Phys.* **27**, 380–382. MR 56 #7615

[25] Cowling, M. G. and Price, J. F. (1983). Generalisations of Heisenberg's inequality. *Harmonic Analysis* (G. Mauceri, F. Ricci, and G. Weiss, eds.). *Lecture Notes in Math.* **992**. Springer, Berlin, 443–449. MR 86g:42002b

[26] ———. (1984). Bandwidth versus time concentration: the Heisenberg–Pauli–Weyl inequality. *SIAM J. Math. Anal.* **15**, 151–165. MR 86g:42002a

[27] Cowling, M. G., Price, J. F., and Sitaram, A. (1988). A qualitative uncertainty principle for semisimple Lie groups. *J. Austral. Math. Soc. Ser. A* **45**, 127–132. MR 89d:43005

[28] Daubechies, I. (1992). *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, PA. MR 93e:42045

[29] de Bruijn, N. G. (1967). Uncertainty principles in Fourier analysis. *Inequalities* (O. Shisha, ed.). Academic Press, New York, 55–71. MR 36 #3047

[30] de Jeu, M. F. E. (1994). An uncertainty principle for integral operators. *J. Funct. Anal.* **122**, 247–253. MR 95h:43009

[31] Donoho, D. L. and Stark, P. B. (1989). Uncertainty principles and signal recovery. *SIAM J. Appl. Math.* **49**, 906–931. MR 90c:42003

[32] Dym, H. and McKean, H. P. (1972). *Fourier Series and Integrals*. Academic Press, New York. MR 56 #945

[33] Echterhoff, S., Kaniuth, E., and Kumar, A. (1991). A qualitative uncertainty principle for certain locally compact groups. *Forum Math.* **3**, 355–369. MR 93a:43005

[34] Faris, W. G. (1978). Inequalities and uncertainty principles. *J. Math. Phys.* **19**, 461–466. MR 58 #4066

[35] Fefferman, C. (1983). The uncertainty principle. *Bull. Amer. Math. Soc.* (*N.S.*) **9**, 129–206. MR 85f:35001

[36] Fefferman, C. and Phong, D. H. (1981). The uncertainty principle and sharp Gårding inequalities. *Comm. Pure Appl. Math.* **34**, 285–331. MR 82j:35140

[37] Folland, G. B. (1989). *Harmonic Analysis in Phase Space*. Princeton University Press, Princeton, NJ. MR 92k:22017

[38] ———. (1995). *A Course in Abstract Harmonic Analysis*. CRC Press, Boca Raton, FL.

[39] Fuchs, W. H. J. (1954). On the magnitude of Fourier transforms. *Proc. Intern. Congress Math. 1954, vol. II*. North-Holland, Amsterdam.

[40] Gabor, D. (1946). Theory of communication. *J. Inst. Elec. Engr.* **93**, 429–457.

[41] Garofalo, N. and Lanconelli, E. (1990). Frequency functions on the Heisenberg group, the uncertainty principle and unique continuation. *Ann. Inst. Fourier* (*Grenoble*) **40**, 313–356. MR 91i:22014

[42] Gesztesy, F. and Pittner, L. (1978). Uncertainty relations and quadratic forms. *J. Phys. A* **11**, 1765–1770. MR 81a:81028

[43] Grabowski, M. (1984). Wehrl-Lieb's inequality for entropy and the uncertainty relation. *Rep. Math. Phys.* **20**, 153–155. MR 86f:82008

[44] Gross, L. (1975). Logarithmic Sobolev inequalities. *Amer. J. Math.* **97**, 1061–1083. MR 54 #8263

[45] ———. (1978). Logarithmic Sobolev inequalities—a survey. *Vector Space Measures and Applications* I (R. M. Aron and S. Dineen, eds.). *Lecture Notes in Math.* **644**. Springer, Berlin, 196–203. MR 80f:46035

[46] Grünbaum, F. A. (1981). Eigenvectors of a Toeplitz matrix: discrete version of the prolate spheroidal wave functions. *SIAM J. Alg. Disc. Methods* **2**, 136–141. MR 82m:15013

[47] Grünbaum, F. A., Longhi, L., and Perlstadt, M. (1982). Differential operators commuting with finite convolution integral operators: some nonabelian examples. *SIAM J. Appl. Math.* **42**, 941–955. MR 84f:43013

[48] Havin, V. and Jöricke, B. (1994). *The Uncertainty Principle in Harmonic Analysis*. Springer, Berlin. MR 96c:42001

[49] Hardy, G. H. (1933). A theorem concerning Fourier transforms. *J. London Math. Soc.* **8**, 227–231.

[50] Heinig, H. P. and Smith, M. (1986). Extensions of the Heisenberg-Weyl inequality. *Internat. J. Math. Math. Sci.* **9**, 185–192. MR 87i:26013

[51] Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematic und Mechanik. *Zeit. Physik* **43**, 172–198.

[52] Helffer, B. and Nourrigat, J. (1988). Remarques sur le principe d'incertitude. *J. Funct. Anal.* **80**, 33–46. MR 90b:35254

[53] Hilberg, W. and Rothe, F. G. (1971). The general uncertainty relation for real signals in communication theory. *Inform. and Control* **18**, 103–125. MR 43 #4540

[54] Hirschman, I. I. (1957). A note on entropy. *Amer. J. Math.* **79**, 152–156. MR 19, 622i

[55] Hogan, J. A. (1988). A qualitative uncertainty principle for locally compact Abelian groups. *Miniconferences on Harmonic Analysis and Operator Algebras (Proc. Centre Math. Anal.* 16; M. Cowling, C. Meaney, and W. Moran, eds.). Australian National University, Canberra, 133–142. MR 89k:43006

[56] ———. (1989). *Fourier Uncertainty on Groups*, Ph.D. Dissertation, University of New South Wales.

[57] Hogan, J. A. (1993). A qualitative uncertainty principle for unimodular groups of type I. *Trans. Amer. Math. Soc.* **340**, 587–594. MR 94b:43003

[58] Hörmander, L. (1991). A uniqueness theorem of Beurling for Fourier transform pairs. *Ark. Mat.* **29**, 237–240. MR 93b:42016

[59] Ishigaki, T. (1991). Four mathematical expressions of the uncertainty relation. *Found. Phys.* **21**, 1089–1105. MR 92k:81012

[60] Kacnel'son, V. E. (1973). Equivalent norms in spaces of entire functions. *Mat. Sb.* **92**, 34–54; English transl., *Math. USSR Sb.* **21**, 33–55. MR 49 #3171

[61] Kahane, J.-P., Lévy-Leblond, J.-M., and Sjöstrand, J. (1993). Une inegalité de Heisenberg inverse. *J. Anal. Math.* **60**, 135–155. MR 95a:34124

[62] Kargaev, P. P. (1982/1983). The Fourier transform of the characteristic function of a set, vanishing on an interval. *Mat. Sb.* **117**, 397–411; English transl. *Math. USSR Sb.* **45**, 397–410. MR 83f:42010

[63] Kargaev, P. P. and Volberg, A. L. (1992). Three results concerning the support of functions and their Fourier transforms. *Indiana Univ. Math. J.* **41**, 1143–1164. MR 94d:42018

[64] Kay, I. and Silverman, R. A. (1959). On the uncertainty relation for real signals: postscript. *Inform. and Control* **2**, 396–397. MR 23 #B2091

[65] Kempf, A. (1994). Uncertainty relation in quantum mechanics with quantum group symmetry. *J. Math. Phys.* **35**, 4483–4496. MR 95k:81063

[66] Kennard, E. H. (1927). Zur Quantenmechanik einfacher Bewegungstypen. *Zeit. Physik* **44**, 326–352.

[67] Knapp, A. W. (1986). *Representation Theory of Semisimple Groups*. Princeton University Press, Princeton, NJ. MR 87j:22022

[68] Koppinen, M. (1994). Uncertainty inequalities for Hopf algebras. *Comm. Algebra* **22**, 1083–1101. MR 95a:16051

[69] Kraus, K. (1967). A further remark on uncertainty relations. *Zeit. Physik* **201**, 134–141. MR 35 #2570

[70] Kumar, A. (1992). A qualitative uncertainty principle for hypergroups. *Functional Analysis and Operator Theory* (B. S. Yadav and D. Singh, eds.). *Lecture Notes in Math.* **1511**. Springer, Berlin, 1–9. MR 94a:43012

[71] Lahti, P. J. and Maczynski, M. J. (1987). Heisenberg inequality and the complex field in quantum mechanics. *J. Math. Phys.* **28**, 1764–1769. MR 88g:81011

[72] Landau, H. J. (1985). An overview of time and frequency limiting. *Fourier Techniques and Applications* (J. F. Price, ed.). Plenum, New York, 201–220.

[73] Landau, H. J. and Pollak, H. O. (1961). Prolate spheroidal wave functions, Fourier analysis and uncertainty II. *Bell Syst. Tech. J.* **40**, 65–84. MR 25 #4147

[74] Landau, H. J. and Pollak, H. O. (1962). Prolate spheroidal wave functions, Fourier analysis and uncertainty III: the dimension of the space of essentially time- and band-limited signals. *Bell Syst. Tech. J.* **41**, 1295–1336. MR 26 #5200

[75] Landau, H. J. and Widom, H. (1980). Eigenvalue distribution of time and frequency limiting. *J. Math. Anal. Appl.* **77**, 469–481. MR 81m:45023

[76] Leipnik, R. (1959). Entropy and the uncertainty principle. *Inform. and Control* **2**, 64–79.

[77] Lieb, E. H. (1978). Proof of an entropy conjecture of Wehrl. *Comm. Math. Phys.* **62**, 35–41. MR 80d:82032

[78] ———. (1990). Integral bounds for radar ambiguity functions and Wigner distributions. *J. Math. Phys.* **31**, 591–599. MR 91f:81076

[79] Logvinenko, V. N. and Sereda, Ju. F. (1974). Equivalent norms in spaces of entire functions of exponential type. *Teor. Funksii Funktsional. Anal. i Prilozhen.* **20**, 102–111. MR 57 #17232

[80] Matolcsi, T. and Szücs, J. (1973). Intersection des mesures spectrales conjuguées. *C. R. Acad. Sci. Paris* **277**, A841–A843. MR 48 #4804

[81] Melenk, J. M. and Zimmerman, G. (1996). Functions with time and frequency gaps. *J. Fourier Anal. Appl.* **2**, 611–614.

[82] Meshulam, R. (1992). An uncertainty principle for groups of order pq. *European J. Combin.* **13**, 401–407. MR 93e:20036

[83] Mustard, D. (1991). Uncertainty principles invariant under the fractional Fourier transform. *J. Austral. Math. Soc. Ser. B* **33**, 180–191. MR 93c:42007

[84] Nahmod, A. R. (1994). Generalized uncertainty principles on spaces of homogeneous type. *J. Funct. Anal.* **119**, 171–209. MR 95g:42027

[85] Narcowich, F. J. (1990). Geometry and uncertainty. *J. Math. Phys.* **31**, 354–364. MR 91a:81088

[86] Nazarov, F. L. (1993). Local estimates for exponential polynomials and their applications to inequalities of the uncertainty principle type. *Algebra i Analiz* **5**, 3–66. MR 94k:42004

[87] Nelson, E. (1959). Analytic vectors. *Ann. of Math.* (2) **70**, 572–615. MR 21 #5901

[88] Pati, V., Sitaram, A., Sundari, M., and Thangavelu, S. (1996). An uncertainty principle for eigenfunction expansions. *J. Fourier Anal. Appl.* **2**, 427–433.

[89] Pearl, J. (1973). Time, frequency, sequency, and their uncertainty relations. *IEEE Trans. Inform. Theory* **IT-19**, 225–229. MR 54 #12344

[90] Perlstadt, M. (1986). Polynomial analogues of prolate spheroidal wave functions and uncertainty. *SIAM J. Math. Anal.* **17**, 340–248. MR 87f:42026

[91] Price, J. F. (1983). Inequalities and local uncertainty principles. *J. Math. Phys.* **24**, 1711–1714. MR 85e:81056

[92] Price, J. F., (ed.). (1985). *Fourier Techniques and Applications*. Plenum, New York.

[93] Price, J. F. (1987). Sharp local uncertainty inequalities. *Studia Math.* **85**, 37–45. MR 88f:42029

[94] Price, J. F. and Racki, P. C. (1985). Local uncertainty inequalities for Fourier series. *Proc. Amer. Math. Soc.* **93**, 245–251. MR 86e:42018

[95] Price, J. F. and Sitaram, A. (1988). Functions and their Fourier transforms with supports of finite measure for certain locally compact groups. *J. Funct. Anal.* **79**, 166–182. MR 90e:43003

[96] ———. (1988). Local uncertainty inequalities for compact groups. *Proc. Amer. Math. Soc.* **103**, 441–447. MR 89i:43004

[97] ———. (1988). Local uncertainty inequalities for locally compact groups. *Trans. Amer. Math. Soc.* **308**, 105–114. MR 89h:22017

[98] Shahshahani, M. (1989). Poincaré inequality, uncertainty principle, and scattering theory on symmetric spaces. *Amer. J. Math.* **111**, 197–224. MR 90g:22017

[99] Shannon, C. (1948)/(1949). A mathematical theory of communication. *Bell System Tech. J.* **27**, 379–423 and 623–656. Reprinted: Shannon, C. and Weaver, W. *The Mathematical Theory of Communication*. Univ. of Illinois Press, Urbana. MR 10, 133e

[100] Sitaram, A. and Sundari, M. An analogue of Hardy's theorem for very rapidly decreasing functions on semi-simple Lie groups. *Pacific J. Math.* (to appear).

[101] Sitaram, A., Sundari, M., and Thangavelu, S. (1995). Uncertainty principles on certain Lie groups. *Proc. Indian Acad. Sci. Math. Sci.* **105**, 135–151.

[102] Skoog, R. A. (1970). An uncertainty principle for functions vanishing on a half-line. *IEEE Trans. Circuit Theory* **CT-17**, 241–243. MR 41 #8926

[103] Slepian, D. (1964). Prolate spheroidal wave functions. Fourier analysis and uncertainty IV: extensions to many dimensions; generalized prolate spheroidal wave functions. *Bell Syst. Tech. J.* **43**, 3009–3057. MR 31 #5993

[104] ———. (1976). On bandwidth. *Proc. IEEE* **64**, 292–300. MR 57 #2738

[105] ———. (1978). Prolate spheroidal wave functions. Fourier analysis and uncertainty V: the discrete case. *Bell Syst. Tech. J.* **57**, 1371–1430.

[106] ———. (1983). Some comments on Fourier analysis, uncertainty and modeling. *SIAM Rev.* **25**, 379–393. MR 84i:94016

[107] Slepian, D. and Pollak, H. O. (1961). Prolate spheroidal wave functions, Fourier analysis and uncertainty I. *Bell Syst. Tech. J.* **40**, 43–63. MR 25 #4146

[108] Smith, K. T. (1990). The uncertainty principle on groups. *SIAM J. Appl. Math.* **50**, 876–882. MR 91i:94008

[109] Spera, M. (1993). On a generalized uncertainty principle, coherent states, and the moment map. *J. Geom. Phys.* **12**, 165–182. MR 94m:58097

[110] Strichartz, R. S. (1989). Uncertainty principles in harmonic analysis. *J. Funct. Anal.* **84**, 97–114. MR 91a:42017

[111] ———. (1994). Construction of orthonormal wavelets. *Wavelets: Mathematics and Applications* (J. J. Benedetto and M. W. Frazier, eds.), 23–50. MR 94i:42047

[112] Sun, L. (1994). An uncertainty principle on hyperbolic space. *Proc. Amer. Math. Soc.* **121**, 471–479. MR 94h:43005

[113] Sundari, M. Hardy's theorem for the n-dimensional Euclidean motion group. *Proc. Amer. Math. Soc.* (to appear).

[114] Thangavelu, S. (1990). Some uncertainty inequalities. *Proc. Indian Acad. Sci. Math. Sci.* **100**, 137–145. MR 91h:43004

[115] Uffink, J. B. M. and Hilgevoord, J. (1984). New bounds for the uncertainty principle. *Phys. Lett.* **105A**, 176–178.

[116] Uffink, J. B. M. and Hilgevoord, J. (1985). Uncertainty principle and uncertainty relations. *Found. Phys.* **15**, 925–944. MR 87f:81012

[117] Voit, M. (1993). An uncertainty principle for commutative hypergroups and Gelfand pairs. *Math. Nachr.* **164**, 187–195. MR 95d:43005

[118] Weyl, H. (1928). *Gruppentheorie und Quantenmechanik*. S. Hirzel, Leipzig. Revised English edition; *The Theory of Groups and Quantum Mechanics* Methuen, London, 1931; reprinted by Dover, New York, 1950.

[119] Wiener, N. (1956). *I Am a Mathematician*. MIT Press, Cambridge.

[120] Williams, D. N. (1979). New mathematical proof of the uncertainty relation. *Amer. J. Phys.* **47**, 606–607. MR 80d:81059

[121] Wolf, J. A. (1992). The uncertainty principle for Gelfand pairs. *Nova J. Algebra Geom.* **1**, 383–396. MR 94k:43006

Received June 25, 1996

Department of Mathematics, University of Washington, Seattle, WA 98195-4350
e-mail: folland@math.washington.edu

Statistics and Mathematics Division, Indian Statistical Institute, R. V. College Post Office, Bangalore 560 059 India
e-mail: sitaram@isibang.ernet.in