Proceedings of Symposia in Applied Mathematics
Volume 00, 19xx

Generalized Harmonic Analysis and Gabor and Wavelet Systems

John J. Benedetto

ABSTRACT. Wiener's Generalized Harmonic Analysis (GHA) is outlined and exposited. As a powerful technology in applicable mathematics, GHA and its ramifications are presented in the context of the uncertainty principle, as well as Gabor and wavelet decompositions.

Contents

1. Introduction

Wiener's Generalized Harmonic Analysis (GHA) extends the classical L^1, L^2, and Fourier series theories of harmonic analysis to the harmonic analysis of functions which are neither periodic nor integrable or square integrable on Euclidean space. In the process, GHA provides a mathematical theory to deal with a host of applications which can not be fathomed satisfactorily by those classical methods. These applications include the harmonic analysis of white light and fluctuating voltages, the modelling of noisy systems with the goal of maximizing signal to noise ratios, and the definition of an effective measure of optical coherence, e.g., **[Bas84]**, **[Lee60]**, **[Lev66]**, **[Mas66]**, **[Ric54]**.

1991 *Mathematics Subject Classification.* Primary 42A38, 42A65, 42C15, 42C30.

I would like to thank the three referees, including Professor Robert Strichartz, for their valuable comments. I would also like to acknowledge Professor Pesi Masani for his singular, magisterial knowledge and expertise of Norbert Wiener's contributions. His impact on this paper has been significant. Finally, I want to express my appreciation to the AFOSR (F496209610193) for their generous support of my work.

Section 2 contains an outline of Wiener's GHA along with some recent developments. There is an emphasis on Wiener's rigorous formulation of the notion of autocorrelation, as well as on the Wiener-Plancherel formula, including an explanation of its importance and its relationship to the classical Parseval-Plancherel formula. Finally, background is provided for Wiener's use of the power spectrum in spectral analysis.

Our goal, after *Section 2*, is to relate Wiener's ideas and methods from GHA with subsequent developments in harmonic analysis. From this point of view, it is natural to focus on Gabor's theory of signal reconstruction from 1946, and, in particular, on what we shall call his *local spectral analysis* [**Gab46**].

Section 3 describes the intellectual interregnum between Wiener and Gabor, and highlights an exciting interleaving and interplay of their ideas concerning signal reconstruction formulas, irregular sampling, communication theory, and the uncertainty principle. Even the sidebars are titillating. For example, Gabor's local spectral analysis was created during the same period he conceived of holography for which he received the Nobel Prize in 1971. It turns out that holography depends on the notion of coherent radiation, which stems from Wiener's GHA (and which Gabor duly acknowledges). Further, nowadays, it is not only meaningful to use the Heisenberg group as a mathematical and unifying backdrop for Gabor's signal analysis [**Ben89**], [**HL95**], but also in studying holography [**Sche90**].

The Classical Uncertainty Principle Inequality is proved in *Section 4*, along with an explanation of its use in Gabor's theory of signal reconstruction and local spectral analysis. *Section 5* is devoted to the theory of frames, and provides what might be considered efficient signal reconstruction formulas consistent with Gabor's ideas about local spectral analysis. In particular, in the spirit of Wiener's point of view, we address the relationship between Gabor's theory and stationarity, as well as the roles of short time Fourier transforms (STFTs), power spectrum computations, and the use of spectrograms. We also define and compare Gabor and wavelet decompositions. The material from *Sections 4* and *5* leads to the discussion in *Section 6* of the Balian-Low theorem, uncertainty principle inequalities associated with Gabor and wavelet systems, and a statement of Bourgain's orthonormal basis theorem in the context of such systems.

In *Section 7*, Gabor's signal analysis from *Sections 4 - 6* is integrated with a raison d'être of GHA, viz., the issue of extricating intelligent messages embedded in noisy environments. *Theorem 7.2* is a Gabor signal reconstruction formula for L^1 which requires Wiener's notion from GHA of mean total power. Finally, in *Section 8* we make a spectrogram analysis, in the spirit of Gabor's local spectral analysis, of fairly sophisticated biomedical data. This data and its analysis is associated with Wiener's interest from the 1930s of extracting information from electroencephalograms by means of autocorrelation methods.

Our notation is described in *Section 9*. To begin the paper, however, we do point out that integration over Euclidean space $\mathbb{R}^d$ is designated by "$\int$" instead of "$\int_{\mathbb{R}^d}$", and that we shall usually deal with the case $d = 1$.

2. Wiener's Generalized Harmonic Analysis (GHA)

In 1930, Norbert Wiener [**30a**, Volume II, pages 183-324] proved an analogue of the Parseval-Plancherel formula, $\|f\|_{L^2(\mathbb{R})} = \|\hat{f}\|_{L^2(\hat{\mathbb{R}})}$, for a large class of functions which is not contained in $L^2(\mathbb{R})$ and whose elements need not be periodic. We refer to his formula as the *Wiener-Plancherel* formula, see (2.4). It is a fundamental result in GHA. Wiener chose to have the formula appear on the cover of his autobiography, *I Am a Mathematician* [**56g**].

Two precursors of Wiener on the subject of GHA, whose work Wiener studied, were Sir Arthur Schuster and Sir Geoffrey I. Taylor. Schuster pointed out analogies between the harmonic analysis of light and the statistical analysis of hidden periods associated with meteorological and astronomical data. Taylor conducted experiments in fluid mechanics dealing with the *onset to turbulence*, and formulated a special case of correlation. A third scientist, whose work (1914) vis à vis GHA was not known to Wiener, was Albert Einstein. Einstein writes: "Suppose the quantity y (for example, the number of sun spots) is determined empirically as a function of time, for a very large interval, T. How can one represent the statistical behavior of y?" In his heuristic answer to this question he came close to the notions of autocorrelation and power spectrum [**Ein14**], cf., commentaries by Masani [**Mas86**] and Yaglom [**Yag85**].

The Fourier analysis of $L^1(\mathbb{R})$ or $L^2(\mathbb{R})$ or the theory of Fourier series were inadequate tools to analyze the issues confronting Schuster, Taylor, and Einstein. On the other hand, GHA became a successful device to gain some insight into their problems, as well as other problems where the data and/or noises can not be modelled by the Fourier transform decay, finite energy, or periodicity inherent in the above classical theories, e.g., [**Ars66**, Chapter II], [**Bas84**], [**Ben75**, Chapter 2], [**Ber87**], [**Ric54**], [**Wie76**, Volume II], [**33i**], [**49g**].

DEFINITION 2.1 (Bounded Quadratic Means). The space $BQM(\mathbb{R})$ of functions having *bounded quadratic means* is the set of all functions $f \in L^2_{\text{loc}}(\mathbb{R})$ for which

$$\sup_{T>0} \frac{1}{2T} \int_{-T}^{T} |f(t)|^2 \, dt < \infty.$$

The *Wiener space* $W(\mathbb{R})$ is the set of all functions $f \in L^2_{\text{loc}}(\mathbb{R})$ for which

$$\int \frac{|f(t)|^2}{1+t^2} \, dt < \infty.$$

Wiener noted the second inclusion of the following result [**33i**, Theorem 20], and each of the inclusions is elementary to verify.

THEOREM 2.2 (Inclusions for GHA).

$$L^\infty(\mathbb{R}) \subseteq BQM(\mathbb{R}) \subseteq W(\mathbb{R}) \subseteq \mathcal{S}'(\mathbb{R}),$$

and the inclusions are proper.

Inspired by Schuster's periodogram analysis, and with the goal of eliminating some of its weaknesses [**30a**, page 131], Wiener introduced what we shall call the Wiener transform in [**30a**], cf., [**Bri81**], [**Pri81**] for periodogram analysis.

DEFINITION 2.3 (The Wiener Transform).

a. The *Wiener transform* (integrated Fourier transform) of $f \in W\mathbb{R})$ is defined as the sum $Wf = s = s_1 + s_2$ where

$$s_1(\gamma) = \int_{-1}^{1} f(t) \frac{e^{-2\pi i t\gamma} - 1}{-2\pi i t} \, dt$$

and

$$s_2(\gamma) = \int_{|t|\geq 1} f(t)\frac{e^{-2\pi it\gamma}}{-2\pi it}\,dt.$$

b. Since $f \in L^1[-1,1]$, we see that s_1 is continuous on $\mathbb{R}$ and $|s_1(\gamma)| \leq 2|\gamma|\,\|f\|_{L^1[-1,1]}$. Since $f \in W(\mathbb{R})$, *Theorem 2.2* and the Parseval-Plancherel Theorem allow us to conclude that $s_2 \in L^2(\widehat{\mathbb{R}})$. In particular, $s \in L^2_{\text{loc}}(\widehat{\mathbb{R}}) \cap \mathcal{S}'(\widehat{\mathbb{R}})$.

c. There is a natural analogue of the Wiener transform on $\mathbb{R}^d$, and in this case it is also true that s is an element of $L^2_{\text{loc}}(\widehat{\mathbb{R}}^d) \cap \mathcal{S}'(\widehat{\mathbb{R}}^d)$, see [**BBE89**, Theorem 3.5]. We mention this since in higher dimensions there are technicalities in proof which can be simplified by using Wiener-amalgam spaces and the Bochner-Fourier transform on such spaces, e.g., [**FS85**], [**Ohk79**].

By definition, the Wiener-amalgam space $W^{pq}(\mathbb{R})$ consists of measurable functions f on $\mathbb{R}$ for which

$$\left(\sum_{n=-\infty}^{\infty} |\int_n^{n+1} |f(t)|^p\,dt|^{q/p}\right)^{1/q} < \infty,$$

with the standard adjustment in case p or q is ∞. Wiener introduced this concept by dealing with the space $S_W(\mathbb{R})$ of continuous elements of $W^{\infty,1}(\mathbb{R})$, see [**32a**], [**33i**].

The original conception and proof of *Theorem 2.4* is in [**30a**], with further developments in [**33i**]. Our distributional formulation, which is elementary, was first proved in [**Ben75**, pages 88–89].

THEOREM 2.4 (The Derivative of the Wiener transform). *Let $f \in W(\mathbb{R})$. Then $f \in \mathcal{S}'(\mathbb{R})$ and*

$$s' = \widehat{f},$$

where $s \in L^2_{\text{loc}}(\widehat{\mathbb{R}}) \cap \mathcal{S}'\widehat{\mathbb{R}})$ is the Wiener transform of f, and s' is the (Schwartz) distributional derivative of s.

DEFINITION 2.5 (Autocorrelation and Power Spectrum).

a. The *autocorrelation* R of a function $f : \mathbb{R} \to \mathbb{C}$ is formally defined as

$$R(t) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f(u+t)\overline{f(u)}\,du.$$

To fix ideas, suppose R exists for each $t \in \mathbb{R}$. It is easy to prove that R is positive definite, and so by the Herglotz-Bochner Theorem R is the Fourier transform of some bounded positive measure $S \in M_{b+}(\widehat{\mathbb{R}})$, i.e., $\widehat{S} = R$. S is called the *power spectrum* of f.

Notwithstanding its bearing on the notion of optical coherence mentioned in *Section 1*, we shall not deal with the *cross-correlation* of functions f and g, but refer to [**Kla96**] in this volume for a thorough treatment.

b. Since "power" equals "energy/time", Wiener spoke of

$$R(0) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} |f(t)|^2\,dt$$

as the *mean total power* of the signal f. This explains why S is called the power spectrum of f, see [**49g**, pages 39–40 and 42].

c. The usefulness of the autocorrelation R is that $R(t)$ can be measured in many cases where the underlying signal f is not given by a formula, but where the graph of f is available as in financial or biomedical time series, cf., *Remark 5.9* and *Section 8.*

d. The discrete measure part of the power spectrum S characterizes periodicities in f, e.g., [**61c**, Chapter X]. This fact is illustrated by taking

$$f(t) = \sum_{k=1}^{n} r_k e^{-2\pi i t \lambda_k}, \; r_k \in \mathbb{C}, \; \lambda_k \in \widehat{\mathbb{R}}.$$

In this case the L^2-autocorrelation $\int f(u+t)\overline{f(u)}du$ is not defined, but the autocorrelation is easily computed to be $\sum_{k=1}^{n} |r_k|^2 e^{-2\pi i t \lambda_k}$. Hence, the power spectrum of f is

$$S = \sum_{k=1}^{n} |r_k|^2 \delta_{\lambda_k}.$$

e. At the other end of the spectrum, so to speak, if $f : \mathbb{R} \to \mathbb{C}$ has the property that $\lim_{|t|\to\pm\infty} f(t) = 0$, then $S = 0$. It is elementary to construct examples f for which $S = 0$ whereas $\limsup_{|t|\to\pm\infty} |f(t)| > 0$, e.g., [**33i**, pages 151–154], [**Bas84**, pages99–100], [**Ben75**, pages 84 and 87].

Given the definition of autocorrelation and power spectrum, it is natural to ask if every $S \in M_{b+}(\widehat{\mathbb{R}})$ is the power spectrum of some function $f : \mathbb{R} \to \mathbb{C}$. Wiener and Wintner [**39c**] gave the following positive response in 1939. There have been subsequent significant contributions by Bass and Bertrandias, e.g., [**Bas84**], [**Ber87**]. R. Kerby and I proved the d-dimensional version; one basic construction is given in [**Ben91a**], and two others, which are quite ingenious, are contained in [**Ker90**].

THEOREM 2.6 (Wiener-Wintner Theorem). *Let $S \in M_{b+}(\widehat{\mathbb{R}}^d)$. There is a constructible function $f \in L^\infty_{\text{loc}}(\mathbb{R}^d)$ such that*

$$R(t) = \lim_{T\to\infty} \frac{1}{|B(T)|} \int_{B(T)} f(t+u)\overline{f(u)}\, du$$

exists for all $t \in \mathbb{R}^d$, and $\widehat{S} = R$, where $B(T) = \{t \in \mathbb{R}^d : |t| \leq T\}$. Briefly, if $S \in M_{b+}(\widehat{\mathbb{R}}^d)$, then there is a self-correlated $f \in L^\infty_{\text{loc}}(\mathbb{R}^d)$ whose power spectrum is S.

In order to prove the Wiener-Plancherel formula, Wiener required a Tauberian theorem of the following type, see [**30a**, pages 141–152].

THEOREM 2.7 (Wiener Tauberian Theorem). *Let $g \in L^1(\mathbb{R})$ have a nonvanishing Fourier transform and let $\varphi \in L^\infty(\mathbb{R})$. If*

$$\lim_{t\to\infty} g * \varphi(t) = r \int g(u)\, du, \tag{2.1}$$

then

$$\forall f \in L^1(\mathbb{R}), \quad \lim_{t\to\infty} f * \varphi(t) = r \int f(u)\, du. \tag{2.2}$$

REMARK 2.8 (Wiener Tauberian Theorem).

a. *Theorem 2.7* has the format of a *classical* Tauberian theorem, viz., a boundedness (or some other) condition and "summability" by a certain method yield "summability" by other methods. In *Theorem 2.7*, the boundedness or "Tauberian" condition is the hypothesis that $\varphi \in L^\infty(\mathbb{R})$. The given summability is (2.1), where g represents a so-called "summability method". The conclusion (2.2) of the theorem is summability for a whole class of summability methods, viz., for all $f \in L^1(\mathbb{R})$. A classical and masterful treatment of summability methods is due to Hardy **[Har49]**.

The particular functions used by Wiener to prove his Wiener Tauberian formulas are found in **[33i]**, **[Ben75**, pages 91-92].

b. *Modern* Tauberian theorems have a more algebraic and/or functional analytic flavor to them. For example, the Wiener Tauberian Theorem is a special case of the fact that *if* $\widehat{g} \in A(\mathbb{R}), T \in A'(\mathbb{R})$, *and* $T\widehat{g} = 0$, *then* $\widehat{g} = 0$ *on* $\operatorname{supp} T$. In fact, the generalizations of *Theorem 2.7* are much more far reaching than this. **[Ben75]** gives an extensive treatment of both classical and modern Tauberian theory, as well as the history of the subject, and applications to spectral synthesis and analytic number theory.

Signal f	$\longleftrightarrow$	$\widehat{f} = s'$	s
Autocorrelation $R = \widehat{S}$	$\longleftrightarrow$	S	$\{\frac{2}{\epsilon}\lvert\Delta_\epsilon s\rvert^2\}$

FIGURE 1

The *Wiener-Plancherel formula* is defined to be equation (2.4) of the following result, cf., *Figure 1*.

THEOREM 2.9 (Wiener-Plancherel Formula). *Let* $f \in BQM(\mathbb{R})$, *and suppose its autocorrelation* $R = \widehat{S}$ *exists for each* $t \in \mathbb{R}$.

a.

$$\forall t \in \mathbb{R}, \quad R(t) = \lim_{\epsilon \to 0} \frac{2}{\epsilon} \int |\Delta_\epsilon s(\gamma)|^2 e^{-2\pi i t\gamma}\, d\gamma, \tag{2.3}$$

where $\Delta_\epsilon s(\gamma) = \frac{1}{2}(s(\gamma+\epsilon) - s(\gamma-\epsilon))$.

b. In particular,

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |f(t)|^2\, dt = \lim_{\epsilon\to 0} \frac{2}{\epsilon} \int |\Delta_\epsilon s(\gamma)|^2\, d\gamma. \tag{2.4}$$

The original proof is in **[30a]** *and more conveniently in* **[33i]**, *cf.*, **[Ben75]**.

EXAMPLE 2.10 (Related Formulas).

a. Because of (2.3) and assuming the setup of *Theorem 2.9*, the following formulas are true under the proper hypotheses, e.g., **[Ben75**, page 90], **[Ben91a**, page 847]:

$$\lim_{\epsilon\to 0} \frac{2}{\epsilon} |\Delta_\epsilon s|^2 = S, \tag{2.5}$$

and

$$\int |\widehat{k}(\gamma)|^2 \, dS(\gamma) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |k * f(t)|^2 \, dt$$

$$(2.6) \qquad = \lim_{\epsilon\to 0} \frac{2}{\epsilon} \int |\widehat{k}(\gamma)\Delta_\epsilon s(\gamma)|^2 \, d\gamma.$$

b. Formally, the second equality of (2.6) is (2.4) for the case $k = \delta$. For $k \in C_c(\mathbb{R})$ the first equality of (2.6) is not difficult to prove, e.g., [**Ben91a**, pages 847-848].

REMARK 2.11 (Importance of the Wiener-Plancherel Formula). The Parseval-Plancherel formula, $||f||_{L^2(\mathbb{R})} = ||\widehat{f}||_{L^2(\widehat{R})}$, allows us to define the Fourier transform of a square integrable function, e.g., [**Ben96**, Theorem 1.10.2]; and, at a certain level of abstraction, this formula can be considered to characterize what is meant by an harmonic analysis of f. On the other hand, for most applications in Euclidean space, the Parseval-Plancherel formula assumes the workaday role of an effective tool used to obtain quantitative results. It is this latter role that was envisaged for the Wiener-Plancherel formula in dealing with functions not amenable to classical harmonic analysis methods such as Fourier series or L^1 and L^2 Fourier transforms.

Schwartz' theory of distributions gives the proper definition of the Fourier transform of tempered distributions. Thus, since $BQM(\mathbb{R}) \subseteq S'(\mathbb{R})$ by *Theorem 2.2*, the Wiener-Plancherel formula can only be viewed as a special case of Schwartz' theory when it comes to generalizing the L^1 and L^2 definitions of the Fourier transform. What we often need, however, are methods that are suitable for calculation in cases where the Parseval-Plancherel formula does not apply. The Wiener-Plancherel formula serves this purpose in a variety of applications. As examples of such applications, there is a host of problems which falls under the heading of an harmonic (spectral) analysis of signals containing non-square-integrable noise and/or random components, whether it be for speech recognition, image processing, geophysical modeling, or turbulence in fluid mechanics. These problems can sometimes be attached by Beurling's profound theory of spectral synthesis [**Beu89**], [**Ben75**] or by the extensive multifaceted theory of time series, e.g., [**Pri81**]. However, Beurling's spectral synthesis does not deal with energy and power considerations, i.e., quadratic criteria, and time series relies on a stochastic point of view. On the other hand, the Wiener-Plancherel formula *Theorem 2.4* is intrinsically quadratic and deals with such problems deterministically and quantitatively.

EXAMPLE 2.12 (Multidimensional Wiener-Plancherel Formulas).

a. The *spherical Wiener Plancherel formula* on $\mathbb{R}^d$ is

$$(2.7) \qquad \lim_{T\to\infty} \frac{1}{|B(T)|} \int_{B(T)} |f(t)|^2 \, dt = \lim_{\lambda\to\infty} \frac{c(d,k)(2\pi)^{4k}}{\omega_{d-1}\lambda^{4k-d}} \int |D_\lambda s_k(\gamma)|^2 \, d\gamma,$$

cf., [**Ben91b**] for a precise statement of hypotheses for the validity of (2.7). The function s_k is the *Wiener transform*,

$$s_k = \widehat{f} * E_k,$$

where $\Delta^k E_k = \delta$, ω_{d-1} is the surface area of the unit sphere Σ_{d-1} in $\mathbb{R}^d$, $c(d,k)^{-1}$ is the L^1-norm of a special function related to the Fourier transform of the restriction of surface measure σ_{d-1} to Σ_{d-1},

$$D_\lambda s_k = s_k - M_\lambda s_k,$$

and M_λ is the spherical mean-value operator defined by

$$M_\lambda s_k(\gamma) = \frac{1}{\omega_{d-1}} \int_{\Sigma_{d-1}} s_k(\gamma + \lambda\theta)\, d\sigma_{d-1}(\theta).$$

The integer k is related to the dimension d, and there must be control of the quadratic means of f over spheres in order to verify (2.7).

b. In [**BBE89**], we proved a rectilinear version of (2.7), cf., [**AKM80**]. The rectilinear result is easier to prove than the spherical one, although by no means elementary. Also, in the case of "rectilinear geometry" the hyperbolic operator $\partial_1\partial_2 \ldots \partial_d$ plays a major role, whereas the "spherical geometry" of (2.7) gives rise to the elliptic operator Δ^k. This remark indicates there is a range of Wiener-Plancherel formulas according to the number of degrees of freedom available in various convergence criteria.

c. As in the proof of the Wiener-Plancherel formula on $\mathbb{R}$, Wiener's Tauberian Theorem is required to prove the rectilinear and spherical versions of the Wiener-Plancherel formula on $\mathbb{R}^d$. For example, in the rectilinear case [**BBE89**], we used the dual of a Tauberian theorem associated with the space $S_W(\mathbb{R}^d)$, which was defined on $\mathbb{R}$ in *Definition 2.3c*. $S_W(\mathbb{R}^d)$ is a particular *Segal algebra*, e.g., [**Seg47**]; and although Wiener provided the prototype $S_W(\mathbb{R})$, it was Segal who noticed the importance and generality of its underlying structure.

A *Segal algebra* $S(\mathbb{R}^d)$ is a dense subalgebra of $L^1(\mathbb{R}^d)$ satisfying the following conditions:

i. $S(\mathbb{R}^d)$ is a Banach algebra with norm $||\cdots||_S$ and the natural injection $S(\mathbb{R}^d) \subseteq L^1(\mathbb{R}^d)$ is continuous;

ii. $S(\mathbb{R}^d)$ is translation-invariant and there exist constants $A, B > 0$ such that

$$\forall f \in S(\mathbb{R}^d) \quad \text{and} \quad \forall t \in \mathbb{R}^d,\ A||f||_S \le ||\tau_t f||_S \le B||f||_S;$$

iii. $\forall f \in S(\mathbb{R}^d)$ and $\forall \epsilon > 0$, $\exists \delta > 0$ such that

$$\forall t \in B(\delta), \quad ||\tau_t f - f||_S < \epsilon.$$

Segal algebras have bounded approximate identities, and every closed ideal $I_S \subseteq S(\mathbb{R}^d)$ has the form $I \cap S(\mathbb{R}^d)$ where I is a unique closed ideal of $L^1(\mathbb{R}^d)$, e.g., [**Rei68**, Chapter 6.2].

EXAMPLE 2.13 (Topological Properties of the Wiener Transform). The space $BQM(\mathbb{R})$ is also called the *Besicovitch space* $B_{2,\infty}(\mathbb{R})$. Further, in light of the Wiener-Plancherel formula, it is natural to consider the space of functions g for which

$$||g|| = \sup_{\epsilon>0} \left(\frac{2}{\epsilon} \int |\Delta_\epsilon g(\gamma)|^2 d\gamma \right)^{1/2} < \infty.$$

This space is the *Besov space* $B^{1/2}_{2,\infty}(\mathbb{R})$. Chen and Lau [**CL90**] proved that the Wiener transform is a topological isomorphism $W : B_{2,\infty}(\mathbb{R}) \longrightarrow B^{1/2}_{2,\infty}(\mathbb{R})$, cf., [**Hei96**] for some extensions to other Besicovitch and Besov spaces.

3. In the aftermath of Wiener's GHA

Our purpose in this section is to observe the emergence and interplay of ideas from harmonic analysis in the spectral analyses of Wiener [**30a**] and Gabor [**Gab46**].

Wiener's GHA and, in particular, the Wiener-Plancherel formulas of *Theorem 2.9* and *Example 2.10* provide a relationship between a function and its power spectrum. Among its uses, GHA is a meaningful mathematical model for the *spectral analysis* of signals $f = s + n$, where $n \in BQM(\mathbb{R})$ is some kind of noise and where s is an "intelligent" periodic message.

In a communication system, the receiver records the data $f(u)$; and a classical problem is to extricate $s(u+t)$ from $f(u)$, where n and sometimes s are assumed to be stationary stochastic processes. If $t > 0$, then the problem is a particular *prediction problem.* Wiener solved the basic problem in 1942, and there was an independent profound contribution by Kolmogorov [**Kol41b**], cf., [**DM76**]. Wiener's solution [**49g**] involves optimal filtering in the context of minimizing quadratic means of the type initiated in GHA [**49g**, pages 37–43], [**49g**, Appendix C by N. Levinson]. Wiener proved that the minimization was equivalent to the solution of a Wiener-Hopf integral equation of the first kind, which he then proceeded to solve [**49g**, pages 60–64].

In his conception of a communication system, modeled after real problems with which he was presented, Wiener began only with assumptions about prior information that could be derived from measurements on the received signal f. Gabor has given an incisive comparison of Wiener's and Shannon's approaches to the theory of communication from the point of view of the optimal use of prior information [**Gab65**], cf., [**Ash65**] for another comparison of Wiener's and Shannon's points of view, as well as [**Pro89**] for an update on one aspect of the subject since the 1940s to see how extensive its development has been.

Shannon's own evaluation of Wiener's contribution to communication theory is encapsulated in the following [**Sha49**, pages 52–53]: "Communication theory is heavily indebted to Wiener for much of its basic philosophy and theory. His classic NDRC report [**49g**] contains the first clear-cut formulation of communication theory as a statistical problem, the study of operations on time series. This work, although chiefly concerned with the linear prediction and filtering problem, is an important collateral reference in connection with the present paper".

In a different stream of thought in this same area, Gabor introduced the notion of a *local spectral analysis* in his fundamental paper, "Theory of communication" [**Gab46**], cf., *Remark 4.5* and *Remark 5.9.* His work was inspired by developments in quantum mechanics as well as signal analysis, and it provides a phase-space analysis of "information" and the transmission of information by, for example, speech or television, cf., [**Pie65**, page 43] concerning Gabor's omission of noise analysis. Gabor's key idea revolved around the Classical Uncertainty Principle Inequality (*Theorem 4.1*) and a method of signal decomposition depending on this inequality.

For now we want to emphasize the ideas from harmonic analysis in Wiener's and Gabor's work, more than their contributions to any particular application such as communication theory. For example, Gabor's signal decomposition mentioned above and formulated in (4.10) can be considered a special type of frame decomposition; and the theory of frames is discussed in *Section 5* from the point of view of harmonic analysis.

It turns out that the Classical Sampling Theorem (Shannon Sampling Theorem [**Sha49**, pages 50 and 53–54]) is a particular case of a wavelet (multiresolution) decomposition [**Dau92**] as well as of a Gabor frame decomposition in the case of regularly spaced sampling points [**Ben92**, Theorems 25 and 26, proved jointly with

W. Heller], cf., the sampling formulas documented in [**Boa54**] and first formulated in the 19th century.

In the case of irregularly spaced sampling points, the theory of frames is also used to prove sampling theorems [**Ben92**, Section 8, containing results proved jointly with W. Heller]. The mathematical background for these theorems originates from results of Wiener [**27e**] and of Paley-Wiener [**34d**, Chapter VI] on the completeness of exponential sequences $\{e^{2\pi i t_n \gamma}\}$ in $L^2[-\Omega,\Omega]$. Their work was followed by the penetrating analysis of Levinson [**Lev40**], and by the fundamental paper of Duffin and Schaeffer [**DS52**], which we shall discuss in *Section 5*, cf., the comparably important contributions by Landau [**Lan67**]. A milestone on the completeness problem was attained in the profound theorems of Beurling and Malliavin [**BM62**], [**BM67**].

Not only are the frame theoretic outgrowths of Gabor's decompositions ultimately related to some of Wiener's ideas (about completeness) but, in implementing the Classical Uncertainty Principle Inequality (*Theorem 4.1*), Gabor broached on another subject close to Wiener's heart. In [**56g**, pages 105–107], Wiener discusses his analysis from 1925 of the complicated interaction of time and frequency in music, cf., related musical interpretations of the uncertainty principle in [**Ben96**, Remark 1.1.4] and [**dBr67**]. It was apparently during this period, at a lecture he gave in Göttingen, that Wiener first presented *Theorem 4.1* of the next section [**Bar68**, page 168], [**Bar70**].

4. The classical uncertainty principle and Gabor's idea

An uncertainty principle in harmonic analysis has come to mean a formula, or inequality, or quantitative expression relating the behavior of a function, such as decay or support, with that of its Fourier transform [**Ben90**], [**BF94**, Chapter 7.6], [**HJ94**], [**Pric85**, pages 25–44 and 149–170].

For example, if $f \in L^2(\mathbb{R})\backslash\{0\}$, then f and $\widehat{f}$ cannot both have compact support. The proof of this assertion is elementary. In fact, if f has compact support then $\widehat{f}(z) = \int f(t)e^{-2\pi i t z}dt,\ z \in \mathbb{C}$, is an entire function; and so if $\widehat{f}(\gamma)$ is zero on a subinterval of $\widehat{\mathbb{R}}$ then $\widehat{f}$ is identically zero. Slepian [**Sle76**] has written a celebrated paper starting with this particular uncertainty principle, and addressed the question of whether or not real signals are really bandlimited. In related earlier work, Landau, Pollak, and Slepian of Bell Laboratories [**PS61**], [**LP61**], [**LP62**], [**Pric85**, pages 201–220] dealt with the problem of quantifying how close both a function f and its Fourier transform $\widehat{f}$ can be to having compact supports. For fixed $T, \Omega > 0$, they defined

$$\forall f \in L^2(\mathbb{R}),\quad \alpha_f = \frac{(\int_{-T}^{T}|f(t)|^2dt)^{1/2}}{\|f\|_{L^2(\mathbb{R})}} \quad\text{and}\quad \beta_f = \frac{(\int_{-\Omega}^{\Omega}|\widehat{f}(\gamma)|^2d\gamma)^{1/2}}{\|f\|_{L^2(\widehat{\mathbb{R}})}}.$$

Clearly, $(\alpha_f, \beta_f) \in [0,1]\times[0,1]$, and we just proved that there is no $f \in L^2(\mathbb{R})\backslash\{0\}$ for which $(\alpha_f,\beta_f) = (1,1)$. One of the "Bell Labs Theorems" characterizes a region $R \subseteq [0,1]\times[0,1]$ with the property that if $(\alpha,\beta) \in R$ then there is $f \in L^2(\mathbb{R})$ for which $(\alpha_f,\beta_f) = (\alpha,\beta)$.

The following result is the harmonic analysis uncertainty principle mentioned at the end of *Section 3* in conjunction with Wiener's interests and Gabor's "Theory of communication". It asserts that a function and its Fourier transform cannot both be "concentrated" at any given pair of points $t_0 \in \mathbb{R}$ and $\gamma_0 \in \widehat{\mathbb{R}}$. We refer to

Theorem 4.1 as the Classical Uncertainty Principle Inequality, and, in the context of quantum mechanics, it is also called the Heisenberg uncertainty principle, e.g., [**Wey50**, page 77].

THEOREM 4.1 (The Classical Uncertainty Principle Inequality). *Let* $(t_0, \gamma_0) \in \mathbb{R} \times \widehat{\mathbb{R}}$. *Then*

$$\forall f \in L^2(\mathbb{R}), \quad ||f||^2_{L^2(\mathbb{R})} \leq 4\pi||(t-t_0)f(t)||_{L^2(\mathbb{R})}||(\gamma-\gamma_0)\widehat{f}(\gamma)||_{L^2(\widehat{\mathbb{R}})}, \tag{4.1}$$

and there is equality in (4.1) if and only if f is of the form

$$\varphi(t) = Ce^{2\pi it\gamma_0}e^{-s(t-t_0)^2}, \tag{4.2}$$

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

PROOF. i. The mapping $f(t) \longmapsto f(t+t_0)e^{-2\pi it\gamma_0}$ shows that it is sufficient to verify (4.1) and (4.2) for $(t_0, \gamma_0) = (0,0)$.

ii. Let $f \in \mathcal{S}(\mathbb{R})$. For each $t \in \mathbb{R}$ we have

$$(|f(t)|^2)' \leq 2|\overline{f(t)}f'(t)|. \tag{4.3}$$

The desired inequality (4.1) for $(t_0, \gamma_0) = (0,0)$ is a consequence of the following calculation:

$$\begin{aligned} ||f||^4_{L^2(\mathbb{R})} &= \left(\int t\,(|f(t)|^2)'\,dt\right)^2 \\ &\leq 4\left(\int |t\overline{f(t)}f'(t)|\,dt\right)^2 \\ &\leq 4\int |tf(t)|^2\,dt \int |f'(t)|^2\,dt \\ &= 16\pi^2 \int |tf(t)|^2\,dt \int |\gamma\widehat{f}(\gamma)|^2 d\gamma. \end{aligned} \tag{4.4}$$

The first equality of (4.4) follows from the integration by parts formula, the first inequality results from (4.3), the second inequality is Hölder's Inequality, and the second equality of (4.4) is a consequence of the Parseval-Plancherel Theorem and the fact that $(f')^\wedge(\gamma) = 2\pi i\gamma\widehat{f}(\gamma)$.

iii. Clearly, the inequality (4.1), for the case $(t_0, \gamma_0) = (0,0)$, is valid for all $f \in L^2(\mathbb{R})$ for which $V_f \equiv \max(||tf(t)||_{L^2(\mathbb{R})}, ||\gamma\widehat{f}(\gamma)||_{L^2(\widehat{\mathbb{R}})}) = \infty$. If $V_f < \infty$, then it is possible to choose a sequence $\{f_n : n = 1, \cdots\} \subseteq \mathcal{S}(\mathbb{R})$ such that

$$\lim_{n\to\infty} \max\left(||f-f_n||_{L^2(\mathbb{R})}, ||t(f(t)-f_n(t))||_{L^2(\mathbb{R})}, ||\gamma(\widehat{f}(\gamma)-\widehat{f}_n(\gamma))||_{L^2(\widehat{\mathbb{R}})}\right) = 0, \tag{4.5}$$

e.g., [**Ben90**, Appendix A], thereby attaining (4.1) from parts *i* and *ii*.

iv. Let $\varphi(t) = Ce^{-st^2}$, $s > 0$. It is elementary to compute

$$||\varphi||^2_{L^2(\mathbb{R})} = |C|^2\left(\frac{\pi}{2s}\right)^{1/2}, \quad ||t\varphi(t)||_{L^2(\mathbb{R})} = |C|\left[\left(\frac{\pi}{2s}\right)^{1/2}\frac{1}{4s}\right]^{1/2},$$

and

$$||\gamma\widehat{\varphi}(\gamma)||_{L^2(\widehat{\mathbb{R}})} = |C|\left[\left(\frac{2s}{\pi}\right)^{1/2}\frac{1}{8\pi}\right]^{1/2},$$

e.g., [**Ben96**, Chapter 1]; and thus equality is obtained in (4.1) by Gaussians of the form in (4.2).

In order to prove that Gaussians are the only equalizers of (4.1) in the set $\{f \in L^2(\mathbb{R}) : V_f < \infty\}$, we first note that there are differentiability properties of f inherent in the condition $V_f < \infty$, e.g., *Remark 4.2a*. These properties allow us to modify the proof of (4.4) to deal with the general case in which $V_f < \infty$, cf., the approach in part *iii*.

Taking f for which $V_f < \infty$, and assuming equality in (4.1), we can then use the criterion for equality in Hölder's Inequality as applied to the modified version of (4.4). This criterion yields an elementary differential equation whose only solutions are of the form (4.2). □

REMARK 4.2.

a. In outlining the proof that Gaussians are the only equalizers of (4.1), we alluded to differentiability properties of $f \in L^2(\mathbb{R})$ in the case $V_f < \infty$. What we mean is the following fact [**BHW95**, Theorem 5.2]: if $f \in L^2(\mathbb{R})$ then its distributional derivative ∂f is an element of $L^2(\mathbb{R})$ if and only if $\gamma\hat{f}(\gamma) \in L^2(\mathbb{R})$; in this case, f is locally absolutely continuous, the ordinary pointwise a.e. derivative equals ∂f a.e., and

$$\partial f(t) = (2\pi i\gamma\hat{f}(\gamma))^{\vee}(t).$$

This elementary result is also used in our analysis with Heil and Walnut of the Balian-Low Theorem [**BHW95**, Sections 5 and 6], cf., *Section 6.*

b. The ingredients of (4.4) are integration by parts, Hölder's Inequality, and the Parseval-Plancherel Theorem. Integration by parts can be generalized to various Hardy inequalities, and the Parseval-Plancherel Theorem can be generalized to various weighted Fourier transform norm inequalities. The results of these generalizations are weighted extensions of (4.1), and they are due to Heinig and the author, e.g., [**Ben90**], [**BF94**, Chapter 7.8]. Related weighted uncertainty principle inequalities, as well as local uncertainty principle inequalities, are due to Cowling, Faris, and J. F. Price, e.g., [**Pric85**, pages 31–33].

In light of the relevance of *Theorem 4.1* in quantum mechanics, as well as the operator theoretic formulation of many ideas from quantum mechanics [**vN55**], we state the following version of (4.1).

THEOREM 4.3 (Operator Theoretic Uncertainty Principle Inequality). *Let A, B be self-adjoint operators on a complex Hilbert space H, where A and B need not be continuous. Denote the domain of A by $D(A)$, and define the commutator $[A, B] = AB - BA$ as well as the quantities*

$$\forall x \in D(A), \quad E_x(A) = \langle Ax, x\rangle$$

and

$$\forall x \in D(A^2), \ \sigma_x^2(A) = E_x(A^2) - \{E_x(A)\}^2.$$

If $x \in D(A^2) \cap D(B^2) \cap D(i[A, B])$ and $\|x\| \leq 1$, then

$$\{E_x(i[A, B])\}^2 \leq 4\sigma_x^2(A)\sigma_x^2(B). \tag{4.6}$$

See [**BF94**, Theorem 7.32] *for a proof and* [**vN55**, pages 230–247] *for von Neumann's original presentation.*

Inequality (4.1) is a corollary of *Theorem 4.3* for the case $H = L^2(\mathbb{R})$, where the operators A and B are defined as

$$A(f)(t) = (t - t_0)f(t)$$

and

$$B(f)(t) = i(2\pi i(\gamma - \gamma_0)\widehat{f}(\gamma))^{\vee}(t).$$

DEFINITION 4.4 (Variance and the Uncertainty Principle).

a. If $f \in L^2(\mathbb{R})$ and $\|f\|_{L^2(\mathbb{R})} = 1$, then the *expected values* associated with $|f|^2$ and $|\widehat{f}|^2$ are

$$\tilde{t} = \int t|f(t)|^2 dt \text{ and } \tilde{\gamma} = \int \gamma|\widehat{f}(\gamma)|^2 d\gamma,$$

respectively; and the *variances* associated with $|f|^2$ and $|\widehat{f}|^2$ are

$$\sigma^2 t = \int (t - \tilde{t})^2 |f(t)|^2 dt \text{ and } \sigma^2\gamma = \int (\gamma - \tilde{\gamma})^2 |\widehat{f}(\gamma)|^2 d\gamma,$$

respectively.

b. If $|f|^2$ is the probability density function of a random variable X, then $\sigma^2 t$ is precisely the usual probabilistic notion of the variance σ_X^2 of X. It is in this context that the quantities $E_x(A)$ and $\sigma_x^2(A)$ of *Theorem 4.3* are defined, see **[BF94**, Example 7.33], **[Ben96**, Section 2.8].

c. For a given $s > 0$, choose any $C \in \mathbb{C}$ for which $|C|^2 = (2s/\pi)^{1/2}$. Then, for any $(t_0, \gamma_0) \in \mathbb{R} \times \widehat{\mathbb{R}}$, φ defined in (4.2) has the properties that $\|\varphi\|_{L^2(\mathbb{R})} = 1$,

$$4\pi\|(t - t_0)\varphi(t)\|^2_{L^2(\mathbb{R})} = \pi/s \text{ and } 4\pi\|(\gamma - \gamma_0)\widehat{\varphi}(\gamma)\|^2_{L^2(\widehat{\mathbb{R}})} = s/\pi.$$

In light of part a, we introduce the notation

$$\sigma^2 t_s = 4\pi\|(t - t_0)\varphi(t)\|^2_{L^2(\mathbb{R})} \tag{4.7}$$

and

$$\sigma^2\gamma_s = 4\pi\|(\gamma - \gamma_0)\widehat{\varphi}(\gamma)\|^2_{L^2(\widehat{\mathbb{R}})}, \tag{4.8}$$

so that $\sigma^2 t_s$ and $\sigma^2\gamma_s$ are independent of t_0 and γ_0, and, in fact,

$$\sigma^2 t_s = \frac{\pi}{s} \text{ and } \sigma^2\gamma_s = \frac{s}{\pi}.$$

Hence, the case of equality in *Theorem 4.1* can be written as

$$1 = \sigma^2 t_s \sigma^2 \gamma_s, \tag{4.9}$$

independent of the pair $(t_0, \gamma_0) \in \mathbb{R} \times \widehat{\mathbb{R}}$ chosen in the definition of φ. It should also be pointed out that the numbers t_0 and γ_0 are not necessarily expected values in the definitions of $\sigma^2 t_s$ and $\sigma^2\gamma_s$.

REMARK 4.5 (Gabor's Idea).

a. Using *Theorem 4.1* and the terminology of *Definition 4.4*, we shall expand on our remarks from *Section 3* on Gabor's paper **[Gab46]** in which he formulated a fundamental method for obtaining signal decomposition in terms of "elementary signals". For reasons related to the Classical Uncertainty Principle Inequality (4.1), e.g., part c, Gabor chose as his sequence of elementary signals the Gaussians φ of (4.2) for a prescribed sequence of time-frequency parameters $(t_0, \gamma_0) \in \mathbb{R} \times \widehat{\mathbb{R}}$. His basic idea is encapsulated in his statement that

> "The elementary signals...assure the best utilization of the information area in the sense that they possess the smallest product of effective duration by effective frequency width" [**Gab46**, page 437].

b. Gabor's approach required a tiling of the time-frequency plane $\mathbb{R} \times \widehat{\mathbb{R}}$ by half-open rectangles $R_{m,n}$ centered at $(n\sigma^2 t_s, m\sigma^2\gamma_s) \in \mathbb{R} \times \widehat{\mathbb{R}}$, where $\sigma^2 t_s$ and $\sigma^2\gamma_s$ are defined in (4.7) and (4.8), and where $n, m \in \mathbb{Z}$. In order that the tiling consist of non-overlapping rectangles, the length of each $R_{m,n}$ is $\sigma^2 t_s$ and its width (height) is $\sigma^2\gamma_s$.

For each rectangle $R_{m,n}$, Gabor associated the Gaussian

$$g_{m,n}(t) = e^{2\pi i t m\sigma^2\gamma_s} e^{-s(t-n\sigma^2 t_s)^2}, \tag{4.10}$$

which is the function φ of (4.2) for the pair $(t_0, \gamma_0) = (n\sigma^2 t_s, m\sigma^2\gamma_s)$. By computing $\widehat{g}_{m,n}$, we see that each $g_{m,n}$ is "concentrated" in $R_{m,n}$ in the sense that $g_{m,n}$, resp., $\widehat{g}_{m,n}$, is small outside of an interval centered at $n\sigma^2 t_s$ and of length $\sigma^2 t_s$, resp., centered at $m\sigma^2\gamma_s$ and of length $\sigma^2\gamma_s$. The extent of concentration in the time or frequency direction depends on the value of $s > 0$.

He then derived signal decompositions of the form

$$f = \sum_{m,n} c_{m,n} g_{m,n}, \tag{4.11}$$

from which he could conclude that each coefficient $c_{m,n}$ contains the amount of "information" in f determined by the subset $R_{m,n}$ of the time-frequency plane, cf., von Neumann's decompositions in [**vN55**, pages 405–407].

c. The reason Gabor chose $\{g_{m,n}\}$, and hence the tiling $\{R_{m,n}\}$, was because of *Theorem 4.1* and equation (4.9). From his point of view, the product of temporal and frequency variances was a natural notion of "information area" (his term from the quotation in part a) with which to construct a tiling of $\mathbb{R} \times \widehat{\mathbb{R}}$. Further, because $\sigma^2 t_s \sigma^2\gamma_s$ is the smallest such area, he chose $\{g_{m,n}\}$ as the sequence of "harmonics" in (4.11). In fact, each tile $R_{m,n}$ of the time-frequency plane should be sufficiently localized in time and frequency so that smaller subtiles are not required in order to glean any more variance information from f than is available from the coefficient $c_{m,n}$.

d. Gabor's idea has been the inspiration for a plethora of time-frequency analysis studies. His use of the Gaussian in signal decomposition has been the wellspring for many ideas, some of which argue convincingly against its effectiveness, see *Section 6.*

5. Gabor and wavelet frames

Let H be a separable complex Hilbert space with inner product $\langle x, y\rangle$ and norm $\|x\| = \langle x, x\rangle^{1/2}$.

DEFINITION 5.1 (Bases).

a. A sequence $\{x_n : n \in \mathbb{Z}^d\} \subseteq H$ is a *Schauder basis* or *basis* for H if each $y \in H$ has a unique decomposition $y = \sum c_n(y) x_n$ in H, where $\{c_n(y)\} \subseteq \mathbb{C}$.

b. A basis $\{x_n\}$ for H is an *unconditional basis* for H if

$$\exists C > 0 \text{ such that } \forall F \subseteq \mathbb{Z}^d, \text{ finite, and } \forall \{b_j, c_j : j \in F\} \subseteq \mathbb{C},$$

$$\text{where } |b_j| \leq |c_j| \text{ for each } j \in F, \text{ we have}$$

$$\|\textstyle\sum_{n \in F} b_n x_n\| \leq C \|\sum_{n \in F} c_n x_n\|.$$

An unconditional basis $\{x_n\} \subseteq H$ is *bounded* if

$$\exists A, B > 0 \text{ such that } \forall n \in \mathbb{Z}^d, \ A \leq \|x_n\| \leq B.$$

c. It is well known that separable Hilbert spaces have orthonormal bases (ONBs) **[GG81]**; and it is elementary to see that ONBs are bounded unconditional bases.

DEFINITION 5.2 (Frames).
a. A sequence $\{x_n : n \in \mathbb{Z}^d\} \subseteq H$ is a *frame* for H if there exist $A, B > 0$ such that

$$\forall y \in H, \ A\|y\|^2 \leq \sum |\langle y, x_n \rangle|^2 \leq B\|y\|^2. \tag{5.1}$$

A and B are the *frame bounds*, and a frame is *tight* if $A = B$. A frame is *exact* if it is no longer a frame whenever any one of its elements is removed.

b. The *frame operator* of the frame $\{x_n\}$ is the function $S : H \to H$ defined as $Sy = \sum \langle y, x_n \rangle x_n$ for all $y \in H$.

c. The theory of frames is due to Duffin and Schaeffer **[DS52]**, cf., **[You80]**, **[DGM86]**, **[Dau92]**, **[BF94**, Chapter 3]. An exact frame is a bounded unconditional basis and vice-versa, e.g., **[You80]**.

THEOREM 5.3 (Frame Decomposition Theorem). *Let $\{x_n : n \in \mathbb{Z}^d\} \subseteq H$ be a frame for H with frame bounds A and B.*

a. The frame operator S is a topological isomorphism with inverse $S^{-1} : H \to H$. $\{S^{-1}x_n\}$ is a frame with frame bounds B^{-1} and A^{-1}, and

$$\forall y \in H, \ y = \sum \langle y, S^{-1}x_n \rangle x_n = \sum \langle y, x_n \rangle S^{-1}x_n \ \text{ in } H. \tag{5.2}$$

b. If $\{x_n\}$ is a tight frame for H, if $\|x_n\| = 1$ for all n, and if $A = B = 1$, then $\{x_n\}$ is an orthonormal basis for H.

c. If $\{x_n\}$ is an exact frame for H, then $\{x_n\}$ and $\{S^{-1}x_n\}$ are biorthonormal, i.e.,

$$\forall m, n, \ \langle x_m, S^{-1}x_n \rangle = \delta(m, n),$$

and $\{S^{-1}x_n\}$ is the unique sequence in H which is biorthonormal to $\{x_n\}$.

d. If $\{x_n\}$ is an exact frame for H, then the sequence resulting from the removal of any one element is not complete in H, i.e., the linear span of the resulting sequence is not dense in H.

REMARK 5.4 (Frames and Coherent Light: an Analogy).
a. If $\{x_n : n \in Z^d\}$ is a frame for H with frame bounds A and B and frame operator S, then it is easy to see that

$$\|I - \frac{2}{A+B} S\| \leq \frac{B-A}{A+B} < 1, \tag{5.3}$$

where $I : H \to H$ is the identity operator. The inequality (5.3) allows us to prove that

$$S^{-1} = \frac{2}{A+B}\sum_{k=0}^{\infty}(I - \frac{2}{A+B}S)^k,$$

which, in turn, can be used to prove part *a* of *Theorem 5.3*.

We also mention (5.3) because of the notion of *visibility* V, which is defined as

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}},$$

where $I_{\max}$ and $I_{\min}$ are maximum and minimum light intensities, e.g., [**Kla96**]. In the case of full interference of light waves (for the classical two-slit experiment), one has $I_{\min} = 0$; and, hence, the visibility is 1, a value associated with *coherent light*. In the context of frames, the analogy is that $A = 0$, so that a frame is not obtained.

b. The proof of *Theorem 5.3* can be found in [**BF94**, Chapter 3]. A strong form of *Theorem 5.3b* was originally proved by Vitali (1921), see [**Ben92**, pages 455–456] for details.

THEOREM 5.5 (Characterization of Frames).

a. A sequence $\{x_n : n \in \mathbb{Z}^d\}$ *in* H *is a frame for* H *with frame bounds* A *and* B *if and only if the mapping*

$$\begin{array}{rcl} L : H & \longrightarrow & \ell^2(\mathbb{Z}^d) \\ y & \longmapsto & \{\langle y, x_n\rangle\} \end{array}$$

is a topological isomorphism of H *onto a closed subspace of* $\ell^2(\mathbb{Z}^d)$. *In this case,*

$$\|L\| \le B^{1/2} \text{ and } \|L^{-1}\| \le A^{-1/2},$$

where L^{-1} *is defined on the range* $L(H)$.

b. A sequence $\{x_n : n \in \mathbb{Z}^d\}$ *in* H *is a frame for* H *if and only if there is* $C > 0$ *such that for all* $y \in H$

$$\Sigma|\langle y, x_n\rangle|^2 < \infty,$$

$$\exists c_y = c_n \in \ell^2(\mathbb{Z}^d), \text{ such that } y = \Sigma c_n x_n \text{ in } H,$$

and

$$\|c_y\|_{\ell^2(\mathbb{Z}^d)} \le C\|y\|.$$

Part *a* of *Theorem 5.5* is proved in [**BF94**, Theorem 7.15]; and part *b*, which we observed with David Walnut, is proved in [**BF94**, Remark 3.9].

DEFINITION 5.6 (Gabor and Wavelet Systems).

a. Let $g \in L^2(\mathbb{R})$ and let $a, b > 0$. The *Gabor system* for g and (a, b) is the sequence $\{\varphi_{m,n} : (m,n) \in \mathbb{Z} \times \mathbb{Z}\}$, where

$$\varphi_{m,n}(t) = e^{2\pi i t m b} g(t - na) = e_{mb}(t)\tau_{na}g(t),$$

and where $e_\gamma(t) = e^{2\pi i t \gamma}$ and $\tau_x g(t) = g(t - x)$.

b. Let $\psi \in L^2(\mathbb{R})$. The *affine system* or *wavelet system* for ψ is the sequence $\{\psi_{m,n} : (m,n) \in \mathbb{Z} \times \mathbb{Z}\}$, where

$$\psi_{m,n}(t) = 2^{m/2}\,\psi(2^m t - n).$$

c. The Fourier transforms of $\varphi_{m,n}$ and $\psi_{m,n}$ are easily computed to be

$$\widehat{\varphi}_{m,n}(\gamma) = e^{2\pi i n a m b} e^{-2\pi i n a \gamma} \widehat{g}(\gamma - mb) = \tau_{mb}(e_{-na}\widehat{g})(\gamma)$$

and

$$\widehat{\psi}_{m,n}(\gamma) = 2^{-m/2} e^{-2\pi i n(\gamma/2^m)} \widehat{\psi}(\gamma/2^m) = 2^{-m/2}(e_{-n}\widehat{\psi})(\gamma/2^m).$$

d. Let $H = L^2(\mathbb{R})$, $g \in L^2(\mathbb{R})$, and $a, b > 0$. If the Gabor system $\{\varphi_{m,n}\}$ for g and (a, b) is a frame, then it is a *Gabor frame.*

Let $H = L^2(\mathbb{R})$ and $\psi \in L^2(\mathbb{R})$. If the wavelet system $\{\psi_{m,n}\}$ for ψ is a frame, then it is a *wavelet frame.*

EXAMPLE 5.7 (Decompositions from Various Fields).

a. In **[Gab46]**, Gabor used the Gabor system $\{\varphi_{m,n}\}$, where $g \in L^2(\mathbb{R})$ is the Gaussian $g(t) = e^{-st^2}$, $s > 0$, and where $(a, b) = (\sigma^2 t_s, \sigma^2 \gamma_s)$. The variances $\sigma^2 t_s$ and $\sigma^2 \gamma_s$ were defined in *Definition 4.4*, and Gabor's functions $\varphi_{m,n}$ were the functions $g_{m,n}$ of (4.10). He then developed decompositions (4.11) which have inspired the theory of Gabor frames, e.g., **[Dau92]**, **[BF94**, Chapter 3], **[HW89]**.

b. A *wavelet* is a function $\psi \in L^2(\mathbb{R})$ for which the wavelet system $\{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$. The first wavelet ONB for $L^2(\mathbb{R})$ is due to Haar (1910), who made the construction for $L^2[0, 1]$, cf., **[Dau92]** for $L^2(\mathbb{R})$. The *Haar wavelet* is

$$\psi(t) = \begin{cases} 1, & \text{if } 0 \le t < 1/2 \\ -1, & \text{if } 1/2 \le t < 1 \\ 0, & \text{otherwise.} \end{cases}$$

There were related functions introduced by Rademacher, Walsh, and Franklin through the 1920s, e.g., **[Hig77]**. (Philip Franklin was Wiener's brother-in-law.) Strömberg **[Str83]** constructed a modified Franklin system and other spline systems used as unconditional bases for H^p-spaces. The work of Ciesielski, Carleson, and Maurey on H^1 immediately precedes Strömberg's.

Compactly supported n-times continuously differentiable wavelets were constructed by Daubechies in 1987, see **[Dau92]**. During the same period Mallat and Meyer introduced the concept of multiresolution analysis (MRA) which is fundamental for wavelet decompositions **[Mal89]**, **[Mey90]**. MRAs had important theoretical and practical precedents in the fields of speech and image processing, which use notions such as quadrature mirror filters and Laplace pyramidal schemes, e.g., **[VK95]**.

The Walsh functions (1923) mentioned above give rise to so-called *wavelet packet decompositions*, which were introduced by Coifman, Meyer, and Wickerhauser, e.g., **[CMW92]**.

c. Calderón's formula (1964)

$$f(u) = \int_0^\infty \left(\int \psi_s(u - t)\psi_s * f(t)dt \right) \frac{ds}{s}, \tag{5.4}$$

where $\psi_s(u) = (1/s)\psi(u/s)$, is a *continuous wavelet transform* of $f \in L^2(\mathbb{R})$, where $\psi \in S(\mathbb{R})$ is real and even, is supported by $[-1, 1]$, and satisfies $\int_0^\infty \widehat{\psi}(\gamma)^2 d\gamma/\gamma = 1$, e.g., **[FJW91]**. The function f is represented by the double integral on the right side of (5.4) instead of the double sum of part *b*. The ONB $\{\psi_{m,n}\}$ of dilates and translates of part *b* is replaced by the "redundant" system $\{\psi_s(\cdot - t) : s > 0,\ t \in \mathbb{R}\}$ of dilates and translates.

d. Let $H = L^2[-\Omega, \Omega]$ and let $\{t_n\} \subseteq \mathbb{R}$ be a sequence of distinct points. The sequence $\{e^{2\pi i t_n \gamma}\} \subseteq L^2[-\Omega, \Omega]$ is a *Fourier system*, and results of Duffin and

Schaeffer, H. Landau, and S. Jaffard allow a characterization of *Fourier frames* in terms of a density condition, e.g., [**BF94**, Theorem 7.44]. Fourier frames are useful in the design of auditory models for the purpose of speech compression [**BT93**], and in irregular sampling theory, e.g., [**BF94**, Section 7.7] which includes results we proved with William Heller.

DEFINITION 5.8 (Short Time Fourier Transform (STFT)). Let $g \in L^2(\mathbb{R})$ and define

$$\forall f \in L^2(\mathbb{R}) \text{ and } \forall (t,\gamma) \in \mathbb{R} \times \widehat{\mathbb{R}},\ S_g(f,t)(\gamma) = \int f(u)\overline{g(u-t)}e^{-2\pi i u\gamma}du.$$

If g is concentrated at the origin in some well-defined way, e.g., if $g = \mathbf{1}_{[-T,T]}$, then $S_g(f,t)$ is the *short time Fourier transform* (STFT) of f at $t \in \mathbb{R}$ depending on g.

REMARK 5.9 (Gabor Systems and Spectrograms).

a. Suppose $g \in L^2(\mathbb{R})$ is the Gaussian and $\{\varphi_{m,n}\}$ is a Gabor frame for g and (a,b). Then the coefficients of $\{\varphi_{m,n}\}$ in the definition of the frame operator S operating on $f \in L^2(\mathbb{R})$ are precisely the values $S_g(f,na)(mb)$ of the STFT.

b. If $g \in L^2(\mathbb{R})$ and the Gabor system $\{\varphi_{m,n}\}$ for g and (a,b) is an ONB, then for each $f \in L^2(\mathbb{R})$,

$$f = \sum S_g(f,na)(mb)\varphi_{m,n}.$$

c. Even though a Gabor system $\{\varphi_{m,n}\}$ for a given $g \in L^2(\mathbb{R})$ and (a,b) is not necessarily an ONB, parts *a* and *b* show that the sequence

$$\{|S_g(f,na)(mb)| : (m,n) \in \mathbb{Z} \times \mathbb{Z}\}$$

of amplitude data provides important information in characterizing signals $f \in L^2(\mathbb{R})$. This data can be represented by gray-scale values for each point (na, mb) of the time-frequency plane. This representation is a *spectrogram*, e.g., *Figure 3* of *Section 8*.

d. Historically, a rationale for spectrograms comes from the periodogram analysis mentioned in *Section 2* as well as from the idea behind Michaelson's interferometer, e.g., [**Ben96**, Sections 2.8.6–2.8.10].

We now define the notion of a stationary sequence in order to make a distinction in *Example 5.12* between Gabor and wavelet systems.

DEFINITION 5.10 (Stationary Sequences). A sequence $\{x_n : n \in \mathbb{Z}^d\} \subseteq H$ is *stationary* if for each $n \in \mathbb{Z}^d$ the inner products

$$R_n = \langle x_{n+k},\ x_k \rangle$$

are independent of $k \in \mathbb{Z}^d$. In this case, $R = \{R_n\}$ is the *discrete stochastic autocorrelation* of $\{x_n\}$, cf., [**Kol41a**]. The Fourier series S, defined by $\sum R_n e^{2\pi i n\gamma}$ and denoted by $R^\vee = S$, is the *power spectrum* of $\{x_n\}$, see [**Ben96**, Chapter 3] for Fourier series. In this case, we also write $\hat{S} = R$.

The following result is a fundamental property of stationary sequences. Part *b* can be viewed as a stochastic version of the Wiener-Wintner Theorem (Theorem 2.6).

THEOREM 5.11 (Khinchin-Kolmogorov Theorem).

a. Let $\{x_n : n \in \mathbb{Z}^d\}$ be a stationary sequence having discrete stochastic autocorrelation R. Then $R^\vee = S$ is a positive measure on $\mathbb{T}^d$, i.e., $S \in M_+(\mathbb{T}^d)$.

b. Let $S \in M_+(\mathbb{T}^d)$. There is a stationary sequence $\{x_n : n \in \mathbb{Z}^d\}$ with discrete stochastic autocorrelation R for which $\widehat{S} = R$.

Part *a* of *Theorem 5.11* is a consequence of Herglotz' Theorem, e.g., [**Ben96**, Chapter 2]. Part *b* was proved by Khinchin (1934) on $\mathbb{R}$, by Wold (1938) on $\mathbb{Z}$, and by Kolmogorov in terms of stationary sequences [**Kol41a**], cf., [**Pri81**, pages 221–222]. *Theorem 5.11* is sometimes called the Wiener-Khinchin Theorem, but this is a misnomer. In fact, Wiener began using *Theorem 5.11* as such only after 1950.

EXAMPLE 5.12 (Stationarity and Gabor Systems).
a. Let $g \in L^2(\mathbb{R})$ and let $a, b > 0$, and consider the Gabor system $\{\varphi_{m,n}\}$. *If $ab = 1$, then the sequence $x : \mathbb{Z} \times \mathbb{Z} \to H = L^2(\mathbb{R})$, defined by $x_{m,n} = \varphi_{m,n}$, is a stationary sequence.* In the case of frames, note that $\{\varphi_{m,n}\}$ is exact for the case $ab = 1$.
b. If $\psi \in L^2(\mathbb{R})$, the wavelet system $\{\psi_{m,n}\}$ is *not* stationary.

6. The classical uncertainty principle for Gabor and wavelet frames

Let us begin this section by stating the Balian-Low Theorem.

THEOREM 6.1 (Balian-Low Theorem). *Let $\{\varphi_{m,n}\}$ be the Gabor system for $g \in L^2(\mathbb{R})$ and (a, b), and assume $\{\varphi_{m,n}\}$ is an exact frame for $L^2(\mathbb{R})$. Then $ab = 1$ and*

$$||tg(t)||_{L^2(\mathbb{R})}||\gamma\widehat{g}(\gamma)||_{L^2(\widehat{\mathbb{R}})} = +\infty.$$

See [**BHW95**] for proofs, history, and relevance of the result.

Using the terminology from *Definition 4.4*, we have the following consequence of *Theorem 6.1*.

THEOREM 6.2 (Products of Variances for Gabor Frames). *Let $\{\varphi_{m,n}\}$ be the Gabor system for $g \in L^2(\mathbb{R})$ and (a, b), and assume $||g||_{L^2(\mathbb{R})} = 1$ and $ab = 1$. If $\{\varphi_{m,n}\}$ is a Gabor frame, then*

$$\forall m, n \in \mathbb{Z},\ ||(t - t_{m,n})\varphi_{m,n}(t)||_{L^2(\mathbb{R})}||(\gamma - \gamma_{m,n})\widehat{\varphi}_{m,n}(\gamma)||_{L^2(\widehat{\mathbb{R}})} = \infty,$$

where $t_{m,n}$ and $\gamma_{m,n}$ are the expected values

$$t_{m,n} = \int t|\varphi_{m,n}(t)|^2 dt \quad \text{and} \quad \gamma_{m,n} = \int \gamma|\widehat{\varphi}_{m,n}(\gamma)|^2 d\gamma.$$

In light of Gabor's use of the Gaussian described in *Remark 4.5* and the fact that Fourier transforms of Gaussians are Gaussians, the Balian-Low Theorem asserts that Gabor systems for the Gaussian do not give rise to frame decompositions in the case $ab = 1$.

The following result summarizes the completeness and decomposition properties of Gabor systems of Gaussians.

THEOREM 6.3 (Gabor Systems of Gaussians). *Let $g(t) = (2s/\pi)^{1/4}e^{-st^2}$, for some $s > 0$, and let $a, b > 0$. Consider the Gabor system $\{\varphi_{m,n}\}$ for g and (a, b).*
a. The linear span of $\{\varphi_{m,n}\}$ is dense in $L^2(\mathbb{R})$ if and only if $ab \leq 1$.
b. The Gabor system $\{\varphi_{m,n}\}$ is a frame for $L^2(\mathbb{R})$ if and only if $ab < 1$.

As indicated in *Remark 4.5b* this type of result for the Gaussian was initiated by von Neumann [**vN55**, pages 405–407] in the 1930s. The proof of *Theorem 6.3a* can be found in [**BBGK71**], [**Per71**]. The proof in part *b* that if $ab < 1$ then $\{\varphi_{m,n}\}$ is a frame is due to [**SWa90**], cf., [**DGM86**], [**Dau90**]. The material in [**SWa90**] and the related work in [**Sei92**] deals with the notion of *Beurling density*, which itself was formulated to deal with the completeness problems dealt with by Paley and Wiener and referenced in *Section 3*. The proof that $\{\varphi_{m,n}\}$ is not a frame when $ab = 1$ follows from a direct calculation in [**BF94**, page 257], or a slightly more complicated calculation using Jacobi theta functions [**DGM86**, pages 1274–1275]. This direction of *Theorem 6.3b* is also an immediate consequence of *Theorem 6.1*.

In contrast to *Theorem 6.2* we have the following fact, noted by Meyer in his Seminaire Bourbaki, 1985–1986, number 622, e.g., [**Mey90**].

THEOREM 6.4 (Products of Variances for Wavelet Systems). *Let $\{\psi_{m,n}\}$ be the wavelet system for $\psi \in L^2(\mathbb{R})$, and assume $\|\psi\|_{L^2(\mathbb{R})} = 1$. If $t\psi(t) \in L^2(\mathbb{R})$ and $\gamma\widehat{\psi}(\gamma) \in L^2(\widehat{\mathbb{R}})$, then*

$$\forall n \in \mathbb{Z}, \quad \sup_m \|(t - t_{m,n})\psi_{m,n}(t)\|_{L^2(\mathbb{R})} = \infty,$$

$$\forall n \in \mathbb{Z}, \quad \sup_m \|(\gamma - \gamma_{m,n})\widehat{\psi}_{m,n}(\gamma)\|_{L^2(\widehat{\mathbb{R}})} = \infty,$$

and

$$\sup_{m,n} 4\pi \|(t - t_{m,n})\psi_{m,n}(t)\|_{L^2(\mathbb{R})} \|(\gamma - \gamma_{m,n})\widehat{\psi}_{m,n}(\gamma)\|_{L^2(\widehat{\mathbb{R}})} < \infty,$$

where $t_{m,n}$ and $\gamma_{m,n}$ are the expected values

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

In light of *Theorem 6.2* and *Theorem 6.4*, it natural to ask if there are frames $\{\theta_n\}$ for $L^2(\mathbb{R})$, with each $\|\theta_n\|_{L^2(\mathbb{R})} = 1$ and expected values

$$t_n = \int t|\theta_n(t)|^2 dt \quad \text{and} \quad \gamma_n = \int \gamma|\widehat{\theta}_n(\gamma)|^2 d\gamma,$$

such that

$$\sup_n \|(t - t_n)\theta_n(t)\|_{L^2(\mathbb{R})} < \infty \quad \text{and} \quad \sup_n \|(\gamma - \gamma_n)\widehat{\theta}_n(\gamma)\|_{L^2(\widehat{\mathbb{R}})} < \infty.$$

Bourgain has answered this question, which is essentially due to Balian [**Bal81**], in the following way [**Bou88**].

THEOREM 6.5 (Bourgain Theorem). *For every $\epsilon > 0$ there is an orthonormal basis $\{\theta_{m,n}\}$ for $L^2(\mathbb{R})$ having expected values $(t_{m,n}, \gamma_{m,n}) \in \mathbb{R} \times \widehat{\mathbb{R}}$ and satisfying the inequalities*

$$\sup_{m,n} \|(t - t_{m,n})\theta_{m,n}(t)\|_{L^2(\mathbb{R})} < \frac{1}{2\sqrt{\pi}} + \epsilon$$

and

$$\sup_{m,n} \|(\gamma - \gamma_{m,n})\widehat{\theta}_{m,n}(\gamma)\|_{L^2(\widehat{\mathbb{R}})} < \frac{1}{2\sqrt{\pi}} + \epsilon.$$

Consequently (by Theorem 4.1*),*

$$\forall m, n \in \mathbb{Z}, \quad 1 = \|\theta_{m,n}\|^2_{L^2(\mathbb{R})}$$

$$\leq 4\pi \|(t - t_{m,n})\theta_{m,n}(t)\|_{L^2(\mathbb{R})} \|(\gamma - \gamma_{m,n})\widehat{\theta}_{m,n}(\gamma)\|_{L^2(\widehat{\mathbb{R}})}$$

$$< 4\pi \left(\frac{1}{2\sqrt{2}} + \epsilon\right)^2.$$

7. Gabor decomposition of L^1

The Gabor and wavelet frame decompositions of *Section 5* are in the setting of the Hilbert space $L^2(\mathbb{R})$. On the other hand, as mentioned in *Remark 2.11*, a major problem in signal processing and other scientific areas is to provide an effective analysis of noisy signals. In particular, as mentioned at the beginning of *Section 3*, we would like to extricate intelligent messages embedded in noisy environments. $L^2(\mathbb{R})$ is generally not an adequate setting in which to model a variety of noises.

In this section we shall give a Gabor decomposition for the $L^1 - L^\infty$ duality since it is not unreasonable to model some noises in L^∞. We shall state a Gabor decomposition on $\mathbb{R}$ for continuous Gabor systems (*Definition 7.1*). We can prove discretizations of this decompositions as well as multidimensional generalizations. The autocorrelation defined in *Definition 2.5* plays a role in the main result (*Theorem 7.2*). We have chosen continuous Gabor systems in order to show that *Theorem 7.2* is a generalization of the Fourier inversion formula (*Remark 7.3*).

DEFINITION 7.1 (Continuous Gabor Systems).

a. Let $g \in L^1_{\text{loc}}(\mathbb{R})$. The *continuous Gabor system* for g is the family of functions

$$\{\varphi_g(\,\cdot\,; x, \gamma, c) : (x, \gamma) \in \mathbb{R} \times \widehat{\mathbb{R}}, c \in \mathbb{R}\}, \tag{7.1}$$

where

$$\varphi_g(t; x, \gamma, c) = e^{-2\pi c x \gamma}(e_\gamma \tau_x g)(t) = e^{-2\pi c x \gamma} e^{2\pi i t \gamma} g(t - x) \tag{7.2}$$

is a function of t on $\mathbb{R}$ for any fixed $(x, \gamma) \in \mathbb{R} \times \widehat{\mathbb{R}}$ and $c \in \mathbb{R}$. The modulation and translation functions e_γ and τ_x are defined in *Definition 5.6a* to denote the right side of (7.2).

A continuous Gabor system is also referred to as a *Gabor system of coherent states.*

b. Let $g \in L^1_{\text{loc}}(\mathbb{R})$ and consider the continuous Gabor system $\{\varphi_g(\,\cdot\,; x, \gamma, c)\}$ for the case $c = 0$. The *continuous Gabor transform* of any $f \in L^1_{loc}(\mathbb{R})$ is then formally defined on $\mathbb{R} \times \widehat{\mathbb{R}}$ as

$$G_g(f)(x, \gamma) = \int f(t)\varphi_g(t; x, \gamma)dt, \tag{7.3}$$

where $\varphi_g(t; x, \gamma)$ denotes $\varphi_g(t; c, \gamma, 0)$. The role of this formal inner product in (7.3) becomes apparent in the statement of *Theorem 7.2*.

c. The Heisenberg group of 3×3 matrices arises in defining unitary operators by means of (7.1), and in particular when c ranges over $\mathbb{R}$, e.g., [**Ben89**].

We shall now give an integral representation of $L^1(\mathbb{R})$ in terms of the continuous Gabor transform. To this end, we shall deal with functions $g \in L^\infty(\mathbb{R})$ and sequences $\{\widehat{k}_n\} \subseteq L^1(\widehat{\mathbb{R}})$ with the property that $\{k_n\} \subseteq L^1(\mathbb{R})$ is an L^1-approximate identity. Recall that a sequence $\{k_n\} \subseteq L^1(\mathbb{R})$ is an *L^1-approximate identity* if $\int k_n(t)dt = 1$ for all n, $\sup_n \|k_n\|_{L^1(\mathbb{R})} < \infty$, and

$$\forall r > 0, \lim_{n\to\infty} \int_{|t|\geq r} |k_n(t)|dt = 0.$$

THEOREM 7.2 (Gabor Representation for $L^1(\mathbb{R})$). *Let $g \in L^\infty(\mathbb{R})\backslash\{0\}$, and consider the corresponding Gabor system $\{\varphi_g(\,\cdot\,; x, \gamma)\}$ and continuous Gabor transform (of Definition 7.1b). Assume that*

$$\forall t \in \mathbb{R},\ R_g(t) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} g(t+u)\overline{g(u)}du$$

exists. Then, for each $f \in L^1(\mathbb{R})$,

$$\lim_{n\to\infty} \|f - f_n\|_{L^1(\mathbb{R})} = 0,$$

where

$$f_n(t) = \frac{1}{R_g(0)} \lim_{T\to\infty} \frac{1}{2T} \int \int_{T}^{T} G_g(f)(x,\gamma)\varphi_g(t; x, \gamma)\widehat{k}_n(\gamma)dxd\gamma$$

and where $\{\widehat{k}_n) \subseteq L^1(\widehat{\mathbb{R}})$ has the property that $\{k_n\}$ is an L^1-approximate identity.

For the proof, see [**Ben89**].

REMARK 7.3 (Generalization of the Fourier transform).

a. If $g = 1$ on $\mathbb{R}$, $f \in L^1(\mathbb{R} \times \widehat{\mathbb{R}})$, and $c = 0$ in the definition of φ_g, then $G_1(f)(x,\gamma) = \widehat{f}(\gamma)$ for $(x,\gamma) \in \mathbb{R}\times\widehat{\mathbb{R}}$. In this case, the assertion of *Theorem 7.2* is a Fourier inversion formula.

b. $R_g(0)$ in *Theorem 7.2* is the mean total power of g, e.g., *Definition 2.5b*.

8. An interpretation of spectral bioelectric data

The most common signals derived from brain potentials are electroencephalograms (EEGs). Theoretically, EEG time series should provide quantitative data to fathom and describe normal brain rhythms, as well as aberrations such as epileptic seizures. Norbert Wiener's interest in brain rhythms goes back to the 1930s, e.g., [**Mas90**, pages 233–238]; and he was directly involved in obtaining the power spectra of brain waves from autocorrelations [**56e**][**61c**, Chapter X].

In this section, we shall address a problem in epilepsy using harmonic analysis techniques. These techniques are in the spirit of Wiener's approach, but are both less sophisticated mathematically and more effective for dealing with single time series because of our use of STFTs and spectrograms, e.g., *Definition 5.8* and *Remark 5.9* on Gabor systems and spectrograms. Wiener's original approach was based on GHA (*Section 2*) as well as a longstanding point of view from optics using Michelson's interferometer (*Remark 5.9*). Naturally, we shall use the discrete Fourier transform (DFT) and associated fast Fourier transform (FFT) algorithm in our computation of spectrograms, see [**Ben96**, Sections 3.8 and 3.9] for the theory underlying these methods as well as references for their implementation.

Since EEG signals are measured on the scalp and potentials are on the order of microvolts, electroencephalograms are subject to many complicated influences such as head geometry, propagation of brain waves through the skull, and muscle movement. These effects are often regarded as "noise", and, consequently, the signal-to-noise ratio of EEG time series can be quite low. Time series which have a much higher signal-to-noise ratio are obtained by measuring potentials directly on the surface of the brain. These are called electrocorticograms (ECoGs), and can only be obtained by invasive procedures. Because of the serious problems inherent in such procedures, it should be pointed out that Wiener proposed to use GHA in a fundamental way by dealing with EEG data in terms of averaged crosscorrelations of stimulus and response. This is the direction that was pursued by Rosenblith of MIT.

The bioelectric trace f at the top of *Figure 2* is electrocorticogram (ECoG) data. The analysis of this data was made at The MITRE Corporation, e.g., [**BC95**], [**B-KCJ94**].

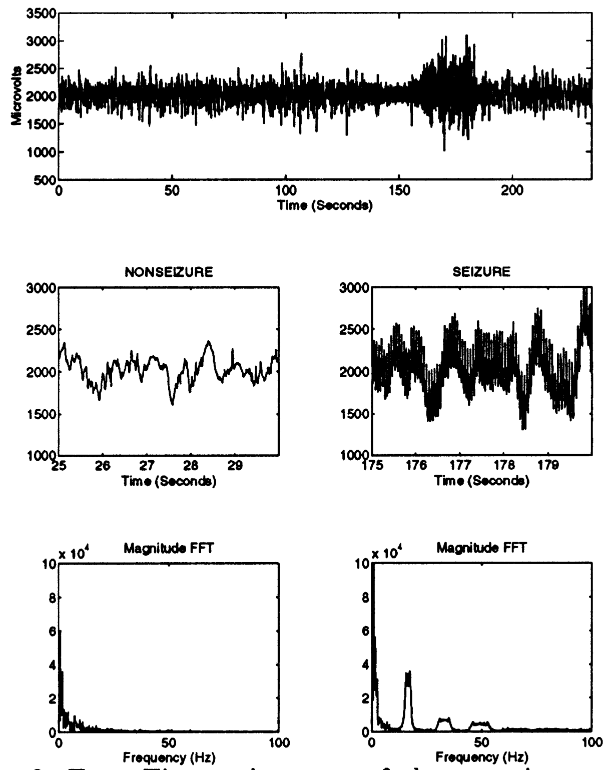

FIGURE 2. Top: Time series trace of electrocorticogram data. Middle: Time series for nonseizure epoch (left) and seizure epoch (right). Bottom: Magnitude FFT for the time series shown in the middle.

The length of the trace is 240 seconds, and it includes (epileptic) seizure activity as well as significant nonseizure activity. Typical of the latter type is activity f_{ns} in the time interval from 25 to 30 seconds, extracted from the original trace f; "ns" designates "nonseizure". f_{ns} is reproduced in the left graph on the second line of *Figure 2*. The ordinate is a microvolt measurement of the potential of the electric field at the surface of the brain relative to a referential electrode. The right graph on the second line of *Figure 2* is activity f_s extracted from seizure data in the original trace f; "s" designates "seizure". (We admit to begging the question of what precisely is seizure data.) The third line of *Figure 2* contains absolute values of the fast Fourier transforms (FFTs) of f_{ns} and f_s, respectively. The *dc* components "$\int f_{ns}(t)dt$" and "$\int f_s(t)dt$" have been removed from the FFTs since, as can be seen from line 2 of *Figure 2* , these values are so large as to affect the readability of nonzero-frequency information. The left graph on the third line of *Figure 2* is indicative of $1/f$ noise, whereas the right graph exhibits some periodic behavior. We shall now look at this latter issue a little more closely.

The right graph on the second line of *Figure 2* exhibits a high frequency almost periodic signal f "riding-on" a low frequency wave. This latter wave is usually in the θ-band $(4-8Hz)$ or the α-band $(8-13Hz)$ of brain activity. The θ-band is the spectral range of drowsiness or light sleep, and the α–band is the spectral range of rhythmic activity in an awake person. The amplitude of the bioelectric trace for the α-band is typically between 5 and 100 microvolts, e.g., [**Nun81**].

A close look at f justifies our labeling of f as an *almost periodic* signal. To see this, first let f_1, resp., f_2, be the restriction of f to a time interval at the beginning, resp., the end, of the seizure. Then, by counting peaks of the trace over the duration of the seizure activity, we find that the period p_1 of f_1 is less than the period p_2 of f_2.

Assuming each $f_j, j = 1, 2$, is periodic on $\mathbb{R}$, even though it is only periodic on a portion of the seizure interval, we have the Fourier series representation,

$$Sf_j(t) = \sum_n a_{j,n} e^{2\pi i n t/p_j}; \tag{8.1}$$

and we further assume that $f_j = Sf_j$. Because of (8.1), the Fourier transform of each f_j, as a distribution on $\mathbb{R}$, is

$$\hat{f_j} = \sum_n a_{j,n} \delta_{n/p_j}. \tag{8.2}$$

As such, if we consider the (t, γ)–phase plane as the domain of the spectrogram, and if we take a value t_1 near the beginning of the seizure, then because of (8.2) we expect nontrivial spectrogram amplitudes at the points $\{(t_1, n/p_1) : n \in \mathbb{Z}\}$. A similar remark holds for other parts of the seizure. In particular, if t_2 is a time value at the end of the seizure, we expect nontrivial spectrogram amplitudes at the points $\{t_2, n/p_2 : n \in \mathbb{Z}\}$.

Finally, for a fixed $n > 0$, we expect the graph of this spectral data in the seizure interval to be a decreasing function of t since $1/p_1 > 1/p_2$, cf., (8.2). This time varying spectral activity can be described in terms of elementary chirps, e.g., [**Ben96**, Section 2.10].

If the previous discussion is too discursive, then experimental data reflected by *Figure 3* tell the same story pictorially. In fact, the region of the (t, γ)-phase plane determined by the seizure time interval [170,180] bears out the previous analysis (and handwaving).

The *prediction problem* is an important aspect of ECoG data analysis. This problem is to predict the onset of seizure sufficiently far in advance in order to take remedial action. Such action can only be effective if there is also a solution to the *localization problem.* The localization problem is to find the region of the brain responsible for the onset of seizure activity. In principle, solutions to the prediction

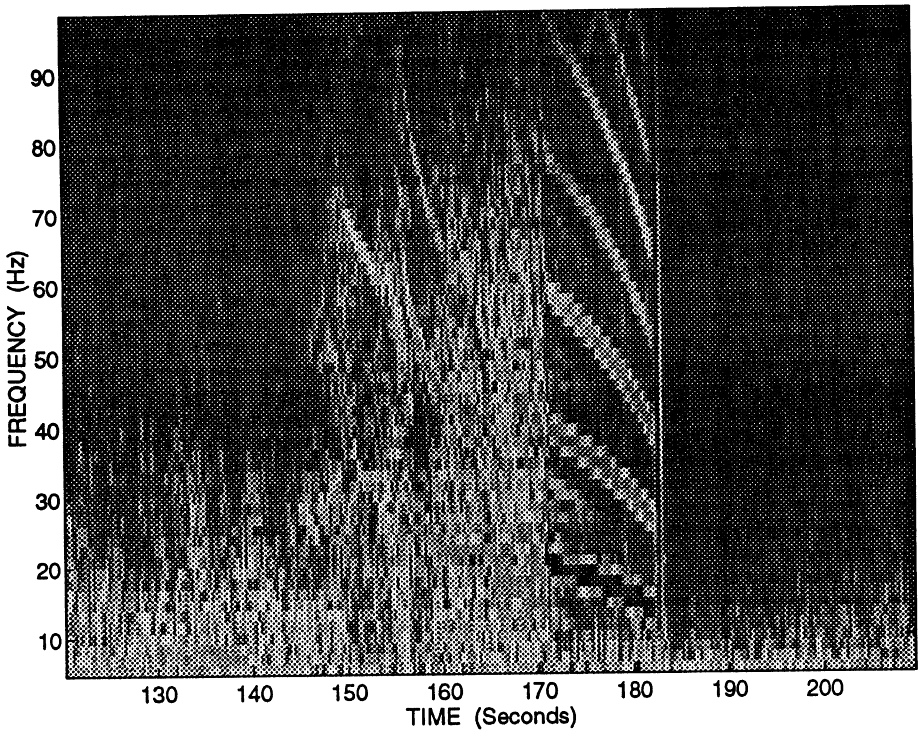

FIGURE 3. Spectrogram for the ECoG time series shown in *Figure 1*.

and localization problems would allow chemical response to specific local parts of the brain in time to temper seizure intensity.

The spectrogram in *Figure 3* provides information about the prediction problem. For example, the definitive chirps in the seizure time interval [170,180] of *Figure 3* have as precursors the periodic chirp-like activity in the time interval [155,162]. One can even detect some such activity in the time interval [150,155]. A close look at the trace on the top line of *Figure 2* shows that these precursors are embedded in low amplitude data.

9. Notation

Besides the standard notation found in the books by Hörmander [**Hör83**], Schwartz [**Sch66**], and Stein and Weiss [**SW71**], we shall also use the following notation and conventions.

The *Fourier transform* $\hat{f}$ of $f \in L^1(\mathbb{R}^d)$ is defined by $\hat{f}(\gamma) = \int f(t)e^{-2\pi i t\cdot\gamma}dt$, where $\mathbb{R}$ is the real line, $\mathbb{R}^d$ is d-dimensional Euclidean space, "$\int$" designates integration over $\mathbb{R}^d$, and $\gamma \in \widehat{\mathbb{R}}^d(= \mathbb{R}^d)$. Similarly, "$\sum$" designates summation over $\mathbb{Z}^d$, where $\mathbb{Z}$ is the ring of integers. $F^\vee$ designates the *inverse Fourier transform* of F. Formally, if $\hat{f} = F$, then

$$f(t) = F^\vee(t) = \int F(\gamma)e^{2\pi i t\cdot\gamma}d\gamma, \tag{9.1}$$

where integration is over $\widehat{\mathbb{R}}^d$, see [**Ben96**, Chapter 1], [**SW71**] for criteria for the validity of (9.1). $A(\widehat{\mathbb{R}}^d)$ is the space of *absolutely convergent Fourier transforms*.

The dual space of $A(\widehat{\mathbb{R}}^d)$, taken with the induced topology from $L^1(\mathbb{R}^d)$, is the space $A'(\widehat{\mathbb{R}}^d)$ of *pseudo-measures.*

$M(X)$ is the space of Radon measures on a locally compact space X. $M_+(X)$, resp., $M_b(X)$, designates the positive, resp., the bounded, elements of $M(X)$. $M_{b+}(X)$ is the space of bounded, positive elements of $M(X)$. If $X = \widehat{\mathbb{R}}^d$, then $M_b(\widehat{\mathbb{R}}^d) \subseteq A'(\widehat{\mathbb{R}}^d)$. Also, $A'(\widehat{\mathbb{R}}^d) = L^\infty(\mathbb{R}^d)^\wedge \subseteq \mathcal{S}'(\widehat{\mathbb{R}}^d)$, where $\mathcal{S}'(\widehat{\mathbb{R}}^d)$ is the space of tempered distributions.

Finally, we use the following notation:

$$(\tau_x f)(t) = f(t-x);$$

$$e_\gamma(t) = e^{2\pi i t\gamma};$$

$$\mathbf{1}_X(t) = \begin{cases} 1, & \text{if } t \in X \\ 0, & \text{if } t \notin X; \end{cases}$$

$$\delta(m,n) = \begin{cases} 1, & \text{if } m = n \\ 0, & \text{if } m \neq n; \end{cases}$$

$$\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d;$$

$$B(T) = \{t \in \mathbb{R}^d : |t| \leq T\}, \ |t| = (t_1^2 + \cdots + t_d^2)^{1/2};$$

supp f is the support of f;

$|X|$ is the Lebesgue measure of $X \subseteq \mathbb{R}^d$.

References

[AKM80] K. Anzai, S. Koizumi, and K. Matsuoka, *On the Wiener formula of functions of two variables*, Tokyo J. Math. **3** (1980), 249–270.

[Ars66] J. Arsac, *Fourier Transforms and the Theory of Distributions*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1966.

[Ash65] R. Ash, *Information Theory*, Interscience Publishers, New York, 1965.

[BBE89] J. J. Benedetto, G. Benke, and W. R. Evans, *An n-dimensional Wiener-Plancherel formula*, Advances in Applied Math, **10** (1989), 457–487.

[BBGK71] V. Bargmann, P. Butero, L. Giradello, and J. Klauder, *On the completeness of coherent states*, Rep. Mod. Phys. **2** (1971), 221–228.

[BC95] J. J. Benedetto and D. Colella, *Wavelet analysis of spectrogram seizure chirps*, SPIE, San Diego, 1995.

[BF94] J. J. Benedetto and M. W. Frazier, editors, *Wavelets: Mathematics and Applications*, CRC Press, Boca Raton, FL, 1994.

[BHW95] J. J. Benedetto, C. Heil, and D. F. Walnut, *Differentiation and the Balian-Low Theorem*, J. Fourier Analysis and Applications **1** (1995), 355–402.

[BM62] A. Beurling and P. Malliavin, *On Fourier transforms of measures with compact support*, Acta Math. **107** (1962), 291–309.

[BM67] A. Beurling and P. Malliavin, *On the closure of characters and the zeros of entire functions*, Acta Math. **118** (1967), 79–93.

[BT93] J. J. Benedetto and A. Teolis, *A wavelet auditory model and data compression*, Applied and Computational Harmonic Analysis **1** (1993), 3–28.

[B-KCJ94] M. Bozek-Kuzmicki, D. Colella, and G. Jacyna, *Feature-based epileptic seizure detection and prediction from ECoG recordings*, IEEE-SP Int'l Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, 1994.

[Bal81] R. Balian, *Un principe d'incertitude fort en théorie du signal ou en mécanique quantique*, C. R. Acad. Sci., Paris **292** (1981), 1357–1362.

[Bar68] J. L. Barnes, *Information theoretic aspects of feedback control systems*, Automatica **4** (1968), 165–185.

[Bar70] J. L. Barnes, *Laplace-Fourier transformation, the foundation for quantum information theory and linear physics*, in Problems in Analysis, R. C. Gunning, editor, Princeton University Press, 1970, pp.157–173.

[Bas84] J. Bass, *Fonctions de Corrélation, Fonctions Pseudo-Aléatoires, et Applications*, Masson, Paris, 1984.

[Ben75] J. J. Benedetto, *Spectral Synthesis*, Academic Press, New York, 1975.

[Ben89] J. J. Benedetto, *Gabor representations and wavelets*, Amer. Math. Soc. Contemp. Math. **91** (1989), 9-27.

[Ben90] J. J. Benedetto, *Uncertainty principle inequalities and spectrum estimation*, Recent Advances in Fourier Analysis and its Applications, J. S. and J. L.Byrnes, editors, Kluwer Academic Publishers, Boston **315C** (1990), 143–182.

[Ben91a] J. J. Benedetto, *A multidimensional Wiener-Wintner theorem and spectrum estimation*, Trans. Amer. Math. Soc. **327** (1991), 833–852.

[Ben91b] J. J. Benedetto, *The spherical Wiener-Plancherel formula and spectral estimation*, SIAM J. Math. Analysis **22** (1991), 1110–1130.

[Ben92] J. J. Benedetto, *Irregular sampling and frames*, in Wavelets–A Tutorial in Theory and Applications, C. K. Chui, editor, Academic Press, Boston, 1992, pp.445–507.

[Ben96] J. J. Benedetto, *Harmonic Analysis and Applications*, CRC Press, Boca Raton, FL, 1996.

[Ber87] J.-P. Bertrandias et al., *Espaces de Marcinkiewicz, Corrélations, Mesures, Systèmes Dynamiques*, Masson, Paris, 1987.

[Beu89] A. Beurling, *Collected Works*, Volumes 1 and 2, Birkhäuser, Boston, 1989.

[Boa54] R. P. Boas, *Entire Functions*, Academic Press, New York, 1954.

[Bou88] J. Bourgain, *A remark on the uncertainty principle for Hilbertian basis*, J. Funct. Analysis **79** (1988), 136–143.

[Bri81] D. R. Brillinger, *Time Series*, Holden-Day, Inc., San Francisco, 1981.

[CL90] Y.-Z. Chen and K.-S. Lau, *Wiener transformation of functions with bounded averages*, Proc. Amer. Math. Soc. **108** (1990), 411–421.

[CMW92] R. R. Coifman, Y. Meyer, and V. Wickerhauser, *Wavelet analysis and signal processing*, in Wavelets and their Applicatons, M.B. Ruskai, et al., editors, Jones and Bartlett Publishers, Boston, 1992.

[DGM86] I. Daubechies, A. Grossmann, and Y. Meyer, *Painless nonorthogonal expansions*, J. Math. Physics **27** (1986),1271–1283.

[DM76] H. Dym and H. P. McKean, *Gaussian Processes, Function Theory and the Inverse Spectral Problem*, Academic Press, New York, 1976.

[DS52] R. J. Duffin and A. C. Schaeffer, *A class of nonharmonic Fourier series*, Trans. Amer. Math. Soc. **72** (1952), 341–366.

[Dau90] I. Daubechies, *The wavelet transform, time-frequency localization and signal analysis*, IEEE Trans. Inf. Theory **39** (1990), 961–1005.

[Dau92] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Conf. Lect. Notes Ser. Appl. Math., vol. 61, Soc. Ind. Appl. Math., Philadelphia, PA, 1992.

[dBr67] N. G. de Bruijn, *Uncertainty principles in Fourier analysis*, in Inequalities, O. Shiska, editor, Academic Press, New York, 1967, pp.57–71.

[Ein14] A. Einstein, *Méthode pour la détermination de valeurs statistiques d'observations concernant des grandeurs soumises à des fluctuations irrégulières*, Archive des Sciences Physiques et Naturelles **37** (1914), 254–255.

[FJW91] M. W. Frazier, B. Jawerth, and G. Weiss, *Littlewood-Paley Theory and the Study of Function Spaces*, CBMS Conf. Lect. Notes, vol. 79, Amer. Math. Soc., Providence, R.I., 1991.

[FS85] J. Fournier and J. Stewart, *Amalgams of L^p and ℓ^p*, Bull. Amer. Math. Soc. **13** (1985), 1–21.

[GG81] I. Gohberg and S. Goldberg, *Basic Operator Theory*, Birkhäuser, Boston, 1981.

[Gab46] D. Gabor, *Theory of communication*, J.IEE London **93**, Part I (1946), 429–457.

[Gab65] D. Gabor, *Wiener and the art of communication*, Wiener Memorial Meeting, Istituto Internazionale delle Communicazioni, Genoa, 1965, 13 pages.

[HJ94] V. Havin and B. Jöricke, *The Uncertainty Principle in Harmonic Analysis*, Springer-Verlag, New York, 1994.

[HL95] J. A. Hogan and J. D. Lakey, *Extensions of the Heisenberg group by dilations and frames*, Appl. Comput. Harmonic Anal. **2** (1995), 174–199.

[HW89] C. Heil and D. Walnut, *Continuous and discrete wavelet transforms*, SIAM Review **31** (1989), 628–666.

[Har49] G. H. Hardy, *Divergent Series*, Oxford University Press, 1949.

[Hei96] C. Heil, *The Wiener transform on the Besicovitch spaces*, preprint.

[Hig77] J. Higgins, *Completeness and Basic Properties of Sets of Special Functions*, Cambridge University Press, 1977.

[Hör83] L. Hörmander, *The Analysis of Linear Partial Differential Operators, Volumes I and II*, Springer-Verlag, New York, 1983.

[Ker90] R. Kerby, *The correlation function and the Wiener-Wintner theorem in higher dimensions*, PhD thesis, University of Maryland, College Park, 1990.

[Kla96] J. R. Klauder, *Optical coherence before and after Wiener*, this volume.

[Kol41a] A. N. Kolmogorov, *Stationary sequences in Hilbert space*, Bull. Moscow State U., Math. 2(6), 1941, 40 pages. Translation by G. Kallianpur.

[Kol41b] A. N. Kolmogorov, *Interpolation und Extrapolation von stationären zuffälligen Folgen*, Izv. Akad. Nauk SSSR **5** (1941), 3–14.

[LP61] H. J. Landau and H. O. Pollak, *Prolate spheroidal wave functions, Fourier analysis and uncertainty II*, Bell System Tech. J. **40** (1961), 65–84.

[LP62] H. J. Landau and H. O. Pollak, *Prolate spheroidal wave functions, Fourier analysis and uncertainty III*, Bell System Tech. J. **41** (1962), 1295–1336.

[Lan67] H. Landau, *Necessary density condition for sampling and interpolation of certain entire functions*, Acta Math. **117** (1967), 37–52.

[Lee60] Y. Lee, *Statistical Theory of Communication*, John Wiley, New York, 1960.

[Lev40] N. Levinson, *Gap and Density Theorems*, Amer. Math. Soc. Colloq. Publ., vol. 26, Amer. Math. Soc., Providence, R.I., 1940.

[Lev66] N. Levinson, *Wiener's life*, Bull. Amer. Math. Soc. **72**, Part II (1966), 1-32. (Wiener Memorial Issue)

[Mal89] S. Mallat, *Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$*, Trans. Amer. Math. Soc. **315** (1989), 69–87.

[Mas66] P. Masani, *Wiener's contributions to Generalized Harmonic Analysis, prediction theory and filter theory*, Bull. Amer. Math. Soc. **72**, Part II (1966), 73–125. (Wiener Memorial Issue)

[Mas86] P. Masani, *Einstein's contribution to generalized harmonic analysis and his intellectual kinship with Norbert Wiener*, Jahrbuch Überblicke Mathematik **19** (1986), 191–209.

[Mas90] P. Masani, *Norbert Wiener : 1894–1964*, Birkhäuser, Boston, 1990.

[Mey90] Y. Meyer, *Ondelettes et Opérateurs*, Hermann, Paris, 1990.

[Nun81] P. Nunez, *Electric Fields of the Brain: The Neurophysics of EEG*, Oxford University Press, 1981.

[Ohk79] H. Ohkubo, *On generalized Fourier Stieltjes transforms*, Tokyo J. Math. **2** (1979), 99–110.

[PS61] H. O. Pollak and D. Slepian, *Prolate spheroidal wave functions, Fourier analysis and uncertainty I*, Bell System Tech. J. **40** (1961), 43–64.

[Pap77] A. Papoulis, *Signal Analysis*, McGraw-Hill Book Company, New York, 1977.

[Per71] A. Perelomov, *On the completeness of the system of coherent states*, Theor. Math. Physics **6** (1971), 156–164.

[Pie65] J. R. Pierce, *Symbols, Signals and Noise*, Harper & Brothers, New York, 1965.

[Pri81] M. B. Priestley, *Spectral Analysis and Time Series*, vol. 1, Academic Press, New York, 1981.

[Pric85] J. F. Price, editor, *Fourier Techniques and Applications*, Plenum Press, New York, 1985.

[Pro89] J. G. Proakis, *Digital Communications*, second edition, McGraw-Hill Book Company, New York, 1989.

[Rei68] M. Reiter, *Classical Harmonic Analysis and Locally Compact Groups*, Oxford University Press, 1968.

[Ric54] S. O. Rice, *Mathematical analysis of random noise*, in Noise and Stochastic Processes, N. Wax, editor, Dover Publications, Inc., New York, 1954. (Bell System Technical J., Volumes 23 and 24).

[SW71] E. Stein and G. Weiss, *An Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971.

[SWa90] K. Seip and R. Wallstén, *Sampling and interpolation in the Bargmann-Fock space*, Mittag-Leffler Technical Report, Djursholm, 1990.

[Sch66] L. Schwartz, *Théorie des Distributions*, Hermann, Paris, 1966. (1st ed., 1950).

[Sche90] W. Schempp, *Holographic image coding and neurocomputer architectures*, in Recent Advances in Fourier Analysis and its Applications, J. S. and J. L. Byrnes, editors, Kluwer Academic Publishers, Boston, MA **315C** (1990), 507–559.

[Seg47] I. E. Segal, *The group algebra of a locally compact group*, Trans. Amer. Math. Soc. **61** (1947), 69–105.

[Sei92] K. Seip, *Density theorems for sampling and interpolation in the Bargmann-Fock space*, Bull. Amer. Math. Soc. **26** (1992), 322–328.

[Sha49] C. E. Shannon, *The Mathematical Theory of Communication*, The University of Illinois Press, 1949.

[Sle76] D. Slepian, *On bandwidth*, Proc. IEEE **64** (1976), 292–300.

[Str83] J.-O. Strömberg, *A modified Franklin system and higher-order spline systems on R^n as unconditional bases for Hardy spaces*, in Conference on Harmonic Analysis in Honor of Antoni Zygmund, W. Beckner, A. P Calderón, R. Fefferman, and P. W. Jones, editors, Wadsworth International Group, Belmont, CA, 1983, pp.475–494.

[VK95] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1995.

[vN55] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955. (1st ed., 1932; 2nd ed., 1949).

[Wey50] H. Weyl, *The Theory of Groups and Quantum Mechanics*, Dover Publications, Inc., New York, 1950.

[Wie76] N. Wiener, *Collected Works*, The MIT Press, 1976. P. Masani, ed. Also 1979 and 1981.

[Yag85] A. M. Yaglom, *Einstein's 1914 paper on the theory of irregularly fluctuating sums of observations*, Problems of Information Transmission (Russian) **21**, no. 4(1985), 101–107. English translation: IEEE ASSP Magazine **4**, no. 4 (October 1987).

[You80] R. M. Young, *An Introduction to Nonharmonic Fourier Series*, Academic Press, New York, 1980.

Norbert Wiener References

[27e] N. Wiener, *On the closure of certain assemblages of trigonometric functions*, Proc. Nat. Acad. Sciences **13** (1927), 27–29.

[30a] N. Wiener, *Generalized harmonic analysis*, Acta. Math. **55** (1930), 117–258.

[32a] N. Wiener, *Tauberian theorems*, Ann. of Math. **33** (1932), 1–100.

[33i] N. Wiener, *The Fourier Integral and Certain of its Applications*, Cambridge University Press, 1933.

[34d] R.E.A.C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, Amer. Math. Soc. Colloq. Publ., vol. 19, Amer. Math. Soc., Providence, R.I., 1934.

[39c] N. Wiener and A. Wintner, *On singular distributions*, J. Math. Phys. **17** (1939), 233–246.

[49g] N. Wiener, *Time Series*, The MIT Press, 1970. (1st edition, 1949).

[56e] N. Wiener, *Brain waves and the interferometer*, J. Phys. Soc. Japan **18** (1956), 499–507.

[56g] N. Wiener, *I Am a Mathematician*, The MIT Press, 1956.

[61c] N. Wiener, *Cybernetics*, second edition, The MIT Press, 1961. (1st edition, 1948).

Department of Mathematics, University of Maryland, College Park, Maryland 20742

E-mail address: jjb@math.umd.edu