JOURNAL OF APPROXIMATION THEORY 48, 272–293 (1986)

Spectral Approximation*

ROBERT K. GOODRICH AND KARL E. GUSTAFSON

Department of Mathematics, University of Colorado, Boulder, Colorado 80309, U.S.A.

Communicated by Paul. G. Nevai

Received December 10, 1984

DEDICATED TO THE MEMORY OF GÉZA FREUD

INTRODUCTION

One is familiar with the terms Spectral Resolution and Spectral Decomposition. In this paper we discuss important problems and recent results on (1) Spectral Factorization, (2) Spectral Synthesis, (3) Spectral Decay, and (4) Spectral Estimation.

We have been recently drawn to these subjects principally from considerations in Signal Analysis but they are also important in Geophysics and Stochastic Time Series, and originally caught our attention from questions in Quatum Mechanics and Statistical Mechanics. Beyond a motivating comment or two here and there, we will not go into the applications, as each would require a more extensive treatment. Suffice it to say that, for example, in the multidimensional Signal Analysis, marvelous problems and questions abound.

We have lumped together all four subjects 1–4, to be treated in the following sections of the same number, under the single term *Spectral Approximation.* All four subjects are concerned with approximating a function from limited spectral data. Two further subjects of similar interest in spectral approximation would have been (5) Spectral Sampling and (6) Spectral Optimization.

An excellent survey of recent mathematical problems and results for what we call Spectral Sampling has been given by Butzer [3] with multidimensional problems discussed by Splettstosser [34]. There are some connections between Spectral Estimation and Spectral Sampling, as will be mentioned in Section 4, but the emphases are different, the former principally posed in the frequency domain, the latter usually in the time

* Partially supported by NATO Grant 889/83.

0021-9045/86 $3.00

domain. By Spectral Optimization we mean both optimization techniques needed in Spectral Estimation (see Section 4) and problems of optimal spectral parameters needed in multidimensional numerical relaxation techniques. These will be treated in separate works.

As we have tried to indicate above, the flavor of this paper will be certain mathematical aspects of these problems, especially toward their setting in two and three dimensions, where most have not been resolved.

1. Spectral Factorization

Spectral factorization comes in many guises. Its main use in applications is to generate a unique minimum phase wavelet function in the time domain from a given spectrum in the frequency domain. It is thus an inverse problem in Spectral Approximation. It is quite crucial in feedback filtering theory, the theory of prediction, and seismic prospecting.

Mathematically it corresponds to factoring an arbitrary Hardy space function h (in the frequency domain) as

$$h = io, \tag{1.1}$$

where i is an inner function and o is an outer function. This may be done for H^2 (disc) or H^{2+} upper half plane). A good treatment of the one dimensional case may be found in Dym and McKean [5], where one may find a proof of the factorization in one dimension. The outer factor of an arbitrary element h in H^{2+} is given by the formula

$$o(\lambda) = \exp \frac{1}{\pi i} \int_{-\infty}^{\infty} \frac{s\lambda + 1}{s - \lambda} \frac{\ln|h_0(s)|}{s^2 + 1} \, ds, \tag{1.2}$$

where h_0 denotes the boundary values of h on the real line. The inner factor may be thought of as a type of Blaschke product which absorbs all upper half plane zeros from h. In this way the difficulty in advancing to a theory of Spectral Factorization in two and three dimensions by going via the theory of functions of two and three complex variables is evident. Although some Hardy space theory for the latter exists, we found that an inner–outer factorization theory did not.

Accordingly we have in Goodrich and Gustafson [7] (see also [6, 8, 9]) an inner–outer factorization for two and three dimensions. It relies on a group representation approach; while we may not know enough about functions of several complex variables, we do know the Euclidean two and three space groups and their representations. This approach goes back to the paper [15] and in effect extends Theorem 1 there to higher dimensions. An initial motivation for [7] was to attempt to extend all three

Theorems 1, 2, 3 of [15] to two and three space dimensions. Some discussion of the higher dimensional situation of Theorems 2 and 3 will be found in Section 3.

A word of clarification here: by two and three dimensions we are referring to the domain (sometimes called parameter) space. For a process this is explified by the representation U_v, where $v=(t_1, t_2)$ or $v=(x_1, x_2, x_3)$. Multidimensional ranges are, from our point of view, much easier and have been treated in the stochastic process literature and elsewhere for many years now.

To go beyond [6–9], let us recall:

DEFINITION 1.1. A function $\psi \in \mathscr{L}_2(R^2)$ is an outer function if

$$\overline{\mathrm{sp}}\{{}^{\vee}\psi(v-w) \mid w=(x, y), x \leqslant 0, y \leqslant 0\} = \mathscr{L}_2((-\infty] \times (-\infty, 0]). \tag{1.3}$$

DEFINITION 1.2. A function $\psi \in \mathscr{L}_2(R^2)$ is a weak outer function if

$$\begin{aligned}&\overline{\mathrm{sp}}\{{}^{\vee}\psi(v-w) \mid w=(x, y), x \leqslant 0, -\infty < y < \infty\}\\ &\quad = \mathscr{L}_2((-\infty, 0] \times (-\infty, \infty))\end{aligned}$$

and (1.4)

$$\begin{aligned}&\overline{\mathrm{sp}}\{{}^{\vee}\psi(v-w) \mid w=(x, y), -\infty < x < \infty, y \leqslant 0\}\\ &\quad = \mathscr{L}_2((-\infty\ \infty) \times (-\infty, 0]).\end{aligned}$$

Here the symbol $^{\vee}$ denotes inverse Fourier transform and $^{\wedge}$ will denote the Fourier transform. Definitions 1.1 and 1.2 are motivated by considering the third quadrant in the plane in the role of negative time or as the basis for, e.g., two dimensional picture prediction. Other geometrics could be considered.

Suppose now that

$$\begin{aligned}&\text{(i) } \psi \text{ has support in the third quadrant, and}\\ &\text{(ii) } \overline{\mathrm{sp}}\{U_{(x,y)}\psi \mid (x, y) \in R^2\} = \mathscr{L}_2(R^2),\end{aligned} \tag{1.5}$$

where $U_{(x,y)}\psi$ denotes the regular representation $U_{(x,y)}\psi = \psi(t-x, s-y)$.

DEFINITION 1.3.

$$E_s = \overline{\mathrm{sp}}\{U_{(x,y)}\psi \mid x \leqslant s, y \text{ arbitrary}\}. \tag{1.6}$$

DEFINITION 1.4.

$$F_t = \overline{\mathrm{sp}}\{U_{(x,y)}\psi \mid y \leqslant t, x \text{ arbitrary}\}. \tag{1.7}$$

DEFINITION 1.5.

$$E_{st} = \overline{\mathrm{sp}}\{U_{(x,y)}\psi \mid x \leqslant s,\ y \leqslant t\}. \tag{1.8}$$

Also assume that

$$\bigcap R(E_s) = \{0\} = \bigcap R(F_t). \tag{1.9}$$

Condition (1.9), where R denotes the range of the operator involved, is a necessary condition for an $\mathscr{L}^2$ analysis and may be called "emptiness of the separate infinitely remote pasts." The conditions on ψ in (1.5) may be seen from the one dimensional theory to be natural necessary conditions on ψ: (i) corresponds to the left half line support that the inverse Fourier transform $\psi = {}^{\vee}o$ of an outer function will have, and (ii) corresponds to the needed cyclicity of $o = \hat{\psi}$.

We could not show that just (1.5) and (1.9) were sufficient for $\hat{\psi}$ to be an outer function. By assuming also

$$E_s F_t = E_{st} \tag{1.10}$$

for all (s, t) in two space R^2, we were able to show that there exists an "inner function" g, $|g| = 1$ a.e., such that

$$o = g\hat{\psi} \tag{1.11}$$

is outer. The inner function g is unique except for a scalar multiple of absolute value one. The same result for weaker outer functions was obtained with (1.10) replaced by the weaker assumption

$$E_s F_t = F_t E_s. \tag{1.12}$$

The geometry of the support of ψ plays a vital role. The characteristic function $\chi_{-1010} = \chi([-1, 0] \times [-1, 0])$ of the right unit square adjacent to the origin in the third quadrant is the inverse transform of an outer function. The tilted 45° third quadrant square χ_{45} yields a weak outer function. To work out a full set of such functions would appear to be a formidable task but would yield factorization theorems for all such functions.

We believe that the commutativity assumption (1.11) is redundant in this theory.

LEMMA 1.1. *Given ψ cyclic, that is, satisfying* (1.5)(ii), *if the factorization*

$$\hat{\psi} = i \cdot o \tag{1.13}$$

with $|i| = 1$ a.e. and ${}^{\vee}o$ satisfying (1.4) *holds, then* (1.12) *is automatically satisfied.*

Proof. First we note that $E_s F_t = F_t E_s$ for all s and t or for no s and t. Observe that

$$\begin{aligned} U_{(0,-t_0)} F_{t_0} U_{(0,t_0)} &= F_0 \\ U_{(-s_0,0)} E_{s_0} U_{(s_0,0)} &= E_0 \\ U_{(0,t)} E_0 U_{(0,-t_0)} &= E_0 \\ U_{(s_0,0)} F_0 U_{(-s_0,0)} &= F_0 \end{aligned} \tag{1.14}$$

for all s_0, t_0. Then (1.13) follows easily from the characterization of a self adjoint projection by its range. From

$$\begin{aligned} F_{t_0} E_{s_0} &= U_{(0,t_0)} F_0 U_{(0,-t_0)} U_{(s_0,0)} E_0 U_{(-s_0,0)} \\ &= U_{(0,t_0)} U_{(s_0,0)} F_0 E_0 U_{(0,-t_0)} U_{(-s_0,0)} \end{aligned} \tag{1.15}$$

we see that $F_{t_0} E_{s_0}$ is unitarily equivalent to $F_0 E_0$ for any t_0, s_0.

The proof of the lemma is then completed by noting three unitary equivalences. First, ψ and $\hat{\psi}$ are unitarily equivalent by Fourier transform, multiplication by i^{-1} gives o, and then inverse Fourier transforms maps the projections E_0 and F_0 back to multiplications by the characteristic functions of the left and lower half planes, respectively. The last two multiplication operators commute, hence so do all E_s and F_t.

LEMMA 1.2. *When $\psi(x, y) = f(x)\, g(y)$, where $\hat{f}$ and $\hat{g}$ are one dimensional outer, then $\hat{\psi}$ is outer.*

Proof. In this case f spans completely to the left, g spans completely down, E_s corresponds to multiplication by $\chi((-\infty, s] \times (-\infty, \infty))$, F_t corresponds to multiplication by $\chi((-\infty, \infty) \times (-\infty, t])$, and hence $E_s F_t = F_t E_s$.

Of interest then toward our conjecture are the cases in which $\psi(x, y) = \psi(y, x)$. An example is χ_{45} above, for which χ_{45} is weak outer. One wonders whether all such cyclic symmetric third quadrant support ψ should be weak outer, or in any case what if any additional conditions would be in order.

To that end, one can establish the following facts.

DEFINITION 1.6. $Sf(x, y) = f(y, x)$ for all f in $\mathscr{L}^2(R^2)$.

LEMMA 1.3. (i) *$SE_0S = F_0$; (ii) $F_0E_0 = SE_0F_0S$; (iii) S intertwines E_0 and F_0: $E_0S = SF_0$; (iv) E_0 and F_0 commute iff $(SE_0)^2$ is selfadjoint.*

Proof. (i) Let (x_1, y_1) be given with $y_1 \leqslant 0$. Then S maps $\psi(x - x_1, y - y_1)$ into $\psi(y - x_1, x - y_1) = \psi(x - y_1, y - x_1)$. E_0 maps this function to itself. Then S maps this function to $\psi(y-y_1, x-x_1)=\psi(x-x_1, y-y_1)$. Thus on a dense subset of the range of F_0 we have $F_0=SE_0S$. This then easily implies $SE_0S=F_0$. Part (ii) follows from (i). (iii) $S^{-1}==S=S$. (iv) The relation is obvious. We note an interesting operator theoretic twist here. SE_0 is not self adjoint, not even normal, yet is a square root of a self adjoint operator whenever E_0 and F_0 commute.

Let $\mathscr{S}=\{f \mid Sf=f\}$, i.e., $\mathscr{S}$ is the vector space of all symmetric functions. Also let $\mathscr{AS}=\{f \mid Sf=-f\}$, i.e., $\mathscr{AS}$ is the vector space of all anti-symmetric functions.

LEMMA 1.4. (i) $S(E_0F_0+F_0E_0)\,S=E_0F_0+F_0E_0$.

(ii) $S(E_0F_0-F_0E_0)=-E_0F_0+F_0E_0$.

(iii) $E_0F_0+F_0E_0$ *maps* $\mathscr{S}$ *into* $\mathscr{S}$ *and* $\mathscr{AS}$ *into* $\mathscr{AS}$.

(iv) $E_0F_0-F_0E_0$ *maps* $\mathscr{S}$ *into* $\mathscr{AS}$ *and* $\mathscr{AS}$ *into* $\mathscr{S}$.

(v) $E_0F_0=F_0E_0$ *if and only if* E_0F_0 *maps* $\mathscr{S}$ *into* $\mathscr{S}$ *and* $\mathscr{AS}$ *into* $\mathscr{AS}$.

Proof. We see that (i) and (ii) are immediate consequences of (ii) in Lemma 3. Then (iii) and (iv) follow from (i) and (ii) and the fact that every $\mathscr{L}_2$ function can be uniquely decomposed into the sum of a symmetric and anti-symmetric function.

To prove (v) note that if E_0F_0 maps $\mathscr{S}$ to $\mathscr{S}$ and $\mathscr{AS}$ to $\mathscr{AS}$ then from (ii) of Lemmas 3 E_0F_0 and F_0E_0 agree on $\mathscr{S}$ and $\mathscr{AS}$ and hence agree on $\mathscr{L}_2$.

Finally, if $E_0F_0=F_0E_0$ then $E_0F_0+F_0E_0=2E_0F_0$ and E_0F_0 must map $\mathscr{S}$ to $\mathscr{S}$ and $\mathscr{AS}$ to $\mathscr{AS}$ by (iii) above.

In [32, Chap. 19] Rudin discusses the conjecture that inner functions do not exist on the unit ball for more than one dimension, and draws a number of conclusions depending on the truth or falseness of the conjecture. As Rudin point out [32, Preface]:

> The fact that they [such open problems] are still unsolved shows quite clearly that we have barely begun to understand what really goes on in this area of analysis....

It turns out that a number of investigations shortly thereafter found the existence of *inner functions*: f of norm $\|f\|_\infty=1$ such that

$$|f^*|=1 \text{ almost everywhere, } f^* \text{ the radial limit.} \tag{1.16}$$

See, for example, Løw [25] for a discussion of those results.

These higher dimensional inner functions are extremely oscillatory near the boundary, and it was known that there are $\mathscr{H}^\infty$ functions for which inner–outer factorization fails on each higher dimensional ball. From attempts to get some sort of factorization came the notion of *internal functions*: f of norm $\|f\|_\infty = 1$ such that

$$h, h^{-1} \text{ bounded analytic, } |f| \leqslant |h| \leqslant 1 \text{ everywhere} \Rightarrow h \text{ is constant.} \tag{1.17}$$

The advantages of this notion is that no radial limits enter into the definition, inner functions being known to have extremely oscillatory behavior near the boundary. By also redefining the notion of outer to external, factorization results have been obtained; see Rubel [28].

There is also an earlier version of such functions, called *interior functions*; see Rubel and Shields [29], for which factorization in higher dimensions also failed. An interior function f was

$$f \text{ in } \mathscr{H}^\infty, \mathscr{I}(f) \text{ is } w^* \text{ closed.} \tag{1.18}$$

Here we have used $\mathscr{I}(f)$ to denote the principal ideal generated by f.

Thus there are now four notions of "inner function" in higher dimensions, along with some factorization result, pro or con, for each. For lack of a better name we will continue to call our version "inner functions" with the understanding that our main focus is factorized rather than some other analytic property.

Because our approach is functional analytic whereas that for internal functions is complex function analytic, it would be interesting to compare the two theories and the properties of the inner and outer functions of each.

2. Spectral Synthesis

The celebrated closure theorem of N. Wiener [38] states that the nonvanishing of the Fourier transform of an integrable function ϕ is necessary and sufficient for the translates of ϕ to span

$$\mathscr{L}^1(R^1) = \overline{\mathrm{sp}}\{\phi(x-y) \mid y \in R^1\}. \tag{2.1}$$

More generally R^1 can be replaced by any nondiscrete locally compact abelian group G with integration taken with respect to a Haar measure on the group. Then the necessary and sufficient condition for denseness of the linear combinations of the translates of ϕ is that $\hat{\phi}$ does not vanish on the character group $\hat{G}$ of G.

If ϕ is in $\mathscr{L}^\infty(G)$ then in the w^* topology of $\mathscr{L}^\infty(G)$, that is, the $\mathscr{L}^1$ topology on $\mathscr{L}^\infty$, the closure of the span of translates of ϕ contains at least one character. Such characters may be called the spectrum of ϕ. The problem of spectral synthesis as originally formulated was to consider any nonvanishing bounded measurable function ϕ and determine if its span contains enough characters such that their span contains ϕ itself:

$$\phi \in \overline{\mathrm{sp}}\{\chi \in \hat{G},\ \chi \in \overline{\mathrm{sp}}\{\phi_y(x) \mid y \in G\}\}. \tag{2.2}$$

From this problem there evolved a number of results and a rather extensive literature. See Hewitt and Ross [17] (Chapter 10 there is devoted entirely to questions of spectral synthesis), and see also Graham and McGehee [13].

Within this general framework, and as one of the early papers on the subject, Beurling [1, 2] established *harmonic spectral synthesis* for weighted $\mathscr{L}_1$ spaces: for every $f \in \mathscr{L}_\infty(R)$ there exists a trigonometric polynomial $p(x) = \sum_{k=1}^{n} \alpha_k e^{it_k x}$, with all t_k in the spectrum of f, such that $\|(f-p)\, w\|_1$ is arbitrarily small. This result holds for any nonnegative even nonincreasing weight $w \in \mathscr{L}_1(R)$. We quote Hewitt and Ross [17]:

> The papers Beurling [1] and [2] remain something of a mystery, and a thorough exegesis of their ideas for general locally compact Abelian groups would be most welcome....

Our approach to Spectral Factorization may thus be viewed as harmonic spectral synthesis for weighted $\mathscr{L}_2$ spaces on a semigroup S or just on a geometric substructure S of a locally compact Abelian group: for every $f \in \mathscr{L}_2(S)$ when does there exist a weighted trigonometric polynomial $p(s) = \sum_{k=1}^{n} \alpha_k e^{i(r_k, s)}\hat{\phi}(s)$, with all r_k in S, such that $\|f-p\|_2$ is arbitrarily small? The notion of spectrum here is left imprecise due to the fact that $\mathscr{L}_2(S)$ is not generally an algebra and also because the set S may no longer be a group. That is, we just allow r access to all values in the set S. So far in our investigations we have only considered for specific S, quadrants, octants, or hyperspaces, e.g., semigroups.

Because the closure of the subspace generated by all translates of a set of one or more functions $\phi_1, ..., \phi_n$ in $\mathscr{L}_1(R)$ is the same as the closed ideal generated by that set of functions, the problem of spectral synthesis may in that case be cast in terms of ideals: when is a closed ideal the intersection of the regular maximal ideals which contain it? From this point of view, $\mathscr{C}(X)$ is therefore seen to be the most natural (easiest) setting for resolution of spectral synthesis questions, by use of the Gelfand–Naimark–Segal construction and the correspondence of closed ideals to closed subsets of X. Much of the abstract work on spectral synthesis seems influenced by this point of view. Going to $\mathscr{L}_1(G)$, one still has an algebra and ideals, but the question of spectral synthesis becomes more difficult and the understanding of it did not really begin until L. Schwartz gave his counter-example [33] for $\mathscr{L}^1(R^3)$. We will describe this example in another context in the next section.

In going to $\mathscr{L}_2(S)$ one loses algebraic structure in both the functional and underlying spaces. But one gains the inherent dualities of $\mathscr{L}_2$ spaces and the general strength and applicability of least-squares approximation methods.

Consider, for example, the above spectral synthesis problems modified by replacing the set of all translates of an $\mathscr{L}^1$ function by the set of translates $\{\phi_y \mid y \in SG\}$, where SG is a semi-group in the group. If we try to apply the theory of commutative Banach algebras to this problem, we immediately run into difficulties. To illustrate these difficulties, let us take the group to be R, the real numbers, and let SG be the set of all non-negative numbers. In this case and in the problems above, we are most interested in understanding which elements of $\mathscr{L}^1$ are in $\overline{\mathrm{sp}}\{\phi_y \mid y \geqslant 0\}$ The spectral synthesis problems in $\mathscr{L}^\infty$ then can be viewed as a dual problem similar to the one under consideration above.

LEMMA 2.1.

$$\overline{\mathrm{sp}}\{\phi_y \mid y \geqslant 0\} = \overline{\mathrm{sp}}\left\{f \mid f(x) = \int_0^\infty \phi(x-y)\, h(y)\, dy \text{ for some } h \in \mathscr{L}^1\right\}.$$

Proof. Let $V^+ = \{f \mid f(x) = \int_0^\infty \phi(x-y)\, h(y)\, dy$ for some $h \in \mathscr{L}^1\}$. We show a continuous linear functional F on $\mathscr{L}^1$ vanishes on V^+ if and only if F vanishes on $V = \{\phi_y \mid y \geqslant 0\}$. Then by the Hahn–Banach theorem this implies $\overline{\mathrm{sp}}\ V = \overline{\mathrm{sp}}\ V^+$.

Let $F(f) = \int_{-\infty}^\infty f(x)\, q(x)\, dx$ for some q in $\mathscr{L}^\infty$. If $f \in V^+$ then $F(f) = \int_0^\infty h(y)(\int_{-\infty}^\infty \phi(x-y)\, q(x)\, dx)\, dy$. If F vanishes on V^+ then $0 = \int_0^\infty h(y) \int_{-\infty}^\infty \phi(x-y)\, q(x)\, dx$ for all h in $\mathscr{L}^1$. Thus $\int_{-\infty}^\infty \phi(x-y)\, q(x)\, dx = 0$ almost everywhere on $(0, \infty)$. But this integral is a continuous function in y since q is in $\mathscr{L}^\infty$. So $\int_{-\infty}^\infty \phi(x-y)\, q(x)\, dx = 0$ for all y in $(0, \infty)$. Thus F vanishes on V. Conversely, if F vanishes on V one can reverse the above steps to show F vanishes on V^+.

Another way of stating this result is that

$$\overline{\mathrm{sp}}\{\phi_y \mid y \geqslant 0\} = \overline{\mathrm{sp}}\{\phi * h \mid h \in \mathscr{L}^1 \text{ and the support of } h \subseteq (0, \infty)\}.$$

It is very easy to see $\overline{\mathrm{sp}}\{\phi_y \mid y \in R\} = \overline{\mathrm{sp}}\{\phi * h \mid h \in \mathscr{L}^1\}$. This last set is an ideal in $\mathscr{L}^1$ under the convolution product. Thus if we are considering *all* translates of ϕ we may apply the theory of commutative Banach algebras to obtain information about this ideal.

In the case where we are only considering $\overline{\mathrm{sp}}\{\phi_y \mid y \geqslant 0\}$, it will not in general be an ideal. This is so because $f * h$ may not have support in $[0, \infty)$ when h has support in $[0, \infty)$ and no restriction is made on the support of f. It would seem that restricting the set of translates to a semi-group, that is, trying to solve the problem with less spectral data, greatly reduces the possibility of algebraic (Banach algebra) techniques for the problem. However, in the special case where one also knows that the support of ϕ is contained in $[0, \infty)$ then the set $\overline{\text{sp}}\{\phi * h \mid h \in \mathscr{L}^1[0, \infty)\}$ forms an ideal in $\mathscr{L}^1[0, \infty)$, where the convolution product is defined in the usual way, i.e., $\phi * h(x) = \int_{-\infty}^{\infty} \phi(x-y)\, h(y)\, dy$, remembering that ϕ and h have supports in $[0, \infty)$. The above set is an ideal because if h has compact support then support $(f * h) \subseteq$ support (f) + support $(h) \subseteq [0, \infty)$ because $[0, \infty)$ is a semi-group (see, e.g., Hormander [19, Theorem 1.6]). Also, if f and h are in $\mathscr{L}^1[0, \infty)$ then if we pick a sequence $\{h_n\}$ in $\mathscr{L}^1$ with each h_n having compact support in $(0, \infty)$ and $\{h_n\} \to h$ in $\mathscr{L}^1$ then $\{f * h_n\} \to f * h$ in $\mathscr{L}^1$ and so the support of $f * h$ is contained in $[0, \infty)$ and the set is an ideal in $\mathscr{L}^1[0, \infty)$.

The maximal ideal space of the Banach algebra $\mathscr{L}^1(0, \infty)$ is computed to be the functionals $\alpha(f) = \int_0^\infty f(s)\, \bar{\chi}(s)\, ds$, where $\chi(s_1 + s_2) = \chi(s_1) \cdot \chi(s_2)$ for s_1 and $s_2 \geq 0$, and χ is in $\mathscr{L}^\infty$ (see, e.g., Loomis [24, p. 73]). Because $\chi(s) = e^{as}$ for some complex a, in order that χ be in $\mathscr{L}^\infty[0, \infty)$ we must have $\operatorname{Re} a \leq 0$. Thus the maximal ideal space of the algebra is the left half complex plane including the imaginary axis.

$\mathscr{L}^1(0, \infty)$ is reminiscent of Beurling's algebra (see [24, p. 180] for a discussion). The above argument holds out the possibility of studying the translates $\{\phi_y \mid y \geq 0\}$ of an $\mathscr{L}^1$ function whose support is contained in the right half line by using the theory of ideals for the algebra $\mathscr{L}^1[0, \infty)$.

In higher dimensions we would then consider translates of a function $\{\phi_y \mid y \in SG\}$ for a function with the support of ϕ contained in a semi-group SG. Motivated by both the above discussion (and, incidentally, the theory of picture processing) one would study this problem with the semi-group of half-planes and quadrants. One could also consider the dual problem of considering the translates $\{\phi_y \mid y \in SG\}$ of an $\mathscr{L}^\infty$ function ϕ whose Fourier transform has support in SG. One would hope these analogies to the known literature of spectral synthesis would yield insight into spectral synthesis problems and Tauberian theorems on semi-groups.

3. Spectral Decay

As mentioned in Section 1 above, the extension of the following two theorems (Theorems 2 and 3, respectively of Gustafson and Misra [15]) to higher (e.g., two and three) dimensions should be considered.

THEOREM. (SSKKKPW). *A stationary process $x(t)$ is regular if and only if its spectral density $f(\lambda)$ satisfies the condition*

$$\int_{-\infty}^{\infty} \frac{\ln f(\lambda)}{1+\lambda^2}\, d\lambda > -\infty \tag{3.1}$$

THEOREM (Unstable Particle). *A unitary evolution U_t is that of a regular stationary process if and only if it admits a one dimensional decaying subspace without regeneration.*

A word about the theorems. Theorem (SSKKKPW) is rather famous and is found variously under the names Szegö, Smirov, Kolmolgorov, Krein, Krylov, and Paley, Wiener. Theorem 1 of [15] provided a new short proof. Theorem (Unstable Particle) grew out of mathematical questions about models for meson decay. Both are related to the condition (3.1), which for efficiency we will just call the Szegö condition (see Sezgö [36]).

The Szegö condition for the unit ball becomes

$$\int_{-\pi}^{\pi} \ln f(\theta) > -\infty. \tag{3.2}$$

A positive $\mathscr{L}_2(-\pi, \pi)$ function $f(\theta)$ is the absolute value of the boundary values of an H^2 function if and only if (3.2) is satisfied. The condition (3.2) may also be seen to be equivalent to the existence of an $\mathscr{L}^2(-\pi, \pi)$ function $f(e^{i\theta})$ with $|f(e^{i\theta})| = f(\theta)$ and with all negative Fourier coefficients vanishing. In signal filtering theory $f(\theta)$ is often taken as a prescribed gain $|B(\theta)|$ and the vanishing negative Fourier coefficients signify causality.

We do not have two and three dimensional versions of the two theorems above and we are not aware of any. Both would depend on a higher dimensional Szegö condition (3.1). Any such condition would help in understanding higher dimensional outer functions.

There was a great amount of work done in the thirties and fourties on decay rates of a function f and its Fourier transform $\hat{f}$. Weiner made the important remark (see Hardy [16]):

> a pair of transforms f and g cannot both be very small.

In other words, f and $g = \hat{f}$ cannot both be very small at infinity. Hardy [16] responded with

THEOREM (Hardy and Weiner). *If*

$$f(t) = 0(|t|^m e^{-\alpha t^2}) \qquad \text{at} \quad |t| \to \infty \tag{3.3}$$

and

$$\hat{f}(\lambda) = 0(|\lambda|^m e^{-\beta\lambda^2}) \qquad \text{as} \quad |\lambda| \to \infty, \tag{3.4}$$

m a nonnegative integer, α and β positive and satisfying

$$\alpha\beta \geqslant \tfrac{1}{4}, \tag{3.5}$$

then $\hat{f}(\lambda) = p_m(\lambda)\, e^{-\beta\lambda^2}$ *where* $p_m(\lambda)$ *is a polynomial of degree less than or equal to m.*

In particular, if $\alpha = \beta = \frac{1}{2}$, then both f and $\hat{f}$ have the same form, and when $m = 0$, both are constant multiples of the Gaussian $e^{t^2/2}$

We now wish to move from this Theorem in two ways. First, we want to consider smallness in the sense of the Szegö condition. Secondly we want to discuss the importance of "look-alike" f and $\hat{f}$.

Our interest in Spectral Decay in the sense of Szegö came not only from [15] but from a conjecture concerning irreversibility and K-flows in Statistical Mechanics (see Goodrich, Gustafson, and Misra [11]).

Specifically, the conjecture was that the following could not hold for a function $f \in \mathscr{L}^2(-\infty, \infty)$:

$$\begin{aligned} &\text{(i) } f(t) \text{ real and nonnegative,} \\ &\text{(ii) } \int_{-\infty}^{\infty} \frac{\ln f(t)}{1+t^2}\, dt > -\infty, \\ &\text{(iii) } \hat{f}(\lambda) \text{ real and even,} \\ &\text{(iv) } \int_{-\infty}^{\infty} \frac{\ln|\hat{f}(\lambda)|}{1+\lambda^2}\, d\lambda > -\infty. \end{aligned} \tag{3.6}$$

Note that f real $\Rightarrow \hat{f}$ even and $\hat{f}$ real $\Rightarrow f$ even but this additional information is only incidental to the question: can a spectral density f and its transform $|\hat{f}|$ both satisfy the Szegö condition? Although at first our intuition favored the conjecture (viz. the remark of Wiener), it is false.

LEMMA 3.1. *The smallness conditions on f and* $\hat{f}$ *of* (3.6) *can be satisfied by "look-alike" f and* $\hat{f}$.

Proof. The idea is to ignore phase. Let

$$g(t) = (1+t^2)^{-1}$$

and

$$f(t) = g(t) + e^{-|t|}. \tag{3.7}$$

Then because

$$f(t) = g(t) + \hat{g}(t) \tag{3.8}$$

we have $\ln f(t) > \ln g(t)$ and

$$\int_{-\infty}^{\infty} \frac{\ln f(t)}{1+t^2} > \int_{-\infty}^{\infty} \frac{\ln g(t)}{1+t^2} > -\infty. \tag{3.9}$$

Also note that f is L^2 and $\hat{f} = \hat{g} + {}^{\vee}\hat{g} = f$ so that (iii) and (iv) are also satisfied.

COROLLARY 3.2. *If f is a Szegö spectral derivative, i.e., satisfies* (3.1), *then*

$$\begin{aligned} &\text{(i) so does } f + g \text{ for any nonnegative } g \text{, and} \\ &\text{(ii) so does } f^p \text{ for any } 0 < p < \infty. \end{aligned} \tag{3.10}$$

Thus from any Szegö spectral density there are many more. Correspondingly, one regular stationary stochastic process $x(t)$ generates a large related family of such processes.

From Lemma 3.1 one is led to ask: does the failure of a spectral density f to satisfy the Szëgo condition imply that its transform $\hat{f}$ must satisfy it (viz. the Wiener remark)? Again "look-alike" f and $\hat{f}$ give a counterexample.

LEMMA 3.3. *Szegö largeness conditions on f and $\hat{f}$, i.e., simultaneous diverge in* (3.6)(ii) *and* (iv), *can be satisfied.*

Proof. Take $f = e^{-t^2}$; then $\hat{f} = (\text{const})\, e^{-\lambda^2/4}$.

Thus the Szegö condition governs Spectral Decay not just at infinity but everywhere on the real axis: $f(t)$ cannot be too small. Thinking of $f(t)$ as the spectral density, i.e., in the case when it is the derivative of a distribution function for a stochastic process or the derivative $d(E(\lambda)\,\phi, \phi)$ of a spectral family, Szegö Spectral Decay is that of a lower bound rather than upper bound on decay, not only at infinity but everywhere. This lower bound on smallness at infinity (and everywhere) can be simultaneously satisfied by a function f and $\hat{f}$.

As a second instance of the importance of look-alike f and $\hat{f}$, we recall the Uncertainty Principle in signal theory: f and $\hat{f}$ cannot both be of short duration. Specifically, in one dimension, if $f(t) = O(t^{-1/2})$ at $|t| \to \infty$ and if α, β, where

$$\begin{aligned} \alpha^2 &= \int_{-\infty}^{\infty} t^2 |f(t)|^2 dt \Big/ \int_{-\infty}^{\infty} |f(t)|^2 dt, \\ \beta^2 &= \int_{-\infty}^{\infty} \lambda^2 |\hat{f}(\lambda)|^2 d\lambda \Big/ \int_{-\infty}^{\infty} |\hat{f}(\lambda)|^2 d\lambda, \end{aligned} \tag{3.11}$$

measure the durations of f, then the Uncertainty Principle states that

$$\alpha\beta \geqslant \tfrac{1}{2}. \tag{3.12}$$

Sharpness of this estimate occurs only at Gaussians $f(t) = e^{-\gamma t^2}$.

The relation of the Uncertainty Principle to the Hardy–Wiener theorem is clear. Only the weight functions differ. Sharpness comes at a "look-alike" function. Probably there is a similar uncertainty principle for the Szëgo functional (weight $(1+t^2)^{-1}$, $\ln f$). Other measures of duration are used in signal processing time–bandwidth considerations, yielding other f, $\hat{f}$ spectral growth and decay limitations.

A third use of "look-alike" f and $\hat{f}$ may be found in the counterexample of L. Schwartz [33] to Spectral Synthesis in $\mathscr{L}_1$. Let S_2 denote the unit sphere in three-space, let $\mathscr{I}$ denote the closed ideal of functions f in $\mathscr{L}_1(R^3)$ such that $\hat{f}(S_2)=0$, let $\mathscr{I}_0$ denote functions in $\mathscr{I}$ for which also $\partial\hat{f}/\partial y_1(S_2)=0$, and let $\overline{\mathscr{I}_0}$ denote the closure of $\mathscr{I}_0$. Then $\overline{\mathscr{I}_0}$ is a translation invariant subspace of $\mathscr{L}_1(R^3)$ for which spectral synthesis fails. This is demonstrated by exhibiting a bounded linear functional which separates $\overline{\mathscr{I}_0}$ and $\mathscr{I}$. Letting

$$x'f=\int_{S_2}(\partial\hat{f}/\partial y_1)\,d\sigma \tag{3.13}$$

it can be checked that $x'(\overline{\mathscr{I}_0})=0$, $x'(f)\neq 0$, where

$$f(x)=2^{3/2}e^{-|x|^2}-e^{(1/4-|x|^2/2)}. \tag{3.14}$$

Note that

$$\hat{f}(\lambda)=e^{-|\lambda^2|/4}-e^{(1/4-|\lambda|^2/2)},$$

that is, f and $\hat{f}$ are "look-alike" functions. Duality in f, $\hat{f}$ serves to measure nonduality in $\mathscr{L}_1$. We formalize this observation as follows.

PROPOSITION 3.4. *"Look-alike" functions f and $\hat{f}$ serve as limits in Spectral Decay statements.*

It would seem (perhaps this has been done somewhere) that a study of the functional dependence

$$\hat{f}(\lambda)=af(b\lambda+c) \tag{3.15}$$

of transforms on important groups such as R^n and C^n would be useful. One could allow a, b, c the two classes of being constant or being functions of λ. Group effects are already indicated in the R^1 scaling law

$$af(at)\leftrightarrow\hat{f}(\lambda/a). \tag{3.16}$$

Some further known examples of "look-alike" functions in one dimension are

$$
\begin{aligned}
&f = |t|^{-1/2}, && \hat{f} = (2\pi)^{1/2}|\lambda|^{-1/2}\\
&f = (\operatorname{sgn} t)|t|^{-1/2}, && \hat{f} = i(\operatorname{sgn}\lambda)(2\pi)^{1/2}\lambda^{-1/2};\\
&f = e^{-zt^2}, && \hat{f} = (\pi z^{-1})^{1/2}e^{-\lambda^2/4z}
\end{aligned}
\tag{3.17}
$$

$$\text{where } z = a + ib,\ z > 0,\ \operatorname{Re}\sqrt{z} > 0;$$

$$f = \cos(at^2), \qquad \hat{f} = (\pi a^{-1})^{1/2}\cos(\lambda^2/4a - \pi/4);$$

$$f = e^{-t^2/2} Her_n(2^{1/2}t), \qquad \hat{f} = (2\pi)^{1/2}(i)^n e^{-\lambda^2/2} Her_n(2^{1/2}\lambda).$$

As mentioned above, all such functions should be characterized.

4. Spectral Estimation

In many applications of Spectral Approximation (by our definition, the extracting of some approximation or other information from limited spectral data), one does not even know the exact spectra data itself. That is, one must first estimate it. We will take this as a general description of the problem of Spectral Estimation.

An important instance of this is power spectrum estimation. Given some incomplete time series data $x(t_n)$, what frequencies are present in it and more importantly what frequencies were present in the complete original source $x(t)$? If we approximate $x(t)$ by $f(t)$ in the time domain, then $\hat{f}(\lambda)$ in the frequency domain would hopefully indicate those frequencies present. For this reason the amplitude $|\hat{f}(\lambda)|$ is called the power spectrum. A rather periodic incoming signal will show up in the frequency domain as a power spectrum concentrated on just those participant incoming frequencies and dropping to zero beyond them, whereas a turbulent incoming signal will transform to a continuous spectrum extending into high frequencies.

A main goal of this last section is to establish basic and intimate connections between Spectral Estimation and the other three aspects of Spectral Approximation brought forth in this paper. Indeed let us go even further and first tie in, albeit briefly, the other two aspects of Spectral Approximation that we mentioned in the Introduction.

The Sampling Theorem of Spectral Sampling asserts that if the incoming signal $f(t)$ is band-limited, namely,

$$|\hat{f}(\lambda)| = 0 \qquad \text{for} \quad |\lambda| \geqslant \sigma, \tag{4.1}$$

then $f(t)$ can be reconstructed exactly from its values sampled at the times $\pm n\pi/\sigma$ by the representation

$$f(t) = \sum_{n=-\infty}^{\infty} \frac{\sin \sigma(t - n\pi/\sigma)}{\sigma(t - n\pi/\sigma)} f\left(\frac{n\pi}{\sigma}\right). \tag{4.2}$$

Note that to so reconstruct $f(t)$ you need to sample over an infinite time duration. In practice you have sampling only over a finite time interval $[-T, T]$. From the Spectral Decay limitations of Section 3 you have no right to limit the data to this interval, transform it, and then expect $\hat{f}(\lambda)$ to represent the power spectrum. Not only will power extend beyond the band $|\lambda| \geqslant \sigma$ but also the effect of this dispersion on the spectrum within the band is a priori unclear. This is the basis of the problem of the estimation of power spectrum. The difficulties are compounded by statistical uncertainties in the sampled data. To effectively remove the latter would require averaging over a great many realizations whereas in practice you have only one (or a few) realizations $x(t)$. From this has grown a large literature on spectral estimators. For example, for the discrete case the exact representation

$$\hat{f}(\lambda) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{-i\lambda n} C(n) \tag{4.3}$$

of $\hat{f}(\lambda)$ in terms of the covariances $C(n)$ yields the estimate

$$\hat{f}_e(\lambda) = \frac{1}{2\pi} \sum_{-\infty}^{\infty} e^{-i\lambda n} w(n)\, C_e(n) \tag{4.4}$$

where $w(n)$ is a so-called lag window vanishing for all $n \geqslant$ some N, and where $C_e(n)$ is a maximal likelihood (or other) estimate of the unknown true Fourier coefficients $C(n)$. This type of estimate goes back to Tukey [37]. The problem is that no matter how large N is, even though $\hat{f}_e(\lambda)$ converges in the mean to $f(\lambda)$, $\hat{f}_e(\lambda)$ has a large variance, greater than, for example, the square of the expected value of $\hat{f}_e(\lambda)$. To reduce this variance one must smooth the power spectrum estimator, thereby giving up resolution for reliability. Such smoothings will be limited by the uncertainty principles of Section 3.

The uncertainty principle (see Section 3, Spectral Decay) says that if $x(t)$ contains most of its energy in $(-T, T)$, then its transform $f(\lambda)$ contains most of its energy in the band $|\lambda| > \text{const}/T$. If we assume that in fact $x(t) = 0$ outside $(-T, T)$, then indeed the best frequency resolution will be const/T. But if instead we assume $x(t)$ continues or may be continued outside $(-T, T)$ in a good way, then we can increase the frequency resolution. Therefore we would like to extrapolate $x(t)$ to the left and right of its sampling interval in a way which maximizes the frequency resolution. A good way to do this that has been developed by Levinson, Burg, and others (e.g., see Lang and McClellan [21, 22]) is: in place of estimating the autocorrelations r_k directly from the sampled data $x(t_n)$, instead estimate a minimum phase prediction error filter directly from the data. This corresponds to a Spectral Factorization.

Recall the Spectral Estimation problem: estimate the power spectrum $f(\lambda)$ of a given stationary random process $x(t)$. Here $f(\lambda)=\hat{r}(t)$, where $r(t)$ is the autocorrelation function

$$r(t)=x(t)*\overline{x(-t)}=\int_{-\infty}^{\infty} x(t+s)\,\overline{x(s)}\,ds. \tag{4.5}$$

If we assume that a realization $x(t)$ is available only over an interval $(-T, T)$, a common estimator of the autocorrelation is

$$r_T(t)=\frac{1}{2T-|t|}\int_{-T+|t|/2}^{T-|t|/2} x(s+t/2)\,x(s-t/2)\,ds. \tag{4.6}$$

This estimator is symmetric, unbiased, and converges to $r(t)$ as $T\to\infty$. But its Fourier transform $\hat{r}(t)$, as a would-be approximation to the power spectrum $f(\lambda)$, converges to a random variable whose variance equals $f^2(\lambda)$. There is a long history of this problem, starting with the above (essentially, the so-called periodogram) approach, evolving through an introduction of a number of windowing and Fast Fourier Transform techniques, and now employing a number of smoothed estimators of $r(t)$. The latter idea is based on smoothing $r(t)$ so that fewer frequencies are present, which yields a sharper power spectrum. An excellent reference for more information is Koopmans [20].

One can now relate Spectral Estimation and Spectral Factorization as follows.

PROPOSITION 4.1. *Spectral Factorization gives Spectral Estimation in one dimension.*

Proof. Spectral Factorization can be formulated as follows: given a function $r(t)$ whose transform $f(\lambda)$ is nonnegative, find a causal function $x(t)$ such that

$$r(t)=x(t)*\overline{x(-t)}=\int_{-\infty}^{\infty} x(t+s)\,\overline{x(s)}\,ds. \tag{4.7}$$

The comparison of (4.5) with (4.7) above is intentional. Here causal is taken to the right: $x(t)=0$ for all $t\leqslant 0$. If moreover $r(t)$ is of limited time duration, e.g., if it is from a finite sampling interval $(-T, T)$, then factorization also can be obtained with $x(t)=0$ for $t \geqslant T$. This argument uses the Fejer–Riesz and Akhiezer–Krein moment theorems; see Papoulis [26]. The connection to Spectral Factorization as we have formulated it in this paper follows from the fact that $x(t)$ will be outer and hence of minimum phase.

In practice one usually will have available from measurement only a discrete set of correlation (also called autocorrelation; note that they are convolutions) values r_k, $k=0, \pm,..., \pm m$. Because we want to also discuss the multidimensional case here, let the r_k values be known for k in $\Delta = \{0, \pm\delta_1,..., \pm\delta_m\}$, i.e., Δ is a symmetric set of vectors about the origin in R^n. We also assume that the power spectrum $f(\lambda)$ has compact support K, and that $f(\lambda) \geqslant 0$.

Let us make a connection to Spectral Synthesis now before we go on. In Spectral Synthesis if one knows all translates of a function then under ideal circumstances one can reconstruct the function from the characters contained in that span. Here we do not have the whole span but we do have some convolutions, namely they r_k, and from this we want a best possible reconstruction of the function.

Remembering that

$$r_\delta = \int_K f(\lambda)\, e^{-jk\cdot\delta} dk \tag{4.8}$$

for all δ in the sampling set Δ, one way in practice to estimate the unknown $x(t)$ is to settle just for an estimation of its power spectrum $f(\lambda)$. If this is successful, then one at least can determine those frequencies present in the signal $x(t)$. Spectral Estimation in practice concentrates on this latter partial problem.

One important method is known as the maximum entropy method. In this method one finds an approximate power spectrum according to the criteria

$$\max_{f \geqslant 0} \int_K \ln f(\lambda)\, d\lambda \tag{4.9}$$

where f satisfies the constraints (4.8). In one dimension the solution takes form

$$f(\lambda) = 1/p(\lambda), \tag{4.10}$$

where $p(\lambda)$ is a positive trigonometric polynomial.

Optimization theories clearly become important in such methods. The choices of objective functional, algorithm, and solution properties are somewhat unclear in higher dimensions and are investigated in Goodrich and Gustafson [10].

The question as to when a finite sequence of convolution vectors $\{r_\delta \mid \delta \in \Delta\}$ can be represented as $r_\delta = \int_K e^{i\delta \cdot k} f(k)\, dk$ for some $f \geqslant 0$ has itself an interesting history. If we consider instead the more general question as to when $r_\delta = \int e^{i\delta \cdot k} d\mu(k)$ for some measure $\mu \geqslant 0$, we find that we have made connection to an important classical extension problem, in one dimension, called the trigonometric extension problem [4, 14, 21, 30, 35].

For one dimension, $\Delta = \{0, \pm 1, ..., \pm m\}$, $K = [-\pi, \pi]$, the solution, as is well known, is that the corresponding correlation matrix R

$$R = \begin{bmatrix} r_0 & \cdots & r_m \\ r_{-1} r_0 & \cdots & r_{m-1} \\ r_{-m} & \cdots & r_0 \end{bmatrix}, \qquad r_k = \bar{r}_{-k}, \tag{4.11}$$

be positive definite, i.e., the function r_δ is a positive definite function on Δ.

In higher dimensions the positive definiteness of r on Δ is not sufficient to imply the existence of a measure μ. This was shown by Rudin [30], and also by Calderon and Papinsky [4]. The reason for this rests in the fact that in higher dimensions not every positive polynomial is the sum of the squares of polynomials. The crucial fact about polynomials was established by Hilbert [18], and was used by Rudin, and by Calderon and Pepinsky to establish their results. The latter imply the existence of a positive definite function with no extension, i.e., no $\mu \geqslant 0$ exists such that $r_\delta = \int_K e^{i\delta \cdot k} d\mu(x)$ for $\delta \in \Delta$, for certain Δ and K. Recently, Lang [23] has constructed an explicit example of a positive definite r that has no extension. This example in turn makes strong use of polynomial constructed by Robinson [27] which is not the sum of squares of polynomials. The existence question does not arise in one dimension since in one dimension every positive mth degree trigonometric polynomial can be factored as the square magnitude of an mth degree trigonometric polynomial, according to the Fejer–Riesz theorem.

Thus it is necessary to give conditions for extendibility for higher dimensional r. These are known; first we need some definitions. Note that if r is extendible then $r(\delta) = \int_K e^{jk \cdot \delta} d\mu(k)$ for all $\delta \in \Delta$ and some measure μ. Then $r(\delta)$ has $2m+1$ components and $r(-\delta) = \bar{r}(\delta)$. We think of $(r(\delta))$ as a $2m+1$ vector in R^{2m+1}. Let p be any vector in R^{2m+1}, then we associate a Δ-polynomial $P(k) = \sum_{\delta \in \Delta} p(\delta)\, e^{-jk\delta}$, where $p(-\delta) = \bar{p}(\delta)$. Such a vector p is called positive if $P \geqslant 0$ on K. A scalar product is defined between all correlation vectors r and polynomials p by

$$(r, p) = \sum_{\delta \in \Delta} \bar{r}(\delta)\, p(\delta). \tag{4.12}$$

Then for $r=(r(\delta))=(\int_k e^{jk\cdot\delta}d\mu(k))$ one has $(r, p)=\int_k p(k)\,d\mu(k)$. Let E be the set of all extendible vectors, E^0 its interior. One then has the following characterization of the set of extendible vectors in R^{2m+1}.

THEOREM [4, 21, 30]. *The vector is extendible if and only if* $(r, p)\geqslant 0$ *for all* $p\geqslant 0$.

Returning to the earlier question as to when $r_\delta=\int_k e^{jk\delta}f(k)\,dk$ for some $f\geqslant 0$ we have the result of Lang and McClellan [21].

THEOREM [21]. *If every neighborhood of every point in K has positive* μ-*measure, then*

(1) *If f is uniformly bounded away from zero over K, then* $r=(r(\delta))=(\int_k f(k)\,e^{jk\cdot\delta}d\mu(k))$ *is in* E^0.

(2) *If* $r\in E^0$, *then* $r=(r(\delta))=(\int f(k)\,e^{jk\cdot\delta}d\mu(k))$ *for some continuous strictly positive function f.*

The interior E^0 of E and the boundary ∂E of E have the simple characterizations

$$\begin{aligned} E^0&=\{r\in E\,|\,(r,p)>0 \quad \text{whenever} \quad p\geqslant 0 \text{ and } p\not\equiv 0\}.\\ \partial E&=\{r\in E\,|\,(r,p)=0 \quad \text{for some} \quad p\geqslant 0 \text{ and } p\not\equiv 0\}.\end{aligned} \tag{4.13}$$

It is thus natural to ask in higher dimensions when the $r(\delta)=\int_K(1/P(k))\,e^{i\delta\cdot k}dk$ for some positive trigonometric polynomial $P(\lambda)$ on K. Recall that this is the form of the maximum entropy solution in one dimension.

THEOREM [22]. *Every* $r\in E^0$ *can be written as* $r(\delta)=\int_K(e^{j\cdot\delta}/p(k))\,d\mu(k)$ *for some positive trigonometric polynomial p if and only if* $\int_K(1/p(k))\,d\mu(k)=\infty$ *for all p such that* $p(k)=0$ *for some* $k\in K$.

Woods [39] showed that every $r\in E^0$ may be represented as $r(\delta)=\int_K(1/p(k))\,e^{i\delta\cdot k}dk$ for all $\delta\in\Delta$ for $K=[-\pi,\pi]^2$ and Δ consisting of vectors with integral vectors with integral components and including the standard coordinate vectors. However, if $K=[-\pi,\pi]^n$ and Δ as above with $n\geqslant 3$ then the above theorem can be used to prove the existence of an r in E^0 that cannot be written in the above form. This example is due to Bruce R. Musicus; see [22]. This is an existence argument and as far as we know no construction of an explicit r has been given. Such a construction would be of interest.

Thus in higher dimensions we observe some breakdown of the maximum entropy method. Recently Goodrich and Steinhardt [12] investigated a different solution for these examples.

THEOREM [12]. *Let* $K=[-\pi,\pi]^n$, $n\geqslant 1$, $\Delta=\{0,\pm\delta_1,\ldots,\pm\delta_n\}$, *and take* $r\in E^0$. *Then there exists a unique trigonometric polynomial P such that* $r(\delta)=\int_K \max(P(k),0)\,e^{ik\cdot\delta}dk$ *for* $\delta\in\Delta$.

This solution is also known to exist in the other examples given by Lang and McClellan and is a finite parameterization fot E^0 with the number of unknowns (the coefficients of P) equal to the number of constraints (the number of elements of Δ). One can also show the existence of solutions of the form $\max(P(k),0)^{1/p-1}$ for $1<p<\infty$. This approach, as in the maximum entropy method, is one of Spectral Optimization and depends on minimizing the functional

$$I(f)=\int_K |f(k)|^p dk$$

over all $f\geqslant 0$ on K and such that $r(\delta)=\int_K f(k)\,e^{jk\cdot\delta}dk$ for all $\delta\in\Delta$. See [12] and [10] for further analytical and numerical studies.

REFERENCES

1. A. BEURLING, On the spectral synthesis of bounded functions, *Acta Math.* **81** (1949), 225–238.
2. A. BEURLING, Sur les spectres des fonctions, "Analyse Harmonic," Colloques Internationaux du Centre National de la Recherche Scientifique, no. 15, pp. 9–29, CNRS, Paris, 1949.
3. P. BUTZER, The Shannon Sampling Theorem and some of its generalizations: An overview, *in* "Constructive Function Theory," pp. 258–274, Bulgarian Acad. Sciences, Sofia, 1983.
4. A. CALDERON AND R. PEPINSKY, On the phases of Fourier coefficients for positive real functions, *in* "Computing Methods and the Phase Problem in X-Ray Crystal Analysis," pp. 339–348, X-Ray Crystal Analysis Laboratory, Dept. of Physics, Penn State, 1952.
5. H. DYM AND H. MCKEAN, "Gaussian Processes, Function Theory, and the Inverse Spectral Problem," Academic Press, New York, 1976.
6. R. GOODRICH AND K. GUSTAFSON, Weighted trigonometric approximations in $\mathscr{L}_2(R^n)$, *in* "Quantum Mechanics in Mathematics, Chemistry, and Physics" (K. Gustafson and W. Reinhardt, Eds.), pp. 343–350, Plenum, New York, 1981.
7. R. GOODRICH AND K. GUSTAFSON, Weighted trigonometric approximation and inner–outer functions on higher dimensional Euclidean spaces, *J. Approx. Theory* **31** (1981), 368–382.
8. R. GOODRICH AND K. GUSTAFSON, Higher dimensional spectral factorization with applications to digital filtering, *in* "Spectral Theory of Differential Operators," pp. 199–205, (I. Knowles and R. Lewis, Eds.), North-Holland, Amsterdam, 1981.
9. R. GOODRICH AND K. GUSTAFSON, Regular representation and approximation, *Colloq. Math. Soc. János Bolyai* **35** (1983), 539–550.
10. R. GOODRICH AND K. GUSTAFSON, Spectral Optimization, in preparation.
11. R. GOODRICH, K. GUSTAFSON, AND B. MISRA, On irreversibility and *K*-flows, *J. Statistial. Physics*, to appear.
12. R. GOODRICH AND A. STEINHARDT, L^2 Spectral Estimation, *J. Statistial. Physics*, to appear.

13. C. GRAHAM AND O. MCGEEHEE, "Essays in Commutative Harmonic Analysis," Springer-Verlag, New York, 1979.
14. U. GRENADER AND G. SZEGÖ, "Toeplitz Forms and Their Applications," Univ. of California Press, Berkeley, 1958.
15. K. GUSTAFSON AND B. MISRA, Canonical commutation relation of quantum mechanics and stochastic regularity, *Lett. Math. Phys.* **1** (1976), 275–280.
16. G. HARDY, A theorem concerning Fourier Transforms, *J. London Math. Soc.* **8** (1933), 227–231.
17. E. HEWITT AND K. ROSS, "Abstract Harmonic Analysis," Springer-Verlag, New York, 1970.
18. D. HILBERT, Ueber die Darstellung definiter Formen als Somme von Formenquadraten, *Math. Ann.* **32** (1888), 342–350.
19. L. HORMANDER, "Linear Partial Differential Operators," Academic Press, New York, 1963.
20. L. KOOPMANS, "The Spectral Analysis of Time Series," Academic Press, New York/London, 1974.
21. S. W. LANG AND J. H. MCCLELLAN, Spectral estimation for sensor arrays, *IEEE Trans. Acoust. Speech Signal Process* **ASSP-31**, No. 2 (April 1983) 349–358.
22. S. W. LANG AND J. H. MCCLELLAN, Multidimensional MEM spectral estimation, *IEEE Trans. Acoust. Speech Signal Process* **ASSP-30**, No. 6 (Dec. 1983), 880–887.
23. S. W. LANG, A positive-definite matrix which is not extendible, *IEEE Trans. Acoust. Speech Signal Process* **ASSP-32**, No 4 (August 1984).
24. L. H. LOOMIS, "An Introduction to Abstract Harmonic Analysis," Van Nostrand, New York, 1953.
25. E. LOW, Inner functions and boundary values in $H^\infty(\Omega)$ and $A(\Omega)$ in smoothly bounded pseudoconvex domains, *Math. Z.* **185** (1984), 191–210.
26. A. PAPOULIS, The Factorization Problem for time-limited functions and trigonometric polynomials, *IEEE Circuits and Systems* **CAS**-25 (1978), 41–45.
27. R. M. ROBINSON, Some definite polynomials which are not sums of squares of real polynomials, *in* "Selected Questions of Algebra and Logic," pp. 205–224, Acad. Sci. USSR, Novosibirsk, 1973.
28. L. RUBEL, Internal and external analytic functions of several complex variables, preprint.
29. L. RUBEL AND A. SHIELDS, The failure of interior-exterior factorization in the polydisk and the ball, *Tôhoku Math. J.* **24** (1972), 409–413.
30. W. RUDIN, The extension problem for positive-definite functions, *Illinois J. Math.* **7** (1963), 532–539.
31. W. RUDIN, "Real and Complex Analysis," McGraw–Hill, New York, 1966.
32. W. RUDIN, "Function Theory in the Unit Ball of C^n," Springer–Verlag, New York, (1980).
33. L. SCHWARTZ, Sur une propriété de synthèse spectrale dans les groupes non compacts, *C. R. Acad. Sci. Paris* **227** (1948), 424–426.
34. W. SPLETTSTOSSER, Sampling approximation of continuous functions with multidimensional domain, *IEEE Trans. Inform. Theory* **IT-28** (1982), 809–814.
35. J. STEWART, Positive definite functions and generalizations, an historical survey, *Rocky Mountain J. Math.* **6**, No. 3, 408–433.
36. G. SZEGÖ, Über die Randwerte analytischer Functionen, *Math. Ann.* **84** (1921), 232–244.
37. J. TUKEY, The sampling theory of power spectrum estimates, *in* "Proceedings, Symp. Appl. Autocorrelation Anal. Phys. Problems," NAVEXOS P-735, ONR, Wash. D.C., 1949.
38. N. WIENER, Tauberian Theorems, *Ann. of Math.* **33**(2) (1932), 1–100.
39. J. W. WOODS, Two-dimensional Markov spectral estimation, *IEEE Trans. Inform. Theory* **IT-22** (1976), 552–559.