UNCERTAINTY PRINCIPLE INEQUALITIES
AND
SPECTRUM ESTIMATION

John J. Benedetto
Department of Mathematics
and
Systems Research Center
University of Maryland

NATO Advanced Study Institute

Fourier Analysis and its Applications
Il Ciocco, Italy
July 16-29, 1989

UNCERTAINTY PRINCIPLE INEQUALITIES AND SPECTRUM ESTIMATION

JOHN J. BENEDETTO
Department of Mathematics
and
Systems Research Center
University of Maryland
College Park, MD 20742

ABSTRACT. Uncertainty principle inequalities are useful devices for estimating signal duration and power spectra in signal analysis. The classical uncertainty principle of quantum mechanics, formulated by Heisenberg, Pauli, Weyl; and Wiener, is an example of such an inequality; and it was applied to signal analysis by Gabor.

Weighted and local versions of the classical uncertainty principle inequality are proved in order to provide fine estimation of signal duration. Further, upper bounds in the case of the classical inequality are analyzed for the wavelet bases or frames in which the signals are decomposed. Finally, in the context of the Bell Labs uncertainty principle inequality, a power spectrum estimation method is introduced; and this leads to a multidimensional signal reconstruction technique. The one-dimensional version of this technique is due to Wiener and Wintner and was developed by Bass and Bertrandias.

The setting throughout is $\mathbb{R}^d$; and the methods include weighted Fourier transform norm inequalities, wavelet theory, and Wiener's generalized harmonic analysis.

INTRODUCTION

We shall present a group of weighted Fourier transform norm inequalities. These inequalities are unified by their common heritage from the classical uncertainty principle of quantum mechanics and by their form which provides estimates of signal energy in terms of both time and frequency information. Weighted Fourier transform norm inequalities are motivated by some of the central issues of signal analysis. For example, in linear system theory weights correspond to various filters in energy concentration problems, and in prediction theory weighted L^p-spaces arise for weights corresponding to power spectra of stationary stochastic processes [B1].

Our presentation is organized as follows:

1. The classical uncertainty principle inequality;
2. Weighted uncertainty principle inequalities;

3. Local uncertainty principle inequalities;
4. The classical uncertainty principle and wavelet theory;
5. Spectrum estimation and the Wiener-Wintner theorem;
A. Closure theorems;
B. Notation.

We shall forego a summary of results in this Introduction, but we do make the following remarks.

a. Section 1 not only provides a statement and background of the classical uncertainty principle inequality, but motivates the topics and points of view of the subsequent sections, e.g., the goals of modern wavelet theory were formulated in terms of the classical uncertainty principle during the 1940s, cf., Sections 1.2.4-1.2.6 and Section 4.

b. The inequalities proved in Sections 2 and 3 allow for further analysis to determine maximal spaces for which they are valid. Appendix A indicates the procedure.

c. The important Bell Labs uncertainty principle [LPS] is not one of our topics, but it does play a rôle in the manner we view power spectrum estimation, e.g., Section 5.

d. Our treatment of uncertainty principle inequalities does not depend on the beautiful work of Cowling and Price [CP], but we would be remiss not to reference their contributions.

e. Finally, Section 2 is part of an ongoing project with Hans Heinig.

1. The Classical Uncertainty Principle Inequality

1.1. STATEMENT AND PROOF OF THE INEQUALITY

The *classical uncertainty principle inequality* is

1.1.1. *Theorem.* Given $(t_0,\gamma_0) \in \mathbb{R}\times\hat{\mathbb{R}}$. Then

$$\forall f \in \mathcal{S}(\mathbb{R}),\ \|f\|_2^2 \le 4\pi\|(t-t_0)f(t)\|_2\|(\gamma-\gamma_0)\hat{f}(\gamma)\|_2 \tag{1.1.1}$$

and there is equality in (1.1.1) if

$$f(t) = C\, e^{-s(t-t_0)^2}\, e^{2\pi i t\gamma_0} \tag{1.1.2}$$

for $C \in \mathbb{C}$ and $s > 0$.

Proof: The mapping, $f(t) \longmapsto f(t+t_0)\, e^{-2\pi i t\gamma_0}$, shows that it is sufficient to verify (1.1.1) for $(t_0,\gamma_0) = (0,0)$.

The following calculation gives (1.1.1) for $(t_0, \gamma_0) = (0,0)$:

$$\|f\|_2^4 = \left[\int t\,|f(t)^2|'\,dt\right]^2 \le \left[\int |t|\,|f(t)^2|'\,dt\right]^2 \le 4\left[\int |t\bar{f}(t)f'(t)|\,dt\right]^2$$

$$\le 4\|tf(t)\|_2^2\|f'(t)\|_2^2 = 16\pi^2\|tf(t)\|_2^2\|\gamma\hat{f}(\gamma)\|_2^2. \quad (1.1.3)$$

If f is defined by (1.1.2) then equality is obtain in (1.1.1) by direct calculation.

q.e.d.

1.1.2. *An Elementary Closure Question.* Is (1.1.1) valid for each $f \in L^2(\mathbb{R})$? Obviously, the answer is "yes" if either of the factors on the right side is infinite. If $\|tf(t)\|_2 + \|\gamma\hat{f}(\gamma)\|_2 < \infty$ there is something to check, but the answer is still "yes". The proof follows from Theorem A.1.2 in Appendix A.1 for the case $L^2_{1,1}(\mathbb{R})$. As we shall see, the problem of extending an inequality, proved for functions in a convenient space, to all appropriate functions is not always routine.

1.1.3. *The Classical Inequality for Odd Functions.* DeBruijn [De, Theorem 3.4] has proved that

$$3\|f(t)\|_2^2 \le 4\pi\|tf(t)\|_2\,\|\gamma\hat{f}(\gamma)\|_2$$

for all odd functions $f \in \mathcal{S}(\mathbb{R})$.

1.1.4. *Underlying Inequalities.* The main ingredients in (1.1.3) are integration by parts, Hölder's inequality, and the Plancherel theorem. We shall extend and refine Theorem 1.1.1 in several ways in Sections 2 and 3. The main ingredients of our proofs will be the same: integration by parts or conceptually similar ideas such as generalizations of Hardy's inequality; Hölder's inequality; and weighted norm inequalities for the Fourier transform, of which the Plancherel theorem is a special case.

1.2. HISTORY AND MOTIVATION

The classical uncertainty principle inequality was developed in the context of quantum mechanics. Our results in Sections 2 and 3 will be interpreted in the spirit of Gabor's view of communication theory [G]. We now address this transition.

1.2.1. *Concert Pitch and the Classical Uncertainty Principle Inequality.* In I am a mathematician (pp. 105-107), Wiener comments on his 1925 lecture at Göttingen. It was apparently at this lecture that Theorem 1.1.1 was first proved [Ba]; and, of course, Heisenberg's (as well as Pauli's, Schrödinger's, Weyl's, etc.) profound contributions on the classical uncertainty principle were being made during the same period - a human counterexample to the very same principle!

In any case, Wiener explicitly conceived of the analogy between the laws of physics and musical notation in the sense of normal behavior becoming unpredictable when normal time intervals are sufficiently compressed. For example, let us consider the following idealized piano experiment. The standard for concert pitch is that the A above middle C should have 440 vibrations per second. Thus, the A four octaves down (and the last key on the piano) should have 27.5 vibrations per second. Suppose we could strike this last key for a time interval of 1/30 seconds, i.e., the hammer strikes the string and 1/30 seconds later the damper returns to the string, thereby stopping the sound. We have very precise time information but correspondingly imprecise frequency information since the emitted sound is anything but the desired pure periodic pitch of this low A.

This piano experiment has the flavor of the classical uncertainty principle inequality, and we can quantify it in terms of Theorem 1.1.1. In fact, if a sound (signal) f is emitted at time t_0 and lasts a very short time, then (1.1.1) asserts that the frequency range for f is quite broad. In particular, f is not close to a pure tone of frequency γ_0, for, if it were, then $\|(\gamma-\gamma_0)\hat{f}(\gamma)\|_2$, as well as $\|(t-t_0)f(t)\|_2$, would be small in contrast to the "loudness" $\|f\|_2$.

By comparison, the relevance of Theorem 1.1.1 for quantum mechanics can be illustrated by considering a freely moving mass point with varying location $x \in \mathbb{R}$. The term $\|tf(t)\|_2^2$ represents the average distance of x from its expected value $t_0 = 0$. In fact, the position x is interpreted as a random variable depending on the state function f; more precisely, the probability that x is in a given region $A \subseteq \mathbb{R}$ is defined as $\int_A |f(t)|^2 dt$, and $\|tf(t)\|_2^2$ is the variance of x.

1.2.2. *Wiener on Linear Operators and Quadratic Means.* Quantum mechanics has been the spawning ground for the two topics in our title: uncertainty principle inequalities and spectrum estimation.

We commented on the former well-known relationship in Section 1.2.1. For the latter we require Wiener's brilliant insights from the late 1920's. He and Max Born were among the first (some say the first) to associate linear operators on function spaces to physical quantities. They did this in terms of arithmetic mean integral operators in their study of the Heisenberg theory [W, Volume 3]. Wiener was then able to apply their idea for quantum mechanics to many other topics including filtering and prediction. The mathematical theory for this point of view is Wiener's generalized harmonic analysis, i.e., the Fourier analysis of non-periodic undamped signals [W, Volume 2]. A critical component for effecting this analysis is the space of functions having bounded quadratic means. Windowing methods in the spectrum estimation problem can be viewed as applications of Wiener's theory. This is the topic of Section 5, where we also illustrate the role of the uncertainty principle in spectrum estimation.

1.2.3. *Gabor and the Fundamental Principle of Communication.* In 1946, Dennis Gabor formulated and analyzed the "fundamental principle of communication" [G]. The initial idea is much like Wiener's described in Section 1.2.1: the more a signal f is concentrated, the longer the bandwidth. Thus, local information about f at any given time is inextricably contained in all of the bandwidth. More precisely, if a good description of f in terms of $\hat{f}$ is required then we must have good information about $\hat{f}$ on all of its domain. This point of view can be quantified to a certain extent by effective sampling; but the principle with which Gabor initiates his study is the intrinsic irreconcilability of achieving high definition reconstruction or transmission (of f) and of obtaining sufficient bandwidth sampling (of $\hat{f}$).

As an example, suppose the mode of transmission is by means of Fourier series and that we wish to transmit f on $[-T,T]$ and have available the frequency band $[-\Omega,\Omega]$. The Fourier series of f on $[-T,T]$ considered as a 2T-periodic function on $\mathbb{R}$ has the form $\Sigma c_n e^{\pi int/T}$, so that the transmission of f is equivalent to the transmission of $\{c_n\}$. Because of the prescribed bandwidth there are effectively $2T\Omega$ "spectral lines" $\gamma_n = n/T$ available to transmit f on $[-T,T]$. Since each datum c_n is associated with a specific spectral line (in the Fourier series), we can only expect to transmit $2T\Omega$ data. The problem (and the principle) is that high resolution of f by N points in $[-T,T]$, say, may require many more than these $2T\Omega$ data and their independent combinations.

Gabor's landmark paper completed a line of research begun by Carson, Nyquist, Küpfmüller, and Hartley, concurrent with the quantum mechanics formulation of the classical uncertainty principle. One of Gabor's basic arguments is that Theorem 1.1.1 is at the root of the "fundamental principle of communication."

1.2.4. *Gabor on the Classical Uncertainty - Principle and Wavelets.* Because of its minimization property in the classical uncertainty principle inequality, Gabor reasoned that the modulated probability pulse in (1.1.2) is the "natural basis on which to build up an analysis of signals in which both time and frequency are recognized as references" [G, p.435]. His analysis of signals, the "Gabor representation," is the origin of the important Weyl-Heisenberg frame - decompositions [DGM] which play such an important role in wavelet theory, e.g., [FG; FJ; HW; M1; M2] for diverse ideas and technology as well as extensive bibliographies on wavelets. Wavelet theory provides discrete signal reconstruction by means of time and frequency localization to the extent allowable by the uncertainty principle, cf., Section 4. The joint localization stands in contrast to the point of view effected by the "fundamental principle of communication." Because of the uncertainty principle there are no contradictions in this contrast.

1.2.5. *Definition/Representation.* a. Given $g \in L^1_{loc}(\mathbb{R})$. The Gabor wavelet $\psi = \psi_g$ is defined as

$$\psi_g(u;t,\gamma) = g(u-t)e^{2\pi i(u\gamma - ct\gamma)}$$

for fixed $c \in \mathbb{R}$. The Gabor wavelet transform of $f \in L^1_{loc}(\mathbb{R})$ is the function

$$F_\psi(f)(t,\gamma) = F_\psi(t,\gamma) = \int f(u)\,\overline{\psi(u;t,\gamma)}du$$

defined on $\mathbb{R}\times\hat{\mathbb{R}}$.

b. As indicated in Section 1.2.4, Gabor considered the case $g(u) = e^{-su^2}$, $s > 0$. The resulting wavelet can be interpreted in terms of the parameters: s is the sharpness of the pulse, t, the epoch of its peak, and γ and $\varphi = ct\gamma$ are the frequency and phase constants of modulating oscillation.

c. Given $g \in L^2(\mathbb{R})$ and $a > 0$. Let $\alpha = 1/a$ and $c = 1$. Then $\psi_{m,n}(u) = \psi_g(u;ma,n\alpha)$ is an orthonormal basis of $L^2(\mathbb{R})$ under certain conditions on g, e.g., if the Zak transform of g has modulus 1. In this case we have the Gabor representation,

(1.2.1) $$\forall f \in L^2(\mathbb{R}),\ f = \sum_{m,n} c_{m,n}(f)\ \psi_{m,n} \quad \text{in} \quad L^2(\mathbb{R}).$$

d. The orthonormal case in part c gives a false impression of versatility since orthonormality is never obtained for compactly supported continuous g or for the Gaussian. On the other hand, the von Neumann lattices $\{(ma,n\alpha)\}$, $\alpha = 1/a$, can be perturbed and/or the orthonormality can be weakened so that the Gabor representation is valid for many different Gabor wavelets.

In light of our discussion in Section 5, we have proved the following continuous Gabor representation for bounded Gabor wavelets ("Gabor representations and wavelets" AMS Contemporary Math. Series, 1989).

1.2.6. *Theorem.* Given $g \in L^\infty(\mathbb{R})$ with continuous and not identically zero autocorrelation,

$$G(u) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} g(u+v)\overline{g(v)}dv.$$

Then, for each $f \in L^1(\mathbb{R})$, $\lim_{n\to\infty} \|f_n - f\|_1 = 0$ where

$$f_n(u) = \frac{1}{G(0)} \lim_{T\to\infty} \frac{1}{2T} \iint_{-T}^{T} F_\psi(f)(t,\gamma)\psi(u;t,\gamma)\rho_n(\gamma)dtd\gamma$$

and $\{\check{\rho}_n\} \subseteq L^1(\mathbb{R})$ is an L^1-approximate identity.

1.2.7. *Remark.* Implicit in considering non-Gaussian Gabor wavelets and in veering towards signal analysis with its host of data windows and filters and prediction theoretic weights, we are viewing Theorem 1.1.1 as one of many possible uncertainty principle inequalities. One expects analogous inequalities and interpretations for other norms and weights besides $\|\cdots\|_2$ and t^2 or γ^2; and this is the subject of Sections 2 and 3.

1.3. EXAMPLES

1.3.1. *Unboundedness in the Classical Uncertainty Principle Inequality.* Suppose $(t_0, \gamma_0) = (0,0)$. If $f(t) = \chi_{(T)}(t)$ then $\hat{f}(\gamma) = [\sin(2\pi T\gamma)]/[\pi\gamma]$ and the right side of (1.1.1) is infinite. Similarly, if $f \in L^2(\mathbb{R})$ behaves like $|t|^a$ as $|t| \longrightarrow \infty$, where $a \in [-3/2, -1/2)$ then the right side of (1.1.1) is infinite. ($a < -1/2$ ensures $f \in L^2(\mathbb{R})$ and $a \geq -3/2$ ensures $\|tf(t)\|_2 = \infty$.)

Theorem 1.1.1 does not provide useful information for these signals. This fact, and the importance of comparing time and frequency energy concentrations in signal analysis, lead to the Bell Labs uncertainty principle [LPS], cf., Proposition 5.1.5.

1.3.2. *The Classical Uncertainty Principle Inequality and Energy Concentration.* Given $T, \Omega > 0$. For each $f \in L^2(\mathbb{R})$,

$$E_T(f) = \left(\int_{-T}^{T} |f(t)|^2 dt\right) \Big/ \|f\|_2^2 \quad \left[\text{resp.,} \quad E_\Omega(\hat{f}) = \left(\int_{-\Omega}^{\Omega} |\hat{f}(\gamma)|^2 d\gamma\right) \Big/ \|f\|_2^2\right]$$

represents the proportion of the total energy of f in $[-T,T]$ (resp., of $\hat{f}$ in $[-\Omega,\Omega]$). It is clear that $E_T(f)\, E_\Omega(\hat{f}) \leq 1$ and that

$$(1.3.1) \qquad \forall \varepsilon > 0,\ \exists f_\varepsilon \in L^2(\mathbb{R}) \text{ such that } E_T(f_\varepsilon)\, E_\Omega(\hat{f}_\varepsilon) < \varepsilon.$$

Since $\operatorname{supp} f_\varepsilon$ can be taken as a subset of $[-T,T]$ in the verification of (1.3.1), we consider the following ratios vis a vis Theorem 1.1.1.

For each T-time limited $f \in L^2(\mathbb{R})$, let $V(\hat{f}) = \|\gamma\hat{f}(\gamma)\|_2^2 / \|f\|_2^2$. Then, for such f, we have

$$\frac{1}{16\pi^2 T^2} < \inf\{V(\hat{f})\} \leq \sup\{V(\hat{f})\} = \infty$$

by Theorem 1.1.1.

1.3.3. *The Classical Uncertainty Principle Inequality for Hilbert Spaces.* We close Section 1, as we began it, with a statement of the classical uncertainty principle inequality. We give the conventional Hilbert space formulation, cf., the remarks in Section 1.2.2.

Theorem. Let A, B be self-adjoint operators on a Hilbert space H (A and B need not be continuous). Define the commutator $[A,B] = AB - BA$, the expectation $E_f(A) = (Af, f)$ of A at $f \in D(A)$ ($D(A)$ is the domain of A), and the variance $\sigma_f^2(A) = E_f(A^2) - \{E_f(A)\}^2$ of A at $f \in D(A^2)$. If $f \in D(A^2) \cap D(B^2) \cap D(i[A,B])$ and $\|f\| = 1$ then

$$\{E_f(i[A,B])\}^2 \le 4\,\sigma_f^2(A)\,\sigma_f^2(B).$$

The verification is routine, and (1.1.1) is a corollary for $H = L^2(\mathbb{R})$ and for $A(f)(t) = tf(t)$ and $B(f)(t) = i(2\pi i\gamma\hat{f}(\gamma))^\vee(t)$, $f \in \mathcal{S}(\mathbb{R})$.

2. Weighted Uncertainty Principle Inequalities

2.1. FIRST METHOD: HARDY AND FOURIER TRANSFORM NORM INEQUALITIES

2.1.1. *Two Fundamental Inequalities.* a. The <u>Hardy operator</u> is the positive linear operator P_d defined as

$$P_d(f)(x) = \int_0^{x_d}\cdots\int_0^{x_1} f(t_1,\cdots,t_d)\,dt_1\cdots dt_d = \int_{<0,x>} f(t)\,dt$$

for Borel measurable functions f on $\mathbb{R}^{+d}$. The <u>dual Hardy operator</u> P'_d is defined as

$$P'(f)(x) = \int_{x_d}^{\infty}\cdots\int_{x_1}^{\infty} f(t_1,\cdots,t_d)\,dt_1\cdots dt_d = \int_{<x,\infty>} f(t)\,dt,$$

where $x > 0$, i.e., each $x_j > 0$ for $x = (x_1,\cdots,x_d)$.

<u>Hardy's inequality</u> (1920) is

(2.1.1) $$\int_0^{\infty} P_1(f)(t)^p t^{-p}\,dt < \left(\frac{p}{p-1}\right)^p \int_0^{\infty} f(t)^p\,dt,$$

where $p > 1$ and $f \ge 0$ ($f \not\equiv 0$) is Borel measurable.

b. The <u>Hausdorff-Young inequality</u> is

(2.1.2) $$\forall f \in \mathcal{S}(\mathbb{R}^d), \quad \|\hat{f}\|_{p'} \le B_d(p)\|f\|_p,$$

where $B_d(p) = (p^{1/p}(p')^{-1/p'})^{d/2}$ and $1 < p \le 2$.

The extension of Hausdorff-Young's inequality for Fourier series to the case of Fourier transforms is due to Titchmarsh (1924). $B_d(p)$ is the best constant and (2.1.2) is an equality for $f(t) = e^{-\pi|t|^2}$

(Babenko (1961) and Beckner (1975)). Finally, (2.1.2) allows an extension to $L^p(\mathbb{R}^d)$ since $\mathcal{S}(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$; in particular, the Fourier transform is well-defined for each $f \in L^p(\mathbb{R}^d)$, $1 < p \le 2$.

2.1.2. *Proposition.* Given $1 < p \le 2$.

$$(2.1.3) \qquad \forall f \in \mathcal{S}(\mathbb{R}), \quad \|f\|_2^2 \le 4\pi B_1(p)\|tf(t)\|_p \|\gamma\hat{f}(\gamma)\|_p.$$

The proof is similar to the proof of Theorem 1.1.1: the L^p-version of Hölder's inequality is used instead of the L^2-version, and the Hausdorff-Young inequality replaces the Plancherel theorem.

2.1.3. *Proposition* [HS, Theorem 1.1]. Given $1 < p \le 2$.

$$(2.1.4) \qquad \forall f \in \mathcal{S}_o(\mathbb{R}), \quad \|f\|_2^2 \le 2\pi p B_1(p)\|tf(t)\|_p \|\gamma\hat{f}(\gamma)\|_p.$$

The constant in (2.1.4) is sharper than that in (2.1.3) for $1 < p < 2$. The closure properties of $\mathcal{S}_o(\mathbb{R})$ are discussed in Appendix A.3. The proof of (2.1.4) is similar to the proof of Theorem 1.1.1 but depends on Hardy's inequality in the following way:

$$\int_0^\infty |\hat{f}(\gamma)|^2 d\gamma \le \left[\int_0^\infty |\gamma\hat{f}(\gamma)|^p d\gamma\right]^{1/p} \left[\int_0^\infty |\tfrac{1}{\gamma}\hat{f}(\gamma)|^{p'} d\gamma\right]^{1/p'}$$

$$= \left[\int_0^\infty |\gamma\hat{f}(\gamma)|^p d\gamma\right]^{1/p} \left[\int_0^\infty |P_1((\hat{f})')(\gamma)|^{p'} \gamma^{-p'} d\gamma\right]^{1/p'}$$

$$\le p\left[\int_0^\infty |\gamma\hat{f}(\gamma)|^p d\gamma\right]^{1/p} \left[\int_0^\infty |(\hat{f})'(\gamma)|^{p'} d\gamma\right]^{1/p'},$$

cf., Sections 2.1.7-2.1.9 for a fuller treatment in a more general setting.

2.1.4. *Definition/Theorem.* Definition. Given even non-negative Borel measurable functions u and v on $\hat{\mathbb{R}}$ and $\mathbb{R}$, respectively. Suppose there is $K > 0$ such that

$$(2.1.5) \qquad \sup_{s>0} \left[\int_0^{1/s} u(x)dx\right]^{1/q} \left[\int_0^s v(x)^{-p'/p} dx\right]^{1/p'} = K,$$

where $1 < p \le q < \infty$. In this case we write $(u,v) \in F(p,q)$.

We have the following generalization of the Hausdorff-Young inequality.

Theorem [BH1]. Given $1 < p \le q < \infty$ and even weights u and v for which $(u,v) \in F(p,q)$. Assume $1/u$ and v are increasing on $(0,\infty)$. Then there is $C = C(K)$ such that

$$\forall f \in \mathcal{S}(\mathbb{R}^d) \cap L^p_v(\mathbb{R}^d), \quad \|\hat{f}\|_{q,u} \le C\|f\|_{p,v} < \infty.$$

After completing [BH1] in 1982, we realized we could use Theorem 2.1.4 to prove the following weighted uncertainty principle inequality, e.g., [B2, p.408; HS].

2.1.5. *Theorem.* Given $1 < p \le q < \infty$ and even non-negative Borel measurable functions v and w which are increasing on $(0,\infty)$. Assume $(1/w, v) \in F(p,q)$ with constant K (as in (2.1.5)). Then there is $C = C(K)$ (the same as in Theorem 2.1.4) such that

$$(2.1.6) \qquad \forall f \in \mathcal{S}(\mathbb{R}), \quad \|f\|_2^2 \le 4\pi C(K)\|tf(t)\|_{p,v}\, \|\gamma\hat{f}(\gamma)\|_{q',w^{q'/q}}.$$

Proof: By means of the first part of (1.1.3) (for $\hat{f}$ instead of f), Hölder's inequality, and Theorem 2.1.4 we have the estimate,

$$\|f\|_2^2 = \|\hat{f}\|_2^2 \le 2\int|\gamma(\hat{f})^-(\gamma)(\hat{f})'(\gamma)|d\gamma$$

$$= 2\int|\gamma(\hat{f})^-(\gamma)w(\gamma)^{1/q}||(\hat{f})'(\gamma)w(\gamma)^{-1/q}|d\gamma$$

$$\le 2\left[\int|\gamma(\hat{f})^-(\gamma)|^{q'}\; w(\gamma)^{q'/q}d\gamma\right]^{1/q'}\left[\int|(\hat{f})'(\gamma)|^{q}\; w(\gamma)^{-1}d\gamma\right]^{1/q}$$

$$\le 2C\|((\hat{f})')^{\vee}(t)\|_{p,v}\, \|\gamma\hat{f}(\gamma)\|_{q',w^{q'/q}},$$

and the result is obtained since $((\hat{f})')^{\vee}(t) = 2\pi i t f(t)$. q.e.d.

2.1.6. *Remark.* a. Naturally, the right side of (2.1.6) is not necessarily finite, although it is easy to specify a convenient subspace of $\mathcal{S}(\mathbb{R})$ where it is finite. We shall not discuss the closure question for (2.1.6), that is, the question of characterizing the appropriate space of functions for which (2.1.6) is valid and the right side is finite, cf., Section 1.1.2.

b. The monotonicity and symmetry hypotheses in Theorem 2.1.4 and 2.1.5 can be weakened at the expense of complicating condition (2.1.5) with criteria formulated in terms of rearrangements; [BH2] provides remarks and bibliography for contributions in this direction, as well as for higher dimensions.

If $p = 1$ and $q > 1$ then Theorem 2.1.4 is true for any positive Borel measurable weight u. In this case, the proof is routine and the constant C is explicit [BH1, pp.272-273]. If $p > 1$ the constant C is less explicit, but it can be estimated by examining the proof of Calderón's rearrangement inequality (Studia Math. 26(1966), 273-299) which we use in proving Theorem 2.1.4.

2.1.7. *Weighted Hardy Inequalities in* $\mathbb{R}^d$.

Lemma [He, Theorem 3.1]. Given $1 < p \le q < \infty$ and non-negative Borel measurable functions u and v on $X \subseteq \mathbb{R}^d$. Suppose $P : L^p_v(X) \to L^q_u(X)$ is a positive linear operator with canonical dual operator $P' : L^{q'}_{u^{-q'/q}}(X) \to L^{q'}_{v^{-p'/p}}(X)$ defined by the duality $\int_X P(f)(x)g(x)dx = \int_X f(x)P'(g)(x)dx$. Assume there exist $K_1, K_2 > 0$ such that

$$\forall g \in L^{(q/p)'}(X), \quad \text{for which} \quad g \ge 0 \quad \text{and} \quad \|g\|_{(g/p)'} \le 1,$$

there are non-negative functions,

$$f_1 \in L^p_v(X), \quad h_1 \in L^p_{u^{p/q}g}(X), \quad f_2 \in L^{p'}_{u^{-p'/q}g}(X), \quad h_2 \in L^{p'}_{v^{-p'/p}}(X),$$

with the properties

$$P(f_1) \le K_1 h_1 \quad \text{and} \quad P'(f_2 g) \le K_2 h_2 \tag{2.1.7}$$

and

$$v = f_1^{-p/p'} h_2 \quad \text{and} \quad u = h_1^{-q/p'} f_2^{q/p}.$$

Then $P \in \mathcal{L}(L^p_v(X), L^q_u(X))$, $P' \in \mathcal{L}(L^{q'}_{u^{-q'/q}}(X), L^{p'}_{v^{-p'/p}}(X))$, and $\|P\|, \|P'\| \le K_1^{1/p'} K_2^{1/p}$.

Setting

$$f_1 = v^{-p'/p} P_d(v^{-p'/p})^{-1/p},$$

$$h_1 = P_d(v^{-p'/p})^{-1/p'},$$

$$f_2 = u^{p/q} P'_d(u)^{-p/(qp')},$$

$$h_2 = P_d(v^{-p'/p})^{-1/p'},$$

it is easy to verify (2.1.7) for any non-negative $g \in L^{(q/p)'}(\mathbb{R}^{+d})$, for which $\|g\|_{(q/p)'} \le 1$, as long as (2.1.8), (2.1.9), and (2.1.10) are assumed. As a result, Hernandez obtained the following version of Hardy's inequality on $\mathbb{R}^{+d}$.

Theorem [He, Section 4.2]. Given $1 < p \le q < \infty$ and non-negative Borel measurable functions u and v on $\mathbb{R}^{+d}$. Assume there exist K, $C_1(p)$, $C_2(p) > 0$ such that

$$\sup_{s>0} \left[\int_{\langle s,\infty\rangle} u(x)dx \right]^{1/q} \left[\int_{\langle 0,s\rangle} v(x)^{-p'/p} dx \right]^{1/p'} = K, \tag{2.1.8}$$

(2.1.9) $$\forall x \in \mathbb{R}^{+d}, \quad P_d(v^{-p'/p}(P_d(v^{-p'/p})^{-1/p})(x) \le C_1(p)\, P_d(v^{-p'/p})^{-1/p'},$$

and

(2.1.10) $$\forall x \in \mathbb{R}^{+d}, \quad P_d'(u(P_d'u)^{-1/p'})(x) \le C_2(p)^{q/p}(P_d'u)^{1/p}.$$

Then $P_d \in \mathcal{L}(L_v^p(\mathbb{R}^{+d}),\ L_u^q(\mathbb{R}^{+d}))$, $P_d' \in \mathcal{L}(L^{q'}_{u^{-q'/q}}(\mathbb{R}^{+d}),\ L^{p'}_{v^{-p'/p}}(\mathbb{R}^{+d}))$, and $\|P_d\|,\ \|P_d'\| \le KC_1(p)^{1/p'}C_2(p)^{1/p}$.

Remark. Condition (2.1.8) is necessary and sufficient for weighted Hardy inequalities on $\mathbb{R}$ and necessary on $\mathbb{R}^d$, $d > 1$. Conditions (2.1.9) and (2.1.10) are automatically satisfied on $\mathbb{R}$. Conditions (2.1.8), (2.1.9), and (2.1.10) are sufficient but not necessary on $\mathbb{R}^d$, $d > 1$. Sawyer has given a characterization for $d = 2$.

2.1.8. *Regrouping Lemma.* Let Ω be the subgroup of the orthogonal group whose corresponding matrices with respect to the standard basis are diagonal with ± 1 entries. Each element $\omega \in \Omega$ can be identified with an element $(\omega_1, \cdots, \omega_d) \in \{-1,1\}^d$, and $\omega\gamma = (\omega_1\gamma_1, \cdots, \omega_d\gamma_d)$. Thus,

$$\int F(\gamma)d\gamma = \sum_{\omega\in\Omega} \int_{\hat{\mathbb{R}}^{+d}} F(\omega\gamma)d\gamma,$$

and since

$$\sum_{\omega\in\Omega} a_\omega^{1/r}\, b_\omega^{1/r'} \le \left[\sum_{\omega\in\Omega} a_\omega\right]^{1/r} \left[\sum_{\omega\in\Omega} b_\omega\right]^{1/r'},$$

for $1 < r < \infty$ and $a_\omega, b_\omega \ge 0$, we have -

Lemma. Given $1 < r < \infty$ and suppose $F \in L^r(\hat{\mathbb{R}}^d)$, $G \in L^{r'}(\hat{\mathbb{R}}^d)$. Then

$$\sum_{\omega\in\Omega} \left[\int_{\hat{\mathbb{R}}^{+d}} |F(\omega\gamma)|^r d\gamma\right]^{1/r} \left[\int_{\hat{\mathbb{R}}^{+d}} |G(\omega\gamma)|^{r'} d\gamma\right]^{1/r'} \le \|F\|_r\ \|G\|_{r'}.$$

2.1.9. *Definition/Uncertainty Principle Inequality.*

Definition. $\mathcal{S}_{oa}(\mathbb{R}^d) = \{f \in \mathcal{S}(\mathbb{R}^d):\ \hat{f}(\gamma) = 0 \text{ if some } \gamma_j = 0\} \subseteq \mathcal{S}_o(\mathbb{R}^d)$. Thus, $f \in \mathcal{S}(\mathbb{R}^d)$ is an element of $\mathcal{S}_{oa}(\mathbb{R}^d)$ if $\hat{f} = 0$ on the coordinate axes.

Combining Hardy's inequality (Theorem 2.1.7) and the regrouping lemma (Lemma 2.1.8) we obtain the following uncertainty principle inequality.

<u>Theorem</u>. Given $1 < r < \infty$ and non-negative Borel measurable weights v and w. Suppose $u = w^{-r'/r}$, and assume that, for all $\omega \in \Omega$, the weights $u(\omega\gamma)$ and $v(\omega\gamma)$ satisfy conditions (2.1.8), (2.1.9) and (2.1.10) on $\hat{\mathbb{R}}^{+d}$ for $p = q = r'$ and constants $K(\omega)$, $C_1(p,\omega)$, and $C_2(p,\omega)$. If $C = \sup_{\omega\in\Omega} K(\omega)C_1(p,\omega)^{1/p'}C_2(p,\omega)^{1/p}$ then

$$\text{(2.1.11)} \qquad \forall f \in \mathcal{S}_{oa}(\mathbb{R}^d), \ \|f\|_2^2 \le C\|\hat{f}\|_{r,w}\|\partial_1,\cdots,\partial_d\hat{f}\|_{r',v}.$$

2.1.10. *Method.* At this point, generalizations of <u>Proposition 2.1.3</u> and <u>Theorem 2.1.5</u> can be stated by applying d-dimensional versions of <u>Theorem 2.1.4</u> to the factor $\|\partial_1,\cdots,\partial_d\hat{f}\|_{r',v}$, on the right side of (2.1.11), e.g., <u>Remark 2.1.6b</u>. We shall avoid a baroque extravaganza with all forms of rearrangements, and confine ourselves to the following section.

2.1.11. *Corollaries.* . .

If $v = 1$ and $p = q = r'$ then (2.1.8) has the form,

$$\text{(2.1.12)} \qquad \sup_{s>0} (s_1\cdots s_d)^{1/r}\left[\int_{\langle s,\infty\rangle} u(y)dy\right]^{1/r'} = K,$$

and (2.1.9) is satisfied for $C_1(r') = r^d$.

<u>Corollary</u>. Given $1 < r \le 2$ and let the non-negative Borel measurable weight w be invariant under the action of Ω. Assume $K < \infty$ (in (2.1.12)) for $u = w^{-r'/r}$ and that

$$\text{(2.1.13)} \qquad P'_d(w^{-r'/r}(P'_d(w^{-r'/r})^{-1/r}) \le C_2(r')(P'_d(w^{-r'/r})^{1/r'}.$$

Then

$$\forall f \in \mathcal{S}_{oa}(\mathbb{R}^d), \quad \|f\|_2^2 \le (2\pi)^d r^{d/r}KC_2(r')^{1/r'}B_d(r)\|t_1\cdots t_d f(t)\|_r\|\hat{f}\|_{r,w}$$

$$\le (2\pi)^d r^{d/r}d^{-d/2}KC_2(r')^{1/r'}B_d(r)\|\,|t|^d f(t)\|_r\|\hat{f}\|_{r,w}.$$

The weight $w(\gamma) = |\gamma_1\cdots\gamma_d|^r$, $1 < r \le 2$, is Ω-invariant, $K = (r'-1)^{-d/r'}$ in (2.1.12), and (2.1.13) is satisfied for $C_2(r') = (r(r'-1))^d$. Consequently, we obtain the following d-dimensional generalization of <u>Proposition 2.1.3</u>.

<u>Corollary</u>. Given $1 < r \le 2$. Then

$$\forall f \in \mathcal{S}_{oa}(\mathbb{R}^d), \quad \|f\|_2^2 \le (2\pi r)^d\, B_d(r)\|t_1\cdots t_d f(t)\|_r\|\gamma_1\cdots\gamma_d\hat{f}(\gamma)\|_r.$$

2.2. SECOND METHOD: A_p-WEIGHTS, AND WEIGHTED GRADIENT AND RIESZ TRANSFORM INEQUALITIES

2.2.1. *Weighted Gradient Inequalities.*

Theorem [Si, Theorem 4.1]. Given $1 < q < \infty$ and non-negative Borel measurable functions u and v on $\mathbb{R}^d$.

a. There is a constant $C > 0$ such that

$$\forall g \in C_c^\infty(\mathbb{R}^d), \quad \|g\|_{q,u} \le C\|t\cdot\nabla g(t)\|_{q,v} \tag{2.2.1}$$

if and only if

$$\sup_{s\in\mathbb{R}^d}\left[\int_0^1 u(xs)x^{d-1}dx\right]^{1/q}\left[\int_1^\infty (v(xs)x^d)^{-q'/q}x^{-1}dx\right]^{1/q'} = K < \infty. \tag{2.2.2}$$

The constants C and K satisfy the inequalities,

$$K \le C \le K\, q^{1/q}(q')^{1/q'}.$$

b. There is a constant $C > 0$ such that

$$\forall g \in C_c^\infty(\mathbb{R}^d) \text{ for which } g(0) = 0, \quad \|g\|_{q,u} \le C\|t\cdot\nabla g(t)\|_{q,v} \tag{2.2.3}$$

if and only if

$$\sup_{s\in\mathbb{R}^d}\left[\int_1^\infty u(xs)x^{d-1}dx\right]^{1/q}\left[\int_0^1 (v(xs)x^d)^{-q'/q}x^{-1}dx\right]^{1/q'} = K < \infty. \tag{2.2.4}$$

2.2.2. *A_p-weights and Fourier Transform Norm Inequalities.*

Definition. Given $1 < p < \infty$ and a non-negative Borel measurable function w on $\mathbb{R}^d$. w is an A_p-weight, written $w \in A_p$, if

$$\sup_Q \left[\frac{1}{|Q|}\int_Q w(x)dx\right]\left[\frac{1}{|Q|}\int_Q w(x)^{-p'/p}dx\right]^{p/p'} = K < \infty.$$

Theorem. Given $1 < p \le q \le p' < \infty$ and let w be a non-negative Borel measurable radial function on $\mathbb{R}^d$. Assume $w(|t|)$ is increasing on $(0,\infty)$. There is a constant $C > 0$ such that

$$\forall f \in C_c^\infty(\mathbb{R}^d), \quad \left[\int |\hat{f}(\gamma)|^q\, |\gamma|^{d(\frac{q}{p'}-1)}\, w\!\left(\frac{1}{|\gamma|}\right)^{q/p} d\gamma\right]^{1/q} \le C\|f\|_{p,w} \tag{2.2.5}$$

if and only if $w \in A_p$.

(The reference for the case d = 1 is J. Benedetto, H. Heinig, R. Johnson, "Fourier inequalities with A_p-weights" ISNM 80 (1987), 217-232. The d > 1 version stated above is in [HS1, Theorem 2.10].)

Remark. We view Theorem 2.2.2 as the culmination of some interesting classical analysis. Take d = 1. In case p = q and w = 1, (2.2.5) is the Hardy, Littlewood, Paley theorem (1931),

$$\int |\hat{f}(\gamma)|^p |\gamma|^{p-2} d\gamma \le C\|f\|_p.$$

If $q = p'$ and $w = 1$, (2.2.5) is the Hausdorff-Young theorem. If $w(t) = |t|^\alpha$, $0 \le \alpha < p-1$, then (2.2.6) reduces to Pitt's theorem (1937),

$$\left[\int |\hat{f}(\gamma)|^q |\gamma|^{-\beta} d\gamma \right]^{1/q} \le C\left[\int |f(t)|^p |t|^\alpha dt \right]^{1/p},$$

where $\beta = \frac{q}{p}(\alpha+1) + 1 - q$. The fact that Fourier transform inequalities are characterized in terms of A_p-weights was initially surprising since the A_p-condition was associated with maximal function and singular integral norm inequalities.

2.2.3. *Definition/Theorem on Riesz Transforms.*

Definition. The d-dimensional Riesz transforms are the d singular integral operators $R_1, \cdots, R_d$ defined by the odd kernels $k_j(x) = \Omega_j(x)/|x|^d$, $j = 1, \cdots, d$, where $\Omega_j(x) = c_d x_j/|x_j|$ and $c_d = \Gamma(\frac{d+1}{2})/\pi^{(d+1)/2}$. In fact,

$$(R_j f)(x) = \lim_{T^{-1}, \varepsilon \to 0} \int_{\varepsilon \le |t| \le T} f(x-t) k_j(t) dt$$

exists a.e. for each $f \in L^p(\mathbb{R}^d)$, $1 < p < \infty$, and there is $C = C(p)$ such that

$$\forall f \in L^p(\mathbb{R}^d), \quad \|R_j f\|_p \le C\|f\|_p,$$

$j = 1, \cdots, d$. $C = C(p)$ does not depend on d [GR, p.223]. Also, we compute

$$\hat{k}_j(\gamma) = -i \frac{\gamma_j}{|\gamma|}, \quad j = 1, \cdots, d$$

Theorem (Hunt, Muckenhoupt, and Wheeden, 1973). Given $1 < p < \infty$ and suppose $w \in A_p$. Then $R_j \in \mathcal{L}(L^p_w(\mathbb{R}^d), L^p_w(\mathbb{R}^d))$, $j = 1, \cdots, d$, e.g., [GR, pp.196, 204, 411-413].

2.2.4. *Uncertainty Principle Inequality.*

Theorem. Given $1 < r \le 2$ and a non-negative radial weight $w \in A_r$ on $\mathbb{R}^d$ for which $w(|t|)$ is increasing on $(0,\infty)$. Assume

$$\sup_{s\in\mathbb{R}^d}\left(\int_0^1 \frac{w(xs)^{-r'/r}}{|xs|^{r'}}\, x^{d-1}dx\right)^{1/r'} \times \tag{2.2.6}$$

$$\times\left(\int_1^\infty w\left(\frac{1}{|xs|}\right)^{-1}|xs|^r\, x^{-\frac{dr}{r'}-1}dx\right)^{1/r} = K < \infty.$$

Then there is a constant $C = C(K) > 0$ such that

$$\forall f \in C_c^\infty(\mathbb{R}^d),\quad \|f\|_2^2 \le C\||t|f(t)\|_{r,w}\||\gamma|\hat{f}(\gamma)\|_{r,w}. \tag{2.2.7}$$

Proof: a. For $1 < r < \infty$ we have

$$\|f\|_2^2 \le \||t|f(t)\|_{r,w}\|f\|_{r',u}. \tag{2.2.8}$$

where

$$u(t) = |t|^{-r'}\, w(t)^{-r'/r}.$$

b. The second factor on the right side of (2.2.8) is estimated by means of Theorem 2.2.1a where q and v in (2.2.1) are $q = r'$ and

$$v(t) = |t|^{-r'}\, w\left(\frac{1}{|t|}\right)^{r'/r},$$

respectively. Thus,

$$\|f\|_{r',u} \le C_1\|t\cdot\nabla f(t)\|_{r',v} \tag{2.2.9}$$

if and only if (2.2.6) holds.

c. By Minkowski's inequality the right side of (2.2.9) is bounded by

$$C_1\sum_{j=1}^d\left(\int|(R_jG_j)^\vee(t)|^{r'}\, w\left(\frac{1}{|t|}\right)^{r'/r} dt\right)^{1/r'}, \tag{2.2.10}$$

where $G_j^\vee(t) = \partial_j f(t)$. Combining (2.2.9) and (2.2.10), and applying Theorem 2.2.2 for the case $p = r$ and $q = p'$ (so that $1 < r \le 2$), we obtain

$$\|f\|_{r',u} \le C_1C_2\sum_{j=1}^d\|R_jG_j\|_{r,w}. \tag{2.2.11}$$

d. Finally, combining (2.2.8) and (2.2.11) and applying Theorem 2.2.3 to the right side of (2.2.11) we have the estimate

$$(2.2.12) \qquad \|f\|_2^2 \le C_1C_2C_3\| \, |t|f(t)\|_{r,w} \sum_{j=1}^{d} \|(\partial_j f)^{\wedge}(\gamma)\|_{r,w}$$

$$\le 2\pi d^{1/r'} \, C_1C_2C_3\| \, |t|f(t)\|_{r,w}\left[\int |\hat{f}(\gamma)|^r \left(\sum_{j=1}^{d} |\gamma_j|^r\right) w(\gamma)d\gamma\right]^{1/r}$$

$$\le 2\pi d^{1/2} \, C_1C_2C_3\| \, |t|f(t)\|_{r,w} \, \| \, |\gamma|\hat{f}(\gamma)\|_{r,w}.$$

q.e.d.

2.2.5. *Corollary.* Given $1 < r \le 2$ and $d > r'$. Then there is $C > 0$ such that

$$(2.1.13) \qquad \forall f \in \mathcal{S}(\mathbb{R}^d), \quad \|f\|_2^2 \le C\| \, |t|f(t)\|_r \| \, |\gamma|\hat{f}(\gamma)\|_r.$$

2.2.6. *Remark.* a. In the notation of (2.2.12) the constant C in Corollary 2.2.5 is of the form

$$C = 2\pi d^{1/2} C_1(r,d) B_d(r) C_3(r).$$

Since it is of interest to measure the growth of C as d increases, we note that $C_1(r,d)$ can be estimated in terms of K in (2.2.6) for any w, cf., the second corollary of Section 2.1.11.

b. Theorem 2.2.1b gives rise to an analogue of Theorem 2.2.4 which, for $w = 1$, yields (2.2.13) for $d < r'$.

3. Local Uncertainty Principle Inequalities

3.1. THE RESULTS OF FARIS AND PRICE

Faris' local uncertainty principle inequality is -

3.1.1. *Theorem* [F, (3.2)]. Given a Borel measurable set $E \subseteq \hat{\mathbb{R}}$. Then

$$(3.1.1) \qquad \forall f \in \mathcal{S}(\mathbb{R}), \quad \int_E |\hat{f}(\gamma)|^2 d\gamma \le 2\pi|E| \, \|tf(t)\|_2 \|f\|_2,$$

cf., Remark 3.1.5b.

Price's generalization of Faris' result is -

3.1.2. *Theorem* [P, Theorem 1.1]. Given a Borel measurable set $E \subseteq \hat{\mathbb{R}}^d$ and $\alpha > d/2$. Then

$$(3.1.2) \qquad \forall f \in \mathcal{S}(\mathbb{R}^d), \quad \int_E |\hat{f}(\gamma)|^2 d\gamma \le C|E| \, \| \, |t|^{\alpha} f(t)\|_2^{\frac{d}{\alpha}} \|f\|_2^{2-\frac{d}{\alpha}}$$

where

$$C = \frac{\omega_{d-1}}{2\alpha}\,\Gamma\left(\frac{d}{2\alpha}\right)\Gamma\left(1-\frac{d}{2\alpha}\right)\left(\frac{2\alpha}{d}-1\right)^{\frac{d}{2\alpha}}\left(1-\frac{d}{2\alpha}\right)^{-1}$$

and $C|E|$ is the smallest possible constant.

3.1.3. *A Local Uncertainty Principle.* Theorem 1.1.1 asserts that if f is concentrated then $\hat{f}$ is spread out. Faris' and Price's theorems quantify the folklore that this spread is smooth, and is not the result of $\hat{f}$ having isolated peaks far from the origin. For example, if $f_n = \sqrt{n}\,\chi_{(1/(2n))}$ then the spectral energy in any fixed band E tends to 0 as $n \to \infty$ since $\|f_n\|_2 = 1$ and $\|tf_n(t)\|_2 = 1/(2n\sqrt{3})$. Similarly, if $\hat{f}$ had a peak about the point γ_0, then $\|\hat{f}\|_{2,E_j}$ would be more or less constant as $|E_j| \to 0$, where $\gamma_0 \in E_j$, thereby contradicting (3.1.1).

Theorem 3.1.2 has been generalized to Heisenberg groups [PS].

3.1.4. *Proposition* (Carlson, 1934).

$$\forall f \in \mathcal{S}(\mathbb{R}),\ \|f\|_1^2 \le \|tf(t)\|_2\|f\|_2.$$

Proof: By Hölder's inequality,

$$\frac{1}{\pi}\|f\|_1^2 \le \frac{1}{\sqrt{t}}\int(t + u^2)|f(u)|^2du = t^{1/2}\|f\|_2^2 + t^{-1/2}\|uf(u)\|_2^2;$$

and the result is obtained by minimizing this function of t.

q.e.d.

3.1.5. *Remark* a. Carlson proved his inequality for power series and obtained the smallest possible constant. This inequality and related ones are used in [B2] in conjunction with the Bell Labs uncertainty principle, e.g., Proposition 5.1.5. Beurling (1938) found a new proof of Carlson's inequality allowing an extension to $\mathbb{R}^d$ [K]. For example, in $\mathbb{R}^2$ the inequality has the form,

$$\|f\|_1^2 \le C(\|\partial_1^2\hat{f}\|_2 + \|\partial_2^2\hat{f}\|_2)\|f\|_2.$$

b. Our reason for discussing Carlson's inequality is to suggest alternative proofs of Theorem 3.1.1 and 3.1.2. (3.1.1) is an immediate consequence of Proposition 3.1.4:

$$\|\hat{f}\|_{2,E}^2 \le |E|\,\|\hat{f}\|_\infty^2 \le |E|\,\|\hat{f}\|_1^2 \le 2\pi|E|\,\|tf(t)\|_2\|f\|_2.$$

Theorem 3.1.2 would require a weighted version of Carlson's inequality on $\mathbb{R}^d$.

3.2. LOCAL UNCERTAINTY PRINCIPLE INEQUALITIES ON $\mathbb{R}$

3.2.1. *Hardy's Inequality.* We shall use Theorem 2.1.7 for $d = 1$. In this case (2.1.9) and (2.1.10) are automatically satisfied and we can take $C_1(p) = p'$ and $C_2(p) = p^{p/q}$. Also, if $p = 1$ or $q = \infty$ then Theorem 2.1.7 is valid and $\|P_1\|$ can be taken as K (defined in (2.1.8)).

The weight u in Theorem 2.1.7 can be replaced by $\mu \in M_+(\mathbb{R})$ [BH2].

Finally, we should mention that Hardy's inequality can be formulated on rather general spaces, not just $\mathbb{R}^{+d}$. In particular, there are formulations on $\mathbb{R}^d$. The reason we have chosen $\mathbb{R}^{+d}$ is illustrated in the following result. In fact, it is natural to apply Theorem 3.2.2 to the weight $v = 1$; and in this case the constant defined by (3.2.1) would be infinite if we employed Hardy's inequality on $\mathbb{R}$. Consequently, we deal with $\mathcal{S}_o(\mathbb{R})$ and $\mathbb{R}^+$ instead of $\mathcal{S}(\mathbb{R})$ and $\mathbb{R}$.

3.2.2. *Theorem.* Given $1 < r < \infty$ and non-negative Borel measurable weights u, v, and w. Define $u_\pm(\gamma) = u(\pm\gamma)$ with similar notation for v and w. Assume

(3.2.1) $$\sup_{s>0} \left[\int_s^\infty \frac{u_\pm(x)}{w_\pm(x)^{r'}}\,dx\right]^{1/r'} \left[\int_0^s v_\pm(x)^{-r/r'}\,dx\right]^{1/r} = K_\pm < \infty.$$

There exist constants $C_\pm > 0$ for which $K_\pm \le C_\pm \le K_\pm (r)^{1/r}(r')^{1/r'}$ and such that

(3.2.2) $$\forall f \in \mathcal{S}_o(\mathbb{R}), \quad \|\hat{f}\|^2_{2,u} \le C\|(\hat{f})'\|_{r',v}\|\hat{f}w\|_{r,u},$$

where $C = \max\{C_+, C_-\}$.

"Proof": The result follows from the estimate,

$$\int_0^\infty |\hat{f}(\gamma)|^2 u(\gamma)\,d\gamma = \int_0^\infty |\hat{f}(\gamma)w(\gamma)\hat{f}(\gamma)\frac{1}{w(\gamma)}|u(\gamma)\,d\gamma$$

$$\le \left[\int_0^\infty |\hat{f}(\gamma)w(\gamma)|^r u(\gamma)\,d\gamma\right]^{1/r} \left[\int_0^\infty P_1|(\hat{f})'|(\gamma)^{r'}\frac{u(\gamma)}{w(\gamma)^{r'}}\,d\gamma\right]^{1/r'},$$

Hardy's inequality, and the regrouping lemma.

"q.e.d."

The following result reduces to Proposition 2.1.3 in case $u(\gamma) = 1$ and $w(\gamma) = |\gamma|$.

3.2.3. *Corollary.* Given $1 < r \le 2$ and non-negative Borel measurable weights u and w. Define $u_{\pm}(\gamma) = u(\pm\gamma)$ and $w_{\pm}(\gamma) = w(\pm\gamma)$. Assume

$$\sup_{s>0} s^{1/r}\left[\int_s^{\infty} \frac{u_{\pm}(x)}{w_{\pm}(x)^{r'}}\,dx\right]^{1/r'} = K_{\pm} < \infty. \tag{3.2.3}$$

Then,

$$\forall f \in \mathcal{S}_0(\mathbb{R}),\quad \|\hat{f}\|_{2,u}^2 \le 2\pi B_1(r)K(r)^{1/r}(r')^{1/r'}\|tf(t)\|_r\|\hat{f}w\|_{r,u}, \tag{3.2.4}$$

where $K = \max\{K_+, K_-\}$.

3.2.4. *Example.* a. If $u = \chi_{(\Omega)}$ and $w = 1$, then Corollary 3.2.3 yields the inequality,

$$\int_{-\Omega}^{\Omega}|\hat{f}(\gamma)|^2 d\gamma \le 2\pi B_1(r)\Omega\left[\int|tf(t)|^r dr\right]^{1/r}\left[\int_{-\Omega}^{\Omega}|\hat{f}(\gamma)|^r d\gamma\right]^{1/r}, \tag{3.2.5}$$

for $1 < r \le 2$ and $f \in \mathcal{S}_0(\mathbb{R})$. (3.2.5) improves on Theorem 3.1.1 for $f \in \mathcal{S}_0(\mathbb{R})$.

b. If $u = \chi_E$ and $w(\gamma) = |\gamma|$ then Corollary 3.2.3 yields the inequality,

$$\int_E|\hat{f}(\gamma)|^2 d\gamma \le 2\pi r B_1(r)\left[\int|tf(t)|^r dt\right]^{1/r}\left[\int_E|\gamma\hat{f}(\gamma)|^r d\gamma\right]^{1/r}, \tag{3.2.6}$$

for $1 < r \le 2$ and $f \in \mathcal{S}_0(\mathbb{R})$. (3.2.6) is a local version of Proposition 2.1.3.

c. In searching for smallest constants in (3.2.6), a direct calculation gives the following result for $E = [-\Omega, \Omega]$ and for the Gaussian $g(t) = (1/\sqrt{\pi})e^{-t^2}$ with dilation $g_\lambda(t) = \lambda g(\lambda t)$:

$$\forall \varepsilon > 0,\quad \exists\, \lambda(\varepsilon,\Omega) > 0 \text{ such that } \forall\lambda > \lambda(\varepsilon,\Omega),$$

$$16\pi^2\int|tg_\lambda(t)|^2 dt\int_{-\Omega}^{\Omega}|\gamma\hat{g}_\lambda(\gamma)|^2 d\gamma \le \left[\int_{-\Omega}^{\Omega}|\hat{g}_\lambda(\gamma)|^2 d\gamma\right]^2 + \varepsilon.$$

3.2.5. *Example.* a. Given $1 < r \le 2$, $1 < p < \infty$, a symmetric set $E \subseteq \hat{\mathbb{R}}$ for which $|E| < \infty$, and $f \in \mathcal{S}_0(\mathbb{R})$. Then

$$\int_E|\hat{f}(\gamma)|^2 d\gamma \le C(p,r)|E|^{\frac{1}{p'r'}}\left[\int|tf(t)|^r dt\right]^{1/r}\left[\int_E|\gamma|^{1+\frac{r}{pr'}}|\hat{f}(\gamma)|^r d\gamma\right]^{1/r}, \tag{3.2.7}$$

where

$$C(p,r) = 2^{1-\frac{1}{p'r'}}\,\pi\left(\frac{r}{pr'}\right)^{\frac{1}{pr'}}(r)^{1/r}(r')^{1/r'}B_1(r).$$

(3.2.7) is a consequence of Corollary 3.2.3 by using Hölder's inequality (for p and p') on (3.2.3) and then setting $u = \chi_E$ and $w(\gamma) = |\gamma|^{\alpha}$ for $\alpha = \frac{1}{r} + \frac{1}{pr'}$.

Naturally, (3.2.7) is just another of many special cases of (3.2.4). We mention it since it involves the measure of E as does Example 3.2.4a and a power weight factor of $\hat{f}$ as does Example 3.2.4b.

b. Letting $p \longrightarrow \infty$, $p \longrightarrow 1$, and $r \longrightarrow 1$, (3.2.7) gives rise to the following inequalities:

$$\int_E |\hat{f}(\gamma)|^2 d\gamma \le 2^{1/r}\pi(r)^{1/r}(r')^{1/r'}B_1(r)|E|^{1/r'} \times$$

$$\times \left(\int |tf(t)|^r dt\right)^{1/r}\left(\int_E |\gamma|\,|\hat{f}(\gamma)|^r d\gamma\right)^{1/r};$$

$$\int_E |\hat{f}(\gamma)|^2 d\gamma \le 2\pi r B_1(r)\left(\int |tf(t)|^r dr\right)^{1/r}\left(\int_E |\gamma|^{1+\frac{r}{r'}}|\hat{f}(\gamma)|^r d\gamma\right)^{1/r};$$

$$\int_E |\hat{f}(\gamma)|^2 d\gamma \le 2\pi\int |tf(t)|dt\int_E |\gamma\hat{f}(\gamma)|d\gamma.$$

3.3. LOCAL UNCERTAINTY PRINCIPLE INEQUALITIES ON $\mathbb{R}^d$

Instead of pursuing the method of Section 3.2 where Hölder, Hardy and Fourier transform inequalities were involved in that order, we shall state a weighted Fourier transform norm inequality (Theorem 3.3.1) particularly suited to making local estimates of spectral energy. The resulting method to obtain local uncertainty principle inequalities is first to use this weighted norm inequality Theorem 3.3.1 to estimate $\|\hat{f}\|^2_{2,u}$, and then to implement Hölder's inequality in the usual ways. We omit statements of the uncertainty principle inequalities which follow from this procedure.

3.3.1. *Theorem* [BH2, Theorem 4.3]. Given $1 < r \le 2$ and non-negative radial weights $u, v \in L^1_{loc}(\mathbb{R}^d)$. Suppose $v^{1-r'} \in L^1_{loc}(\mathbb{R}^d\backslash B(0,T))\backslash L^1(\mathbb{R}^d)$ for each $T > 0$. Assume

$$\sup_{s>0}\left(\int_0^s x^{d+1}\,u(\tfrac{x}{\pi})dx\right)^{1/2}\left(\int_0^{1/s} x^{d-1+r'}v(x)^{1-r'}dx\right)^{1/r'} = K_1 < \infty$$

and

$$\sup_{s>0}\left[\int_s^\infty x^{d-1}\ u(\frac{x}{\pi})dx\right]^{1/2}\left[\int_{1/s}^\infty x^{d-1}v(x)^{1-r'}dx\right]^{1/r'} = K_2 < \infty.$$

There is a constant $C > 0$ such that

$$\forall f \in L_v^r(\mathbb{R}^d),\ \|\hat{f}\|_{2,u} \le C\|f\|_{r,v}, \tag{3.3.1}$$

and C can be chosen as

$$C = 2\pi^{-(d-1)/2}\ \omega_{d-1}^{\frac{1}{2}+\frac{1}{r'}}(r)^{1/2}(r')^{1/r'}(K_1 + K_2).$$

The integrability condition on $v^{1-r'}$ allows us to obtain the inequality (3.3.1) for each $f \in L_v^r(\mathbb{R}^d)$, e.g., Appendix A.2. If supp u is compact and $v^{1-r'}$ is integrable off of a neighborhood of the origin then $K_2 < \infty$, and it is for this reason that Theorem 3.3.1 can be used to obtain local inequalities.

4. The Classical Uncertainty Principle and Wavelet Theory

4.1. WAVELET BASES AND WEYL-HEISENBERG FRAMES

The first problem of wavelet theory is to construct best possible orthonormal bases $\{\psi_n\}$ of $L^2(\mathbb{R})$. "Best possible" means that each ψ_n should be as smooth as possible and should have controllable, preferably compact, support. It also means that each basis should be sparse and "localizable", a notion which indicates that local changes or fine tuning in a signal can be made by adjustments to a small number of basis elements. The ψ_n are wavelets, cf., Remark 4.1.2c.

More general notions to effect decompositions are implemented when orthonormal bases are too complicated for applicability or too restrictive, e.g., Section 4.1.3, cf., Section 1.2.5d.

4.1.1. *Theorem* (Daubechies [D2]). For each $r \ge 1$ there is a compactly supported function $\psi \in C^{(r)}(\mathbb{R})$ (the space of r-times continuously differentiable functions) with the property that $\{\psi_{m,n}\} \subseteq L^2(\mathbb{R})$ is an orthonormal basis of $L^2(\mathbb{R})$ where

$$\forall\ m,n \in \mathbb{Z},\quad \psi_{m,n}(t) = 2^{m/2}\psi(2^m t-n). \tag{4.1.1}$$

4.1.2. *Remark.* a. Theorem 4.1.1 is difficult to prove. The first result of this type is due to Meyer [M1]. The Meyer "analyzing wavelet" ψ is an element of $\mathcal{S}(\mathbb{R})$ for which supp $\hat{\psi}$ is compact; and the resulting orthonormal basis $\{\psi_{m,n}\}$ is an unconditional basis of Sobolev spaces, Besov spaces, etc. Related results are due to Lemarié and Battle.

The notion of "multiresolution analysis", due to S. Mallat, is the most important idea associated with the difficult constructions of "analyzing wavelets" ψ which give rise to bases $\{\psi_{m,n}\}$ defined by (4.1.1), e.g., [M2] or the forthcoming book by Meyer. The original constructions did not benefit from this unifying and underlying concept. Intuitively a system of dilations and translations defined by (4.1.1) can be thought of as a deblurring mechanism. For example, if $f \in L^2(\mathbb{R})$ then

$$f_M = \sum_{m \le M} \sum_n (f, \psi_{m,n}) \psi_{m,n}$$

is a blurred vision of f in the following sense. If $\operatorname{supp} \psi \subsetneq [-\frac{1}{2}, \frac{1}{2}]$ then $\operatorname{supp} \psi_{m,n} \subsetneq [n2^{-m}-2^{-(m+1)}, n2^{-m}+2^{-(m+1)}] = I_{m,n}$ and $|I_{m,n}| = 2^{-m}$; and so if we add the $M+1$ term to f_M we have the effect of bringing to light the behavior of f at intervals of length $2^{-(M+1)}$, and thereby deblurring f_M.

b. Theorem 4.1.1 and the other results mentioned in part a have analogues in $\mathbb{R}^d$.

c. For a given $\psi \in L^2(\mathbb{R})$, the set $\{\psi_{m,n}\}$ defined by (4.1.1) is called an affine system because of the role of the underlying $ax+b$ group cf., Section 4.1.3d where the Heisenberg group plays an analogous role for the Weyl-Heisenberg system. "Analyzing wavelets" ψ for affine systems satisfy the "wave condition" $\int \psi(t)dt = 0$. Consequently, the corresponding wavelets $\psi_{m,n}$ (defined by (4.1.1)) are more appropriately and usually called "wavelets" than are the elements of the Weyl-Heisenberg system.

4.1.3. *Definitions/Representation.* a. Let H be a separable Hilbert space with inner product $(\ldots,\ldots)$. A sequence $\{\psi_n\} \subseteq H$ is a frame for H with frame bounds A and B if

$$\exists\, A,B > 0 \text{ such that } \forall f \in H,$$

$$A\|f\|^2 \le \sum |(f,\psi_n)|^2 \le B\|f\|^2.$$

b. Orthonormal bases in H are bounded unconditional bases which, in turn, are frames.

c. For a given frame $\{\psi_n\}$, the S-operator is the map $S : H \longrightarrow H$, $f \longmapsto \sum (f,\psi_n)\psi_n$. It is a basic fact that $S : H \longrightarrow H$ is an isomorphism (linear bijective topological isomorphism), from which we obtain the frame representation,

$$\forall f \in H, \; f = \sum (f,\psi_n) S^{-1}\psi_n = \sum (S^{-1}f, \psi_n)\psi_n. \tag{4.1.2}$$

d. Given $g \in L^2(\mathbb{R})$ and $a, \alpha > 0$; we define the Weyl-Heisenberg system,

$$\forall m,n \in \mathbb{Z}, \; \psi_{m,n}(t) = \psi_g(t; ma, n\alpha) = g(t-ma)e^{2\pi i t n\alpha}, \tag{4.1.3}$$

cf., Section 1.2.5a,c for $c = 0$. If $\{\psi_{m,n}\}$ is a frame it is designated a Weyl-Heisenberg frame for $L^2(\mathbb{R})$.

4.1.4. *Theorem* [D1], cf., [HW]. Given $g \in L^2(\mathbb{R})$ and $a > 0$. Assume

$$(4.1.4) \qquad 0 < A = \operatorname*{ess\,inf}_{t\in\mathbb{R}} \sum |g(t-ma)|^2 \le \operatorname*{ess\,sup}_{t\in\mathbb{R}} \sum |g(t-ma)|^2 = B < \infty$$

and

$$(4.1.5) \qquad \lim_{\alpha\to 0} \sum_{k\neq 0} \beta\left(\frac{k}{\alpha}\right) = 0,$$

where

$$\beta(s) = \operatorname*{ess\,sup}_{t\in\mathbb{R}} \sum |g(t-ma)||g(t-s-ma)|.$$

Then there is $\alpha_0 > 0$ such that for each $\alpha \in (0,\alpha_0)$, $\{\psi_g(t;ma,n\alpha)\}$ is a Weyl-Heisenberg frame for $L^2(\mathbb{R})$ with frame bounds

$$\alpha^{-1}A - \alpha^{-1}\sum_{k\neq 0}\beta\left(\frac{k}{\alpha}\right) \quad \text{and} \quad \alpha^{-1}B + \alpha^{-1}\sum_{k\neq 0}\beta\left(\frac{k}{\alpha}\right).$$

4.1.5. *Remark*. a. Condition (4.1.4) is a necessary condition in order that $\{\psi_g(t;ma,n\alpha)\}$ be a Weyl-Heisenberg frame. The sufficient condition (4.1.5) has been the subject of an important analysis in terms of Wiener-type spaces by D. Walnut, cf., Section 4.2.2d.

b. The relative merits of the translations/modulations of (4.1.3) and of the translations/dilations of (4.1.1) have been intensely editorialized and analyzed.

4.2. BALIAN'S *THEOREM*

4.2.1. *Theorem* (Balian). Given $g \in L^2(\mathbb{R})$ and $a,\alpha > 0$ for which $a\alpha = 1$. if $\{\psi_g(t;ma,n\alpha)\}$ is a Weyl-Heisenberg frame then

$$(4.2.1) \qquad tg(t) \notin L^2(\mathbb{R}) \quad \text{or} \quad \gamma\hat{g}(\gamma) \notin L^2(\hat{\mathbb{R}}),$$

cf., Section 4.2.5.

4.2.2. *The Zak Transform: Definition/Remarks.*

a. Given $g \in L^2(\mathbb{R})$ and $a,\alpha > 0$. The Zak transform of g is the function,

$$Z(g)(t,\gamma) = a^{1/2}\sum g(t-ka)e^{2\pi i k\gamma/\alpha},$$

defined on the rectangle $R_{a,\alpha} = [-a/2,a/2)\times[-\alpha/2,\alpha/2)$. $Z(g)$ extends to a function on $\mathbb{R}\times\hat{\mathbb{R}}$ satisfying the quasi-double periodicity condition,

$$\forall\ m,n \in \mathbb{Z},\ Z(g)(t+ma,\gamma+n\alpha) = Z(g)(t,\gamma)e^{2\pi i m\gamma/\alpha}.$$

Properties of the Zak transform are exposited in [D1; DGM; HW], and A.J.E.M. Janssen has made some of the most important recent contributions.

b. The Zak transform is a unitary (linear bijective isometry) map, $Z : L^2(\mathbb{R}) \longrightarrow L^2(R_{a,\alpha})$.

c. Given $g \in L^2(\mathbb{R})$ and $a,\alpha > 0$. If Zg is continuous on $\mathbb{R}\times\hat{\mathbb{R}}$ then Zg has a zero in $R_{a,\alpha}$. Examples of continuous Zak transforms Zg on $\mathbb{R}\times\hat{\mathbb{R}}$ are provided by $g \in C_c(\mathbb{R})$. C. Heil has proved that a large class of functions g with continuous Zak transform Zg on $\mathbb{R}\times\hat{\mathbb{R}}$ is the original Segal algebra (due to Wiener),

$$W(\mathbb{R}) = \{f \in L^1(\mathbb{R}) \cap C_b(\mathbb{R}) : \sum \|f\|_{\infty,[n,n+1]} < \infty\},$$

where $\|\ldots\|_{\infty,[n,n+1]}$ is the usual L^∞-norm on the interval $[n,n+1]$. Clearly, the elements of $W(\mathbb{R})$ vanish at $\pm\infty$ and belong to $L^2(\mathbb{R})$. $W(\mathbb{R})$ is also the first of the Wiener-type Banach spaces mentioned in Remark 4.1.5a defined locally and then globally; in the case of $W(\mathbb{R})$ the local norms are $\|\ldots\|_{\infty,[n,n+1]}$ and the global norm is $\|\ldots\|_1$.

d. If $g \in L^2(\mathbb{R})$ and $a\alpha = 1$ then $\{\psi_g(t;ma,n\alpha)\}$ is a Weyl-Heisenberg frame with frame bounds A and B if and only if $0 < A \le |Z(g)(t,\gamma)|^2 \le B < \infty$ a.e. on $R_{a,\alpha}$, cf., part c and Section 1.2.5d.

4.2.3. *Example.* Given $g = \chi_{[0,1)}$ and $a = \alpha = 1$. Define $\psi_{m,n}(t) = \psi_g(t;m,n)$ as in (4.1.3). Then Zg is defined a.e. on $\mathbb{R}\times\hat{\mathbb{R}}$ by the property,

$$\forall n,\quad \forall t \in (n,n+1),\quad \text{and}\quad \forall\gamma,\quad Z(g)(t,\gamma) = e^{2\pi i n\gamma}.$$

Since $|Z(g)(t,\gamma)| = 1$ a.e. on $R_{1,1}$ it is easy to check that $\{\psi_{m,n}\}$ is an orthonormal basis of $L^2(\mathbb{R})$, and hence it is a Weyl-Heisenberg frame. Further, $tg(t) \in L^2(\mathbb{R})$ and g', which exists a.e., is an element of $L^2(\mathbb{R})$, cf., (4.2.1). In this example, $\gamma\hat{g}(\gamma) \notin L^2(\mathbb{R})$ which corroborates Balian's theorem. However, this function g provides a counterexample to the proposed "proof" of Balian's theorem for g' that we've reproduced in Section 4.2.4.

With regard to the previous observation about g' and $\gamma\hat{g}(\gamma)$, recall that if $f \in L^2(\mathbb{R})$ and $\gamma\hat{f}(\gamma) \in L^2(\hat{\mathbb{R}})$ then f' exists a.e., is an element of $L^2(\mathbb{R})$, and $f'(t) = (2\pi i\gamma\hat{f}(\gamma))^\vee(t)$ a.e.

4.2.4. *Analysis.* The following is a plan to verify Balian's theorem for f, $tf(t)$, $f' \in L^2(\mathbb{R})$, $a\alpha = 1$, and $\{\psi_f(t;ma,n\alpha)\}$.

Step 1. Prove that $\partial_1 Zf, \partial_2 Zf \in L^2_{loc}(\mathbb{R}^2)$.

Step 2. Define $Z_r(f)(t,\gamma)$ to be the continuous mean of Zf over a square of side $2r$ centered at (t,γ).

Step 3. Assume $\{\psi_f(t;ma,n\alpha)\}$ is a Weyl-Heisenberg frame with frame bounds A and B. Use Step 1 to prove that $A^{1/2}/2 \le |Z_r(f)(t,\gamma)| \le 2B^{1/2}$ on a specified square R containing $R_{1,1}$.

Step 4. Obtain a contradiction by considering $\log Z_r(f)$.

Most details for these steps are correct and can be found in the literature. The above outline would give the result since the hypotheses $tf(t),\ f' \in L^2(\mathbb{R})$ allow a verification of Step 1. (Recall that g of Example 4.2.3 satisfies $tg(t),\ g' \in L^2(\mathbb{R})$.) The problem with this outline is part of the proof of Step 3 where the following calculation is made for $Z = Z(f)$:

$$Z(t',\gamma') - \frac{1}{4r^2}\int_I\!\!\int Z(t'',\gamma'')dt''d\gamma''$$

$$= -\frac{1}{4r^2}\int_I\!\!\int (Z(t'',\gamma'')-Z(t',\gamma'')+Z(t',\gamma'')-Z(t',\gamma'))dt''d\gamma''$$

$$= -\frac{1}{4r^2}\int_I\!\!\int\left(\int_{t'}^{t''}\partial_1 Z(t''',\gamma'')dt''' + \int_{\gamma'}^{\gamma''}\partial_2 Z(t',\gamma''')d\gamma'''\right)dt''d\gamma'',$$

where $I = \{(t'',\gamma'') : |t''-t'| \le r \text{ and } |\gamma''-\gamma'| \le r\}$. Generally, this calculation fails since the fundamental theorem of calculus is not applicable. For example, in the case of Example 4.2.3,

$$\int_{t'}^{t''}\partial_1 Z(g)(t''',\gamma'')dt''' = 0$$

and

$$Z(g)(t'',\gamma'') - Z(g)(t',\gamma'') = e^{2\pi i n''\gamma''} - e^{2\pi i n'\gamma''};$$

and this causes problems on R when $|n''-n'| > 0$.

4.2.5. *Discussion of Balian's Theorem.* Balian's original treatment was in terms of orthonormal bases [Bal], cf., [HJJ]. An ingenious simple proof is due to Battle [Bat], and, using Battle's idea, it has been possible for several groups to verify Theorem 4.2.1, either by formal calculation or correct proof.

Using the uncertainty principle, Theorem 1.3.3, and assuming $tg(t) \in L^2(\mathbb{R})$ and $\gamma\hat{g}(\gamma) \in L^2(\hat{\mathbb{R}})$, Battle shows that $g = 0$ so that the Weyl-Heisenberg system can not give rise to an orthonormal basis. The uncertainty principle is invoked in the sense that he calculates $E_g(i[A,B]) = 0$, noting that $E_g(i[A,B]) = \|g\|_2^2$.

C. Heil has observed, in light of his result quoted in Section 4.2.2, that Balian's theorem is immediate with the further hypothesis, $g \in W(\mathbb{R})$.

Finally, we note that Balian's theorem leads one to quantify further the notion of functions g being far away from the Gaussian such as reflected by condition (4.2.1). In this regard, we mention DeBruijn's theorem to the effect that if (1.1.1) is almost an equality for a function g then g is almost equal to a function of the form (1.1.2) [De, Theorem 3.3].

4.2.6. *Strong Uncertainty for Weyl-Heisenberg Systems.* Given $g \in L^2(\mathbb{R})\setminus\{0\}$ and $a,\alpha > 0$. Define the Weyl-Heisenberg system $\psi_{m,n}(t) = \psi_g(t;ma,n\alpha)$, $m,n \in \mathbb{Z}$. For any fixed $(t_0,\gamma_0) \in \mathbb{R}\times\hat{\mathbb{R}}$, we have the strong uncertainty property

$$\sup_{m,n}\|(t-t_0)\psi_{m,n}(t)\|_2\|(\gamma-\gamma_0)\hat{\psi}_{m,n}(\gamma)\|_2 = \infty, \tag{4.2.2}$$

cf., (4.3.2).

To verify (4.2.2) we first observe by basic function theory associated with Jensen's or Carleman's theorem and the Paley-Wiener theorem that if g (resp., $\hat{g}$) vanishes on a half-line then $\hat{g}$ (resp., g) cannot vanish on intervals without being identically zero. Next, consider the estimate,

$$\begin{aligned}
&\|(t-t_0)\psi_{m,n}(t)\|_2^2\|(\gamma-\gamma_0)\hat{\psi}_{m,n}(\gamma)\|_2^2 \\
&\geq \int((t-t_0)+ma)^2|g(t)|^2dt \int_{\gamma_0}^{\infty}((\gamma-\gamma_0)+n\alpha)^2|\hat{g}(\gamma)|^2d\gamma \\
&\geq (n\alpha)^2\int((t-t_0)+ma)^2|g(t)|^2dt \int_{\gamma_0}^{\infty}|\hat{g}(\gamma)|^2d\gamma,
\end{aligned} \tag{4.2.3}$$

where the second inequality is valid for $n \geq 0$. Thus, if supp g is contained in a half-line then (4.2.2) follows from (4.2.3). If supp g is not contained in a half-line then the obvious adjustment of (4.2.3) yields (4.2.2).

4.3. BOURGAIN'S THEOREM

4.3.1. *Notation for Expected Values.* Given $\psi \in L^2(\mathbb{R})$ and consider the affine system $\{\psi_{m,n}\}$ defined in (4.1.1). For each m,n the expected values $(t_{m,n},\gamma_{m,n}) \in \mathbb{R}\times\hat{\mathbb{R}}$ are

$$t_{m,n} = \int t|\psi_{m,n}(t)|^2dt \quad \text{and} \quad \gamma_{m,n} = \int\gamma|\hat{\psi}_{m,n}(\gamma)|^2d\gamma,$$

cf., the end of Section 1.2.1 for the origin of this terminology from quantum mechanics.

4.3.2. *Weak Uncertainty for Affine Systems.* Given $\psi \in L^2(\mathbb{R})$ and consider the notation from Section 4.3.1. We have the weak uncertainty property,

$$\sup_{m,n} \|(t-t_{m,n})\psi_{m,n}(t)\|_2 \|(\gamma-\gamma_{m,n})\hat{\psi}_{m,n}(\gamma)\|_2 < \infty, \tag{4.3.2}$$

cf., (4.2.2). The inequality (4.3.2) is a consequence of the equalities

$$\begin{aligned} \|(t-t_{m,n})\psi_{m,n}(t)\|_2 &= 2^{-m}\|(t-t_{0,0})\psi(t)\|_2 \\ \|(\gamma-\gamma_{m,n})\hat{\psi}_{m,n}(\gamma)\|_2 &= 2^{m}\|(\gamma-\gamma_{0,0})\hat{\psi}(\gamma)\|_2, \end{aligned} \tag{4.3.3}$$

which, in turn, follow from easy calculations.

4.3.3. *Problem.* "Weak uncertainty" for wavelet bases such as the affine system described in Theorem 4.1.1 provides simultaneous control of time and frequency information among all the basis elements. It is natural to ask how precise this simultaneity can be. To quantify this question we ask specifically if there is an orthonormal basis $\{\psi_n\}$ of $L^2(\mathbb{R})$ having expected values $(t_n, \gamma_n) \in \mathbb{R}\times\hat{\mathbb{R}}$ so that

$$\sup_n \|(t-t_n)\psi_n(t)\|_2 < \infty \quad \text{and} \quad \sup_n \|(\gamma-\gamma_n)\hat{\psi}_n(\gamma)\|_2 < \infty.$$

Because of (4.3.3) any such refinement of (4.3.2) would have to go outside the realm of affine systems. In any case the following result gives a strong solution to this problem.

4.3.4. *Theorem* (Bourgain [Bo]). For every $\varepsilon > 0$ there is an orthonormal basis $\{\psi_n\}$ of $L^2(\mathbb{R})$ having expected values $(t_n, \gamma_n) \in \mathbb{R}\times\hat{\mathbb{R}}$ and satisfying the inequalities,

$$\sup_n \|(t-t_n)\psi_n(t)\|_2 < \frac{1}{2\sqrt{\pi}} + \varepsilon \quad \text{and} \quad \sup_n \|(\gamma-\gamma_n)\hat{\psi}_n(\gamma)\|_2 < \frac{1}{2\sqrt{\pi}} + \varepsilon.$$

Thus,

$$\forall\ n,\ 1 = \|\psi_n\|_2^2 \le 4\pi\|(t-t_n)\psi_n(t)\|_2\|(\gamma-\gamma_n)\hat{\psi}_n(\gamma)\|_2 < 4\pi\left(\frac{1}{2\sqrt{\pi}} + \varepsilon\right)^2.$$

5. Spectrum Estimation and the Wiener-Wintner Theorem

5.1. SPECTRUM ESTIMATORS AND THE UNCERTAINTY PRINCIPLE

5.1.1. *Problem.* The spectrum estimation problem is to clarify and quantify the statement: find periodicities in a signal $f(t)$ recorded over a fixed time interval $[-T,T]$. In more picturesque language, we want to filter the noise from the incoming signal f in order to determine the intelligent message (periodicities) therein.

5.1.2. *Definition.* In order to quantify this problem we introduce the following mathematical setting.

Let $f(t,\alpha)$ be a stationary stochastic process (SSP), where the sample functions $f(\cdot,\alpha)$ on $\mathbb{R}$ are indexed by α in the underlying probability space X. (For our purposes, an SSP f is characterized the conditions that the expected value $E\{f(t)\}$ is constant, by $E\{f(t+h)\overline{f(u+h)}\} = E\{f(t)\overline{f(u)}\}$ for all $h, t, u \in \mathbb{R}$, and $\lim_{h\to 0} E\{|f(h) - f(0)|^2\} = 0.$)

The autocorrelation of the SSP f is the continuous positive definite function

$$R_f(t) = E\{f(t+u)\overline{f(u)}\};$$

and the power spectrum of f is the positive measure S_f for which $\hat{R}_f = S_f$ (a distributional Fourier transform).

5.1.3. *Definition/Remark.* Given $f \in L^2_{loc}(\mathbb{R}^d)$ and define

$$\forall T > 0, \quad P_{f,T} = \frac{1}{|B(T)|} (f\chi_{B(T)})*(f\chi_{B(T)})^{\sim}$$

so that $P_{f,T} \in L^1_{loc}(\mathbb{R}^d) \subseteq M(\mathbb{R}^d)$. Suppose that there is a continuous positive definite function P_f for which $\lim_{T\to\infty} P_{f,T} = P_f$ in the vague topology $\sigma(M(\mathbb{R}^d), C_c(\mathbb{R}^d))$. Then $P_f \in L^\infty(\mathbb{R}^d)$ is the autocorrelation of f and $\hat{P}_f = \mu_f \in M_{b+}(\hat{\mathbb{R}}^d)$ is the power spectrum of f.

In order to reconcile the two apparently different definitions on $\mathbb{R}$ of both autocorrelation and power spectrum we shall assume that

(5.1.1) $$\forall t \in \mathbb{R}, \quad \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(t+u,\alpha)\overline{f(u,\alpha)}\,du = R_f(t)$$

converges in measure. Processes satisfying condition (5.1.1) are correlation ergodic processes; and verification of correlation ergodicity of $f(t,\alpha)$ requires knowledge of its fourth order moments.

5.1.4. *Definition/Fact.* Given $v \in L^d(\mathbb{R})$ and suppose f is an SSP for which each sample function $f(\cdot,\alpha)$ is an element of $L^\infty(\mathbb{R})$. Then

$$S_f(\gamma,\alpha) = \left|\int f(t,\alpha)v(t)\, e^{-2\pi i t\gamma}dt\right|^2$$

is the periodogram associated with the process f and data window v.

If f is a real SSP and $v = v^{\sim}$ then it is not difficult to verify that

(5.1.2) $$E\{S_v(\gamma)\} = S*V^2,$$

where $\hat{v} = V$. Further if $\{V_T^2 : T > 0\}$ is an L^1-approximate identity and $\hat{v}_T = V_T$ then

(5.1.3) $$\lim_{T\to\infty} E\{S_{v_T}(\gamma)\} = S$$

in the canonical weak topology.

Because of (5.1.3) we can assert that $E\{S_{v_T}\}$ is an asymptotically unbiased estimator of S, thereby providing a relatively naive solution to our vaguely posed spectrum estimator problem.

A special case of the Bell Labs uncertainty principle is -

5.1.5. *Proposition* [LSP]. Given $T, \Omega > 0$. There is $c = c(T\Omega) \in (0,1)$ such that for each Ω-band limited function f, i.e., $\operatorname{supp} \hat{f} \subseteq [-\Omega,\Omega]$,

(5.1.4) $$\|f\chi_{(T)}\|_2 \le c(T\Omega)\|\hat{f}\|_2;$$

$c(T\Omega)^2$ is the largest eigenvalue of the operator BA, where $Af = f\chi_{(T)}$ and $Bg = (\hat{g}\chi_{(\Omega)})^{\vee}$. ($\chi_{(T)}$ is the characteristic function of $[-T,T]$.)

Once the values $c(T\Omega)$ are known, Proposition 5.1.5 answers the question: what is an upper bound of the energies $\|f\chi_{(T)}\|_2^2$ as f ranges through the Ω-band limited signals having a fixed finite energy? The classical uncertainty principle says that the variances of $|f|^2$ and $|\hat{f}|^2$ cannot both be small; Proposition 5.1.5 says that the energies of Ω-band limited signals and of their restrictions to $[-T,T]$ cannot be arbitrarily close, in spite of the Plancherel theorem.

5.1.6. *Calculation.* The uncertainty principle embodied in Proposition 5.1.5 allows us to quantify energy loss in spectrum estimators as the following calculation shows for the estimator $E\{S_v\}$. Besides the above hypotheses we add the realistic conditions that $\operatorname{supp} v \subseteq [-T,T]$, $\|v\|_2 = 1$, $v > 0$ on $(-T,T)$, and $\operatorname{supp} S \subseteq [-\Omega,\Omega]$ for fixed $T, \Omega > 0$. We compute

$$\|E\{S_v\}\|_2 = \|S * V^2\|_2 = \|S^{\vee}(v*v)\|_2$$
$$\le \|S^{\vee}(v*v)\,\frac{1}{v*v}\|_{L^2[-2T,2T]} = \|S^{\vee}\chi_{(2T)}\|_2$$
$$\le c(T\Omega)\ \|S\|_2.$$

One can make similar calculations involving weighted L^2-spaces (and therefore reminiscent of the classical uncertainty principle), where the uncertainty components are high resolution and precision of estimator, respectively [Gr].

5.2. DISCUSSION OF SPECTRUM ESTIMATION

5.2.1 *Deterministic Assumptions.* The path we have chosen in Section

5.1 to quantify the loosely posed Problem 5.1.1 is based on the following deterministic assumptions.

i. The signal f is defined on the product space $[-T,T]\times X$, where X is a probability space and f is the restriction to $[-T,T]\times X$ of some SSP g.

ii. The expectation of the periodogram $S_v(\gamma,\alpha)$ is known, where $\operatorname{supp} v \subseteq [-T,T]$ and $\|v\|_2 = 1$.

iii. The power spectrum S of f is uniquely determined.

Assumption iii is a theorem in many cases involving the experimentally reasonable hypothesis that $\operatorname{supp} S_g$ is compact, where S_g is the power spectrum of the SSP g which, in turn, is an extension of f defined on $[-T,T]\times X$. Mathematically, this assumption allows us to specify S in (5.1.2) and, ultimately, leads to the Beurling - Malliavin theory [B1].

On the other hand, Assumption iii is not universally accepted. In fact, the maximum entropy method (MEM) of spectrum estimation is based on a point of view opposite that of a uniquely determined power spectrum. MEM does not assert the existence of a unique power spectrum and then estimate it; instead, given f on $[-T,T]\times X$ or the autocorrelation R on $[-T,T]$, MEM models autocorrelation data outside $[-T,T]$ by maximizing a certain entropy integral.

5.2.2. *Statistical Assumptions*. Besides the questions surrounding Assumption iii(of Section 5.2.1), we must also note that Assumption i preempts all genuine statistical problems.

A different point of view from that of Section 5.2.1 is the following. We are given some sample paths of finite duration $[-T,T]$ corresponding to a specific experiment. The only stochastic process available may be an idealized process representing the potential output of some underlying mechanism. For example, in speech analysis the generation of a specific sound varies with time and person but is subject to statistical regularities, and this mechanism is the ultimate source of realistic spectrum estimation of frequencies corresponding to the sound.

5.2.3. *Remark*. Spectrum estimation is a multi-faceted, basic, and deep problem with points of view bordering on the philosophical and viable techniques ranging from sophisticated periodogram analysis to various "high resolution" methods such as MEM. We refer to [PI] and [L] for a scholarly presentation of diverse methods and a brilliant new insight, respectively.

The purpose of our discussion has been to present the remaining parts of Section 5 which provide theorems susceptible to transition and interpretation as spectrum estimation algorithms.

5.3. SIGNAL DETERMINISTIC ESTIMATORS ON $\mathbb{R}^d$

Suppose the data characterizing a given signal f is known. In the following result, V can be thought of as a properly shaped window function so that the left side of (5.3.1) represents the power of f in the region supp V. Formula (5.3.1) provides a method for computing this power in terms of the known functions f and v. In practice, then, numerical estimates of the right side of (5.3.1) lead to a spectrum estimation algorithm.

5.3.1. *Theorem* [B3]. Given $f \in L^2_{loc}(\mathbb{R}^d)$ with autocorrelation P_f and power spectrum μ_f. Assume there is an increasing function $i(R)$ on $(0,\infty)$ for which $\sup_{|t|\le T} |f(t)| \le i(T)$ and $\lim_{T\to\infty} i(T)^2/T = 0$. Then

$$(5.3.1) \qquad \forall v \in C_c(\mathbb{R}^d), \quad \int |V(\gamma)|^2 d\mu_f(\gamma) = \lim_{T\to\infty} \frac{1}{|B(T)|} \int_{B(T)} |f*v(t)|^2\, dt,$$

where $\hat{v} = V$.

5.3.2. *Corollary*. Given the hypotheses of Theorem 5.3.1. If $v \in L^p(\mathbb{R}^d)$ and $f \in L^{p'}(\mathbb{R}^d)$, $1 \le p < \infty$, then the tempered distribution $\hat{v} = V$ is a well-defined element of $L^2_{\mu_f}(\mathbb{R}^d)$ and

$$\|V\|_{2,\mu_f} \le \|f\|_{p'} \, \|v\|_p.$$

5.4. AUTOCORRELATION DETERMINISTIC ESTIMATORS ON $\mathbb{R}^d$

Suppose the autocorrelation data for a signal is known and the signal itself is not explicitly known. This is a typical situation and one of the reasons autocorrelations are so important, e.g., [B1]. In the following result, V can be thought of as a properly shaped window function so that the left side of (5.4.1) represents the power of f in the region supp V. Since the norm constant in Proposition 5.4.1 is explicit and computable, we see that (5.4.1) provides a means of estimating an upper bound for the power of f in the region supp V.

5.4.1. *Proposition* [B3]. Given $\mu \in M_{b+}(\hat{\mathbb{R}}^d)$ for which $\mu^{\vee} = P \in L^{p'}(\mathbb{R}^d)$, $p \in [1,\infty]$. Then

$$\forall v \in L^1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d),$$

$$(5.4.1) \qquad \|V\|_{2,\mu} \le \|P\|_{p'}^{1/2}(\|v\|_1\|v\|_p)^{1/2} \le (\tfrac{1}{2}\|P\|_{p'}^{1/2}(\|v\|_1 + \|v\|_p),$$

where $\hat{v} = V$ and $(L^1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d), \|\cdots\|_1 + \|\cdots\|_p)$ is a Banach space.

5.4.2. *Corollary* [B3]. Given $d \ge 2$ and $1 \le p < 2d/(d+1)$. Then $\|\mu_{d-1}\|_{p'} < \infty$ and, for each $v \in L^1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$,

$$(5.4.2)\qquad \left[\int_{\Sigma_{d-1}} |V(\theta)|^2\, d\sigma_{d-1}(\theta)\right]^{1/2} \le (\tfrac{1}{2}\|\check{\mu}_{d-1}\|_{p'}^{1/2}(\|v\|_1 + \|v\|_p),$$

where $\hat{v} = V$.

5.4.3. *Remark.* The corollary is immediate from elementary properties of Bessel functions and Proposition 5.4.1, which itself follows from an elementary calculation. We stress the simplicity of (5.4.2) to compare it with the much deeper Tomas-Stein restriction theorem:

$$(5.4.3)\qquad \forall v \in L^1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d),$$
$$\left[\int_{\Sigma_{d-1}} |V(\theta)|^2\, d\sigma_{d-1}(\theta)\right]^{1/2} \le c(p)\ \|v\|_p,$$

where $\hat{v} = V$ and $1 \le p < 2(d+1)/(d+3)$. In (5.4.3) the constant $c(p)$ is not explicit, $p = 2(d+1)/(d+3)$ is largest possible, and the right hand norm is $\|v\|_p$. In (5.4.2) the constant is explicit and the values of p extend beyond $2(d+1)/(d+3)$, but the right hand norm is $\|v\|_1 + \|v\|_p$.

5.5. THE WIENER-WINTNER THEOREM IN $\mathbb{R}^d$

Underlying the results in Section 5.3 and 5.4 is the question: given $\mu \in M_+(\hat{\mathbb{R}}^d)$, does there exist $f \in L^2_{loc}(\mathbb{R}^d)$ such that $\hat{P}_f = \mu$? The Wiener-Wintner theorem provides an answer to this question and a means of reconstructing signals f on $\mathbb{R}^d$ corresponding to a given power spectrum.

5.5.1. *Preliminaries.* Given $\mu \in M_{b+}(\hat{\mathbb{R}}^d)$ and let δ_ω be the Dirac measure supported by $\{\omega\}$. It is well-known that there is a sequence $\{\mu_n\} \subseteq M_{b+}(\text{supp}\ \mu)$ of positive discrete measures,

$$\mu_n = \sum_{j=1}^{N_n} a_{j,n}\, \delta_{\omega_{j,n}}, \quad a_{j,n} > 0,$$

such that $\{\omega_{j,n} :\ j = 1, \cdots, N_n\} \subseteq \text{supp}\ \mu$ for each n,

$$(5.5.1)\qquad \lim_{n\to\infty} \langle\mu_n, 1\rangle = \langle\mu, 1\rangle,$$

and $\lim_{n\to\infty} \mu_n = \mu$ in the (vague) topology $\sigma(M_b(\hat{\mathbb{R}}^d), C_c(\hat{\mathbb{R}}^d))$. Actually, (5.5.1) and the $\sigma(M_b(\hat{\mathbb{R}}^d), C_c(\hat{\mathbb{R}}^d))$ convergence allow us to conclude that $\lim_{n\to\infty} \mu_n = \mu$ in the "Levy" topology $\sigma(M_b(\hat{\mathbb{R}}^d), C_b(\hat{\mathbb{R}}^d))$.

For a given $\mu \in M_{b+}(\hat{\mathbb{R}}^d)$ and sequence $\{\mu_n\} \subseteq M_{b+}(\text{supp}\ \mu)$ as above, we define

$$f_n(t) = \sum_{j=1}^{N_n} a_{j,n}^{1/2} e^{2\pi i t\cdot\omega_{j,n}}$$

so that

$$\|f_n\|_\infty \le \sum_{j=1}^{N_n} a_{j,n}^{1/2} \le N_n \sup_{1\le j\le N_n} a_{j,n}^{1/2}.$$

5.5.2. *Lemma/Notation.* The proof of the following result depends on elementary properties of Bessel functions.

Lemma. For each n,

$$\lim_{T\to\infty} \frac{1}{|B(T)|} \int_{B(T)} f_n(t+x)\overline{f_n(x)}dx = \check{\mu}_n(t)$$

uniformly on $\mathbb{R}^d$.

Notation. We use the *Lemma* to define a specific sequence $\{T_n\}$ and a specific function f in the following way. From the uniform convergence we know that

$$\forall n \ge 1,\ \exists A_n \ge A_{n-1} \text{ such that } \forall t \in \mathbb{R}^d \text{ and } \forall T \ge A_n,$$
$$\left| \frac{1}{|B(T)|} \int_{B(T)} f_n(t+x)\overline{f_n(x)}dx - \check{\mu}_n(t) \right| < \frac{1}{2^{n+1}}.$$

We set $T_n = (A_1+1)(A_2+2)\cdots(A_n+n)$ so that $T_n \ge n!$ and the sequences $\{T_n\}$, $\{T_{n+1}/T_n\} = \{A_{n+1}+n+1\}$, and $\{T_{n+1}-T_n\} = \{(A_1+1)\cdots(A_n+n)(A_{n+1}+n)\}$ increase to infinity. For this sequence $\{T_n\}$ and for $\{f_n\}$ defined above we define f on $\mathbb{R}^d$ by setting $f(t) = f_n(t)$ for $T_n<|t|<T_{n+1}$, $n \ge 1$, and letting $f(t) = 0$ for $|t| < T_1$. Clearly, $f \in L^\infty_{loc}(\mathbb{R}^d)$ and the values of $f(t)$ for $|t| = T_n$ are not important.

The proof of the following result is long and intricate.

5.5.3. *Theorem* [B3]. Given $\mu \in M_{b+}(\hat{\mathbb{R}}^d)$ with corresponding functions $\{f_n\}$ and $f \in L^\infty_{loc}(\mathbb{R}^d)$. Assume there is $C > 0$ such that for all n

(5.5.2) $$\|f_n\|_\infty^2 \le CT_n.$$

Then, for each $t \in \mathbb{R}^d$,

$$\lim_{T\to\infty} \frac{1}{|B(T)|} \int_{B(T)} f(t+x)\overline{f(x)}dx = \check{\mu}(t).$$

5.5.4. *Remark.* The original proof of Theorem 5.5.3 by Wiener and Wintner on $\mathbb{R}$ is cryptic and there have been important contributions on $\mathbb{R}$ by J. Bass and Bertrandias.

There is left unanswered the problem of characterizing those μ for which $f \in L^\infty(\mathbb{R}^d)$. This problem is not yet solved in case $d = 1$.

A. Closure Theorems

A.1. BI-SOBOLEV SPACES

A.1.1. *Definition.* Given integers $m,n \ge 0$ and $1 \le p \le \infty$. The <u>Sobolev space</u> $L^p_m = L^p_m(\mathbb{R}^d)$ is the Banach space of functions $f \in L^p(\mathbb{R}^d)$ for which

$$\|f\|_{m,p} = \sum_{|\alpha|\le m} \|\partial^\alpha f\|_p < \infty.$$

The <u>weighted space</u> $L^p_{o,n} = L^p_{o,n}(\mathbb{R}^d)$ is the Banach space of functions $f \in L^p(\mathbb{R}^d)$ for which

$$\|f\|_{p,n} = \sum_{|\beta|\le n} \|t^\beta f(t)\|_p < \infty.$$

The <u>Bi-Sobolev space</u> $L^p_{m,n} = L^p_{m,n}(\mathbb{R}^d)$ is the Banach space of functions $f \in L^p_m \cap L^p_{o,n}$ for which

$$\|f\|_{m,p,n} = \|f\|_{m,p} + \|f\|_{p,n} < \infty. \tag{A.1.1}$$

A.1.2. *Theorem.* Given integers $m,n \ge 0$. $C_c^\infty(\mathbb{R}^d)$ is dense in the Hilbert space $(L^2_{m,n}, \|\cdots\|_{m,2,n})$ with inner product

$$[f,g] = \sum_{|\alpha|\le m} (\partial^\alpha f, \partial^\alpha g) + \sum_{|\beta|\le n} (t^\beta f, t^\beta g),$$

where $(\cdots,\cdots)$ is the usual inner product on $L^2(\mathbb{R}^d)$.

<u>Proof</u>: a. It is well-known that $C_c^\infty(\mathbb{R}^d)$ is dense in L^p_m, $1 \le p < \infty$. To see this, choose $f \in L^p_m$, let $\{h_j\} \subseteq C_c^\infty(\mathbb{R}^d)$ be an L^1-approximate identity with each $\operatorname{supp} h_j \subseteq B(0,1)$, and take $u_j \in C_c^\infty(\mathbb{R}^d)$ defined by $u_j(t) = u(t/j)$ where $0 \le u \le 1$ and $u = 1$ on $B(0,1)$.

Fix $|\alpha| \le m$. Not only does $\partial^\alpha(f*h_j) = f*\partial^\alpha h_j$ but, by integration by parts, $\partial^\alpha(f*h_j) = (\partial^\alpha f)*h_j$. Consequently, we can apply Young's theorem to obtain $\|\partial^\alpha(f*h_j)\|_p \le \|\partial^\alpha f\|_p \|h_j\|_1 \le K\|\partial^\alpha f\|_p$. Thus, each $f*h_j$ is an element of L^p_m, as is each $u_j(f*h_j)$.

The desired density will follow from the triangle inequality once we prove

$$\lim_{j\to\infty} \|\partial^{\alpha}[(f*h_j)(u_j - 1)]\|_p = 0 \tag{A.1.2}$$

for each $|\alpha| \leq m$. To this end we first use Leibniz's formula for the estimate,

$$\begin{aligned}\|\partial^{\alpha}[(f*h_j)(u_j - 1)]\|_p &\leq \|(u_j - 1)\partial^{\alpha}(f*h_j)\|_p \; + \\ &\sum_{\substack{\beta\leq\alpha \\ |\beta|\geq 1}} |C_{\alpha\beta}| j^{-|\beta|} \|\partial^{\alpha-\beta}(f*h_j)(t)\partial^{\beta}u|_{(t/j)}\|_p.\end{aligned} \tag{A.1.3}$$

The dominated convergence theorem and Young's theorem allow us to prove that the first term on the right side of (A.1.3) tends to 0 as $j\to\infty$. Young's theorem and the fact that $\lim_{j\to\infty} j^{-|\beta|}\|\partial^{\beta}u\|_{\infty} = 0$ for $|\beta| \geq 1$ show that the remaining terms on the right side of (A.1.3) tend to 0 as $j\to\infty$. (A.1.2) is proved.

This density in L^p_m can also be verified by an equicontinuity argument much like the one we give for $L^2_{o,n}$.

b. It is sufficient to prove that

$$\forall f \in L^2_{o,n}, \quad \lim_{j\to\infty} \|F_{\beta j}(f)\|_2 = 0 \tag{A.1.4}$$

for each $|\beta| \leq n$, where $F_{\beta j}(f) = F_j(f) = t^{\beta}(u_j(f*h_j) - f)$. To this end we first show that

$$\sup_j \|F_j(f)\|_2 = C(f) < \infty. \tag{A.1.5}$$

This is accomplished by the estimate,

$$\|F_j(f)\|_2 - \|t^{\beta}f(t)\|_2 \leq C(\beta) \; \|\partial^{\beta}(\hat{f}\hat{h}_j)\|_2$$

$$\leq \sum_{\gamma\leq\beta} |C_{\beta\gamma}| \|\partial^{\beta-\gamma}\hat{f}\,\partial^{\gamma}\hat{h}_j\|_2 \leq \sum_{\gamma\leq\beta} |C_{\beta\gamma}| \|\partial^{\beta-\gamma}\hat{f}\|_2 \sup_{\gamma\leq\beta} \|\partial^{\gamma}\hat{h}_j\|_{\infty},$$

and the fact, in the case h_j is the dilation $j^d h(jt)$, that

$$|\partial^{\gamma}\hat{h}_j(\lambda)| = |C(\gamma) \int h(u)u^{\gamma} \; j^{-|\gamma|} \; e^{-2\pi i(u/j)\cdot\lambda}du| \leq K(\gamma)j^{-|\gamma|}$$

since $\operatorname{supp} h$ is compact. The estimate used the Plancherel theorem and so we note the fact that the distribution $\partial^{\beta}(\hat{f}\hat{h}_j)$ is an element of $L^2(\mathbb{R}^d)$.

It is easy to check that the elements of $L^2_{o,n}$ having compact support are dense in $L^2_{o,n}$ and that $\{F_{\beta j}\}$ is contained in $\mathcal{L}(L^2_{o,n}, L^2(\mathbb{R}^d))$ for each $|\beta| \leq n$. Because of (A.1.5) we can invoke the uniform boundedness principle and obtain $\sup \|F_j\| = C < \infty$. Thus, $\{F_j\}$ is equicontinuous. On the other hand it is routine to check that $\lim_{j\to\infty} \|F_j(f)\|_2 = 0$ for compactly supported functions $f \in L^2_{o,n}$. This convergence on a dense subset of $L^2_{o,n}$ combined with the equicontinuity yield convergence on $L^2_{o,n}$, and the resulting limit $F(f)$ for $f \in L^2_{o,n}$ determines an element $F \in \mathcal{L}(L^2_{o,n}, L^2(\mathbb{R}^d))$. Therefore, (A.1.4) is obtained.

q.e.d.

A.1.3. *Remark.* Instead of defining $L^p_{m,n}$ to deal with closure questions for the uncertainty principle, we could define the weighted Sobolev space $L^p_{m,w}(\mathbb{R}^d)$ consisting of functions $f \in L^1_{loc}(\mathbb{R}^d)$ for which

$$\|f\|_{\dot{m},p,w} = \sum_{|\alpha|\leq m} \|\partial^\alpha f\|_{p,w(\alpha)} < \infty.$$

In case $d = 1$, $p = 2$, $m = 1$ (and so $\alpha = 0, 1$), and $w = (w(0), w(1))$ with $w(0)(t) = (1+t^2)$ and $w(1)(t) = 1$, we see that $L^2_{1,w}(\mathbb{R})$ is the Bi-Sobolev space $L^2_{1,1}(\mathbb{R})$.

A.2. L^p_v AND DENSE MOMENT SPACES

A.2.1. *Theorem* [BH2]. Given $v \in L^1_{loc}(\mathbb{R}^d)$ where $v > 0$ a.e. and choose $p \in (1,\infty)$.

a. If $h \in L^p_v(\mathbb{R}^d)$ annihilates $\mathcal{S}_o(\mathbb{R}^d) \cap L^p_v(\mathbb{R}^d)$ then h is a constant function.

b. $\overline{\mathcal{S}_o(\mathbb{R}^d) \cap L^p_v(\mathbb{R}^d)} = L^p_v(\mathbb{R}^d)$ or $L^p_v(\mathbb{R}^d) \subseteq L^1(\mathbb{R}^d)$.

c. If $v^{1-p'} \notin L^1(\mathbb{R}^d)$ then $\overline{\mathcal{S}_o(\mathbb{R}^d) \cap L^p_v(\mathbb{R}^d)} = L^p_v(\mathbb{R}^d)$.

A.2.2. *Remark.* a. Theorem A.2.1 requires some effort to prove, but is considerably easier if $L^{p'}_{v^{1-p'}}(\mathbb{R}^d) \subseteq \mathcal{S}'(\mathbb{R}^d)$.

b. The condition, $p > 1$, is necessary in Theorem A.2.1. In fact, if $p = 1$ and $v = 1$ then by a standard spectral synthesis result, the L^1-closure of $\mathcal{S}_o(\mathbb{R}^d)$ is the (closed) maximal ideal $\{f \in L^1(\mathbb{R}^d): \hat{f}(0) = 0\}$.

A.3. $L^2_{1,1}$ AND PROPER MOMENT SPACES

Consider $\mathcal{S}_o(\mathbb{R})$ as a subspace of $L^2_{1,1}(\mathbb{R})$.

A.3.1. *Proposition.* $\mathcal{S}_o(\mathbb{R})^{\perp}$ (as a subspace of $L^2_{1,1}(\mathbb{R})'$) is the set of constant functions on $\mathbb{R}$.

Using this fact and the inclusion, $L^2_{1,1}(\mathbb{R}) \subseteq L^1(\mathbb{R})$, we have

A.3.2. *Proposition.* The closure of $\mathcal{S}_o(\mathbb{R})$ in $L^2_{1,1}(\mathbb{R})$ is $\{f \in L^2_{1,1}(\mathbb{R}) : \hat{f}(0) = 0\}$.

A.3.3. *Remark.* $L^2_{1,1}$ is a Hilbert space so that its dual is isomorphic to $L^2_{1,1}$. As in the case of L^2_m, this isomorphism is complicated and the continuous linear functions on $L^2_{1,1}$ have an alternate explicit representation. For example, the proof of Proposition A.3.1 shows that the constants are elements of $(L^2_{1,1})'$, and so, noting that the elements of $L^2_{1,1}(\mathbb{R})$ are locally absolutely continuous, we see that there is $g \in L^2_{1,1}(\mathbb{R})'$ so that $[f,g] = f(0)$ for each $f \in L^2_{1,1}(\mathbb{R})$.

B. Notation

Besides the usual notation in analysis as found in the books by L. Hörmander, L. Schwartz, and E. Stein and G. Weiss, we use the following conventions and notation.

The integral over $\mathbb{R}^d$ is designated by "$\int$". The Fourier transform of f is $\hat{f}(\gamma) = \int f(t)e^{-2\pi i t\cdot\gamma}dt$, $\gamma \in \hat{\mathbb{R}}^d(=\mathbb{R}^d)$, and $f = (\hat{f})^{\vee}$.

$\mathcal{S}_o(\mathbb{R}^d) = \{f \in \mathcal{S}(\mathbb{R}^d) : \hat{f}(0) = 0\}$ and $\mathcal{S}_{oa}(\mathbb{R}^d) = \{f \in \mathcal{S}(\mathbb{R}^d) : \hat{f}(\gamma_1,\cdots,\gamma_d) = 0$ if some $\gamma_j = 0\}$.

$C_c(\mathbb{R}^d)$ (resp., $C_c^{\infty}(\mathbb{R}^d)$) is the space of continuous (resp., infinitely differentiable) compactly supported functions on $\mathbb{R}^d$. $C_b(\mathbb{R}^d)$ consists of the bounded continuous functions on $\mathbb{R}^d$.

$M(\hat{\mathbb{R}}^d)$ (resp., $M_b(\hat{\mathbb{R}}^d)$, $M_+(\hat{\mathbb{R}}^d)$, $M_{b+}(\hat{\mathbb{R}}^d)$) is the space of Radon measures (resp., bounded, positive, bounded and positive Radon measures) on $\hat{\mathbb{R}}^d$.

$L^p_v(\mathbb{R}^d) = \{f : \|f\|_{p,v} = (\int|f(t)|^p v(t)dt)^{1/p} < \infty\}$ and L^p_E is L^p_v for $v = \chi_E$, the characteristic function of $E \subseteq \mathbb{R}^d$ with Lebesgue measure $|E|$. $\mathcal{L}(X,Y)$ is the space of continuous linear maps between the topological vector spaces X and Y.

σ_{d-1} designates surface measure on $\hat{\mathbb{R}}^d$. Its restriction to the unit sphere Σ_{d-1} is μ_{d-1}, and $\mu_{d-1}(\Sigma_{d-1}) = \omega_{d-1} = 2\pi^{d/2}/\Gamma(d/2)$. Finally, the ball of radius T centered at the origin $0 \in \mathbb{R}^d$ is $B(0,T) = B(T)$ and if $d = 1$ we write $B(T) = \chi_{(T)}$.

REFERENCES

[Bal] Balian, R. (1981) "Un principe d'incertitude fort en théorie du signal on en mécanique quantique," C. R. Acad. Sci., Paris 292, 1357-1362.

[Ba] Barnes, J. (1970) "Laplace-Fourier transformation, the foundation for quantum information theory and linear physics," Problems in analysis (R. Gunning, ed.), Princeton University Press.

[Bat] Battle, G. (1988) "Heisenberg proof of the Balian-Low theorem," Lett. Math. Phys. 15, 175-177.

[B1] Benedetto, J. (1984) "Some mathematical methods for spectrum estimation," and "Fourier uniqueness criteria and spectrum estimation theorems," Fourier techniques and applications (J. Price, ed.), Plenum Press, N.Y.

[B2] Benedetto, J. (1985) "An inequality associated with the uncertainty principle," Rend. Cir. Mat., Palermo 34, 407-421.

[B3] Benedetto, J., "A multidimensional Wiener-Wintner theorem and spectrum estimation," (submitted).

[BH1] Benedetto, J. and Heinig, H. (1983) "Weighted Hardy spaces and the Laplace transform," Springer Lecture Notes 992, 240-277.

[BH2] Benedetto, J. and Heinig, H., "Fourier transform inequalities with measures weights," Advances in Math. (to appear).

[Bo] Bourgain, J. (1988) "A remark on the uncertainty principle for Hilbertian basis," J. of Functional Anal. 79, 136-143.

[CP] Cowling, M. and Price, J. (1984) "Bandwidth versus time concentration: the Heisenberg-Pauli-Weyl inequality," SIAM J. Math. Anal. 15, 151-165.

[D1] Daubechies, I., "The Wavelet transform, time-frequency localization and signal analysis," IEEE Trans. Inf. Theory (to appear).

[D2] Daubechies, I. (1988) "Orthonormal bases of compactly supported wavelets," Comm. Pure and Appl. Math. 41, 909-996.

[DGM] Daubechies, I., Grossmann, A. and Meyer, Y. (1986) "Painless non-orthogonal expansions," J. Math. Phys. 27, 1271-1283.

[De] DeBruijn, N. (1967) "Uncertainty principles in Fourier analysis," Inequalities (O. Shisha, ed.), Academic Press, N.Y. 57-71.

[DJ] Daubechies, I. and Janssen, A., "Two theorems on lattice expansions," IEEE Trans. Inf. Theory (submitted).

[F] Faris, W. (1978) "Inequalities and uncertainty principles," J. Math. Phys. 19, 461-466.

[FG] Feichtinger, H. and Gröchenig, K., "Banach spaces related to integrable group representations and their atomic decompositions I," J. Functional Anal. (to appear).

[FJ] Frazier, M. and Jawerth, B., "A discrete transform and decomposition of distribution spaces," (preprint).

[G] Gabor, D. (1946) "Theory of communication," J. IEE (London) 93, 429-457.

[Gr] Grenander, U. (1951) "On empirical spectral analysis of stochastic processes," Arkiv för Mat. 1, 503-531.

[GR] Garcia-Cuerva, J. and Rubio de Francia (1985) Weighted norm inequalities and related topics, North-Holland, Amsterdam.

[HW] C. Heil and Walnut, D., "Continuous and discrete wavelet transforms," SIAM Review (to appear).

[HSi] Heinig, H. and Sinnamon, G., "Fourier inequalities and integral representations of functions in weighted Bergman spaces over tube domains," Indiana Math. J. (to appear).

[HS] Heinig, H. and Smith, M. (1986) "Extensions of the Heisenberg-Weyl inequality," International J. Math. and Math. Sci. 9, 185-192.

[He] Hernández, E., "Factorization and extrapolation of pairs of weights," (preprint).

[HJJ] Høholdt, T., Jensen, H. and Justesen, J. (1988) "Double series representation of bounded signals," IEEE Trans. Inf. Theory 34, 613-624.

[K] Kjellberg, B. (1943) "Ein momentumproblem," Ark. Mat. Fys. Astr. 29, A2.

[L] Landau, H. (1987) "Maximum entropy and the moment problem," Bull. AMS 16, 47-77.

[LPS] Landau, H., Pollak, H., and Slepian, D. (1961) "Prolate spheroidal wave functions, Fourier analysis and, uncertainty," I-V, Bell System Tech. J. 40, 43-64, 40(1961), 65-84, 41(1962), 1295-1336, 43(1964), 3009-3058, 57(1978), 1371-1430.

[M1] Meyer, Y. (Feb. 1986) "Principe d'incertitude, bases Hilbertiennes, et algèbres d'operateurs," Sem. Bourbaki 662.

[M2] Meyer, Y. (1986) "Ondelettes fonctions splines, et analyses graduées," U. of Torino.

[P] Price, J. F. (1987) "Sharp local uncertainty inequalities," Studia Math. 85, 37-45.

[PI] Proc. IEEE (1982) "Special issue on spectrum estimation," 70.

[PS] Price, J. F. and Sitaram, A. (1988) "Local uncertainty inequalities for locally compact groups," Trans. AMS 308, 105-114.

[Si] Sinnamon, G., "A weighted gradient inequality," Proc. Roy. Soc. Edinburgh (to appear).

[W] Wiener, N., Collected works (P. Masani, ed.) MIT Press.