ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 33, Number 2, Summer 2003

CONTINUED FRACTIONS, WAVELET TIME OPERATORS, AND INVERSE PROBLEMS

KARL GUSTAFSON

ABSTRACT. The spectral properties of the analytic theory of continued fractions, orthogonal functions, and rational approximation, are examined from the point of view of operator theory. New historical and mathematical perspectives are provided. Then a general question is posed: what spectral information is needed to uniquely determine an operator? The time operator induced by an arbitrary wavelet basis is presented as an example. This question is then directed at continued fractions.

1. Introduction. This paper, invited for the *Conference on the analytic theory of continued fractions, orthogonal functions, rational approximation and related topics* in honor of William B. Jones's 70th birthday, will focus on the related topic of *operator theory* and how this fourth topic relates to the first three topics. This investigation grew out of a conversation between this author and Professor Jones a few months prior to the conference. In that conversation I asked, "what is the spectral theory (of the first three topics)?" and a few days later I found the recent paper [**27**] in my mailbox. That paper provides an excellent overview of orthogonal Laurent polynomials, moment theory, continued fractions, Gaussian quadrature, Stieltjes transforms, and linear functionals. As such, [**27**] provides a summary of results of 'spectral' type as seen from the viewpoint of the *continued fraction community* (for convenience and with apologies I will use this term to represent all three principal topics of this conference).

The first goal of this paper is to "answer my own question," to wit: to supplement the viewpoint of [**27**], and some of the extensive literature which it represents, by my own commentary here, as a representative of the *operator theory community.* Of course, to fully answer my question would be a lifetime's work, so here I will only be able to patch the two

Key words and phrases. Continued fractions, spectral theory, wavelets, time operators, inverse problems.
Received by the editors on September 30, 2001, and in revised form on February 20, 2002.

subjects together in a few ways. Still, I believe that I have provided some new and helpful perspectives.

A second goal will be to go further to synthesize a general question: given spectral information, when is there a *natural* associated linear operator A? By 'natural' of course I do not mean for example that one just take A to be the identity operator I on the linear span $\mathrm{sp}\,\{\phi_n\}$, when given a complete set $\{\phi_n\}$ of orthogonal functions. Rather, A is to embody as many special properties of the given spectral information as possible, for example A is to embody deeper underlying mathematical or physical structures which generated the $\{\phi_n\}$ in the first place. I will illustrate this proposed new general inverse problem by our recent results [**2, 3, 16**] in which we show that the 'natural' operator for all wavelet bases is a time operator.

The outline of the paper is as follows. Section 2 examines the issue of whether orthogonal polynomials originated from the theory of continued functions, or from operator theory. Section 3 compares spectral theory as defined and used in the continued fraction community, with its definitions and use in the operator theory community. Section 4 brings forth and exposits the very valuable contribution of Marshall Stone, in whose 1932 book one finds a detailed connecting of the two spectral theory viewpoints within a context of Jacobi matrices. Section 5 introduces the notion, from the operator theory viewpoint, of inverse problem, and recalls two of the most famous inverse problems, drum and interaction potential. Section 6 presents our recent results that all wavelet systems naturally induce a corresponding time operator which incorporates all essential mathematical structure underlying the wavelet system. Section 7 permits me to use the example of Section 6 to propose the new very general inverse problem described above. Section 8 partially answers this question for continued fractions.

2. Historical comment: Orthogonal polynomials. In my book [**18**], the position was taken that the subject of orthogonal functions originated first from physics (vibrating string problem) and then more generally from second order linear self-adjoint differential operator boundary value problems (Sturm-Liouville theory). Let us compare that position to the (not the same) position taken by the continued fraction community. For the latter I have consulted [**4–6, 10, 25–27, 39, 40**].

Indeed, Szegö [**39**, p. 54] asserts "Historically, the orthogonal polynomials $\{p_n(x)\}$ originated in the theory of continued fractions." Here the class of polynomials $\{p_n(x)\}$ under treatment are [**39**, p. 26] those resulting from orthogonalization of $1, x, x^2, \dots, x^n, \dots$ with respect to the inner product

$$\langle f, g\rangle_\alpha = \int_a^b f(x)g(x)d\alpha(x) \tag{2.1}$$

where $\alpha(x)$ is a given nondecreasing weight-function with infinitely many increasing points in the interval (a, b) and such that all "moments"

$$c_n = \int_a^b x^n d\alpha(x) \tag{2.2}$$

exist, $n = 0, 1, 2, \dots$. I note specifically that Szegö [**39**, p. 3] also accepts trigonometric polynomials

$$g(\theta) = a_0 + a_1 \cos\theta + b_1 \sin\theta + \cdots + a_m \cos m\theta + b_m \sin m\theta \tag{2.3}$$

and in particular states that some of these, e.g., Tchebichef polynomials of the first and second kind, "play a fundamental role in subsequent considerations." In other words, polynomials in which the argument x is a trigonometric function are an integral part of the subject of orthogonal polynomials. One may compare these two assertions to similar statements in Szegö [**40**]. Outlining Stieltjes' approach to the moment problem, we find [**40**, p. 7] "thus, the denominators of the approximating fractions of the continued fraction are identical with the orthogonal polynomials. This explains the close relation of the orthogonal polynomials to the theory of continued fractions." Also [**40**, p. 5] "The simplest special case of the Jacobi polynomials is that of the Chebichev polynomials $d\alpha(x) = (1 - x^2)^{-1/2}dx$, $a = -1$, $b = 1 \dots$. Thus the classical Fourier series appear as a special case of the orthogonal polynomial expansions."

On the other hand, in [**18**, p. 23], in connection with the physical vibrating string problem and its solution by separation of variables, I point out that Bernoulli (1755) approximated the solution by a finite Fourier sine series, that Euler (1777) calculated the coefficients, i.e., the moments of the physical solution relative to the trigonometric polynomials with weight function $\alpha(x) \equiv 1$, and Lagrange (1760) similarly employed trigonometric polynomials to solve the problem. The orthogonality of the sine functions goes back to Wallis (1616–1703). I also point out [**18**, p. 25] that Legendre (1785) solved a physical problem of gravitational attraction of ellipsoidal solids of revolution by means of a separation of variables procedure which reduced the PDE (partial differential equation) to an ODE (ordinary differential equation, the one now bearing his name) and thence to Legendre polynomials. Whitaker and Watson [**42**, p. 302] confirm this work, give a precise citation to Legendre's paper, and point out the other, more physically motivated name, zonal harmonics, for his polynomials in that application.

So above are the two somewhat different historical viewpoints. From the continued fraction community viewpoint, e.g., Szegö [**39, 40**]; see also [**25**–**27**]), orthogonal polynomials go back to Stieltjes (1884), also [**27**, p. 52] to Gauss (1814), also [**10**, p. 209] to Tchebichef (1858) and others, as related to continued fractions. From the operator theory viewpoint I take them back to Euler (1777) and others in the same epic. The distinction would seem to be: whether one allows trigonometric arguments, e.g., $x = \cos\theta$. From the physical, i.e., operator theory, point of view, that is certainly justified.

Brezinski [**6**] gives a broader history of continued fractions and Padé approximants, in a truly extraordinary expository effort. He takes continued fractions back more than 2000 years, to Euclid. So certainly continued fractions is a very old subject, far older than that of orthogonal polynomials. However, when he turns to the latter topic [**6**, p. 213], he falls back on the viewpoint that orthogonal polynomials originated from certain types of continued fractions, notably those of Gauss, Jacobi, Christoffel, Mehler, Chebyshev, Heine, and Markov. But he is careful to also note that earlier, before any general theory existed, orthogonal polynomials were discovered by Lagrange, Legendre, Fourier (and later) Laplace, Laguerre, and Hermite, among others. I note that those earlier orthogonal polynomials were generated by ordinary or partial differential equations, i.e., from operator theory. See for example [**18**, Chapter 2].

Extremely valuable is the short history provided by Askey [**4**], and the citations therein. He points out "the analogy between Sturm-Liouville differential equations and the three-term recurrence relation of orthogonal polynomials has been used through the years." Chihara [**10**, pp. 80, 85] points out that Wallis's (1655) recurrence relations for the partial numerators A_n and the partial denominators B_n of a continued fraction's nth approximant $C_n = A_n/B_n$, also lead directly to the connection between orthogonal polynomials and continued fractions. Brezinski [**5**, p. 71] establishes a one-to-one correspondence of orthogonal polynomials and Jacobi matrices.

To conclude this brief historical comment about whether orthogonal polynomials originated from continued fraction theory or from operator theory, I would conclude: both. From the early classical physics of vibrating systems, an orthogonal system of functions $\{\phi_n\}$ arises as the fundamental vibration modes of the system. The ϕ_n are then the operator eigenfunctions corresponding to the eigenvalues λ_n which are the fundamental frequencies of the system. But this operator theory only originates orthogonal polynomial systems $\{\phi_n\}$ which are special functions for some physical differential equation or other eigenvalue problems. The continued fraction origin of polynomial systems $\{\phi_n\}$ need not correspond to any physical system.

3. Spectral theory: Moment problems. The question "what is the spectral theory" of continued fractions, orthogonal functions, and rational approximation, has too many ramifications to fully develop here. However, we may reduce this general question by first turning to the spectral theory of operator theory, which is well-defined. The following may be found in many books, e.g., [**21, 37**] among others. I also point out the recent paper [**17**] which provides many references and also connects the finer points of spectral theory to some interesting physical questions.

Given a bounded linear operator A mapping a complete normed linear space X into itself, then the set of complex numbers λ such that $A-\lambda I$ maps X onto X with bounded inverse $(A-\lambda I)^{-1}$, is called the resolvent set $\rho(A)$ of the operator A. The inverse $(A-\lambda I)^{-1}$ is called the resolvent operator R_λ. The complement $\sigma(A) = \rho(A)^C$ is called the spectrum of A. This can be broken down into point spectrum, continuous spectrum, and residual spectrum, see e.g., [**17**].

The special and important case in which X is a Hilbert space, i.e., when the norm is given in terms of an inner product, and in which A is a self-adjoint operator, possesses its own spectral theory. Then the spectrum $\sigma(A)$ is a subset of the real line, and A can be represented as an integral over the line, $A = \int \lambda dE_\lambda$, where the E_λ are projection valued measures. In practical terms this means that

$$\langle Af, g\rangle = \int_{-\infty}^{\infty} \lambda d\langle E_\lambda f, g\rangle \tag{3.1}$$

for all functions f and g in the Hilbert space. Moreover by the spectral functional calculus one also has

$$\langle A^n f, g\rangle = \int_{-\infty}^{\infty} \lambda^n d\langle E_\lambda f, g\rangle \tag{3.2}$$

for all powers A^n.

On the other hand, in the continued fraction literature, the spectrum is defined [**27**, p. 58] as follows. Let $\alpha(t)$ be a real valued function. Then the spectrum $\sigma(\alpha)$ consists of those points t in the domain of α such that α is strictly increasing on some interval about t. The special case of distribution functions $\Phi^c(a,b)$, namely, functions $\alpha(t)$ which are bounded, nondecreasing, and possessing of infinitely many points of increase in (a,b), are the most important when considering moment problems ([**1, 6, 10, 25**–**27, 35, 39, 40**] and citations therein). The original Stieltjes moment problem was: given real numbers c_n (the moments), find a distribution function α such that

$$c_n = \int_0^{\infty} x^n d\alpha(x), \quad n = 0, 1, 2, \ldots \tag{3.3}$$

There are many other moment problems on different intervals (a,b) but for comparison to operator spectral theory let me restrict attention here to $(a,b) = (-\infty, \infty)$. Then let us consider a general moment problem. Suppose that we are given a Riesz basis of the given Hilbert space X: a family $\{r_n\}$ whose closed linear span $\overline{\text{sp }\{r_n\}} = X$ and which moreover has the property that $\{r_n\}$ is obtainable by an isomorphic linear map V from some orthonormal basis $\{\phi_n\}$. Riesz bases play a key role in the modern theory of wavelets and for more about both, see for example [**11, 29, 38**]. For simplicity at this point let us first specialize to $X = \mathcal{L}^2(\mathbf{R})$. Then a general moment problem is, given real numbers

c_n (the moments) where n runs through any given specified index set, find a distribution function α such that

$$c_n = \int_{-\infty}^{\infty} r_n d\alpha(x) \tag{3.4}$$

One can of course go to Riesz bases $\{r_n\}$ in $\mathcal{L}^2(a,b)$ for any interval and state (3.4) in the same way. A point to be emphasized here is: The polynomials $\{p_n\}$ under treatment in the continued fraction orthogonal polynomial literature, e.g. [**39**], are a special class resulting from orthogonalization of the Riesz basis $1, x, x^2, \ldots, x^n, \ldots$ in $\mathcal{L}^2(0,\infty)$.

More generally, one can consider the moment problem, given real (or complex) numbers c_n, find a vector α in X such that

$$c_n = \langle r_n, \alpha\rangle_\beta. \tag{3.5}$$

Here X can be any Hilbert space and $\{r_n\}$ can be any Riesz basis in it and $\langle,\rangle_\beta$ can be any inner product on X. This includes the special case (2.1) where X is an $\mathcal{L}^2(a,b)$ space and $\alpha = \beta$ is an absolutely continuous function.

Now let us compare the two perspectives.

From the operator theory point of view, the continued fraction spectrum may be regarded as concerned with just one of the distribution functions

$$\alpha(\lambda) = \langle E_\lambda f, f\rangle \tag{3.6}$$

from a spectral representation (3.1). That is, given a bounded self-adjoint operator A on some Hilbert space X and As spectral family of projections E_λ, $-\infty < \lambda < \infty$, see e.g. [**17**], and having chosen nonnull element $f = g$ in X, then $\alpha(\lambda)$ in (3.6) is real valued, nondecreasing, and increasing from $\alpha(\lambda) \to 0$ as $\lambda \to -\infty$ to $\alpha(\lambda) \to 1$ as $\lambda \to +\infty$. Of course $\alpha(\lambda)$ in (3.6) has total variation one and is defined on $(a,b) = (-\infty,\infty)$ but the general case of $\Phi^c(a,b)$ of distribution functions of bounded variation on any interval is essentially the same. In the operator spectral theory, by projection one can restrict the $\alpha(\lambda)$ of (3.3) to any prescribed interval, and one often does [**17**]. When the point λ is a jump increase of $\alpha(\lambda)$, it means f is an eigenfunction of A and λ is an eigenvalue of A and such λ are called the point spectra

$\sigma_p(A)$ of A. The other increase points λ for some f and corresponding $\alpha_f(\lambda)$ are in the continuous spectra $\sigma_c(A)$ of A. If for all f the $\alpha(\lambda)$ are all flat at λ, then λ is in the resolvent set $\rho(A)$. Moreover one can consider moment problems such as (3.3) as special cases of (3.2).

Above, I stated a more general moment problem, for arbitrary Riesz basis $\{r_n\}$. It is less immediate how to associate an operator A so that (3.4) or (3.5) is included in A's spectral representation calculus, and I don't want to go into that here. So instead let us recall that (3.4) or (3.5) defines a linear functional l_α by the inner product action

$$l_\alpha f = \langle f, \alpha \rangle. \tag{3.7}$$

and conversely all linear functionals on X are generated by such vectors α. So all moment problems are, from the point of view of operator function theory, the question: how many vectors f in the Hilbert space X do we need to know to uniquely determine a distribution function α with the properties we want for it.

I may also state moment problems in terms of pure operator theory, without reference to spectral notions. Let X be the Banach space of all functions of bounded variation on the finite interval (a, b). Let Y be the Banach space of all bounded sequences. Define a linear operator A with domain X and mapping into Y, $A\alpha = y$, $\alpha \in X$ and $y = \{c_n\}_{n=0}^{\infty}$, where

$$c_n = \int_a^b t^n d\alpha(t). \tag{3.8}$$

Then the moment problem may be regarded as the question: what is the range $\mathcal{R}(A)$? We know A maps into Y because all moments c_n are bounded in terms of the total variation of the distribution function $\alpha(t)$ and the interval length $b - a$. But $\mathcal{R}(A)$ is not all of Y. One could also describe the moment problem for infinite intervals this way by appropriate modification, e.g., that $\alpha'(t)$ have compact support.

One could roughly recast the considerations above as: classical analysis versus functional analysis. The former came first, the latter is more general. In some sense such a conclusion partially answers my original question stated in the Introduction to this paper. Oversimplistically, the conclusion is that the 'spectral theory' residual to continued fraction theory, orthogonal function theory, and rational approximation theory, is that part of classical analysis which now fits into the modern spectral theory of operator theory. Some of those currently working in the classical frame have noticed this and are clarifying the situation (as I am, in the present paper). I note in particular the work of Sri Ranga [**35**] as reported in [**27**]. Of course parts of the Russian community, e.g., as represented in books such as [**1**], which I will not discuss here, have already taken this unified point of view, combining to some extent the classical and functional formulations. In this connection we should also note that there are large parts of functional analysis which can be crudely described as generalizing the complex numbers z to operators A. For example, just as continued fractions are limits of their finite rational approximations R_n/S_n which may be expanded, in the simple multiplicity case, into partial fractions of the form $\sum \lambda_k^n (z - t_k^n)^{-1}$, one can generalize all of this to operator valued resolvent expressions $\sum \lambda_k (A - \mu_k)^{-1}$ and the like. A physicist might call this first quantization. Then one could work with operators on operators. We might call this second quantization. Certainly one could define and consider moment problems in those general operator theory contexts. However in such abstraction one needs specific physical problems to indicate the right questions.

4. Spectral theory: Jacobi matrices. Brezinski [**6**, pp. 228, 230, 286] accurately describes how Stieltjes' work on continued fractions led to later developments in the spectral theory of operators. He points out [**6**, p. 291] that some of these later developments, especially the connections between continued fractions, Jacobi matrices, and self-adjoint operators with simple spectra, were treated in the book [**37**]. Askey [**4**] also briefly notes Stone's [**37**] book, as does Chihara [**10**]. Gragg [**15**] looks at the one to one correspondence of Laurent series moments with infinite Hankel matrices, but with no reference to Stone's work [**37**]. See also the treatment in [**25**, pp. 249–255]. In any case, here I want to go further.

The first point I would like to make is that Stone's book [**37**] principally constitutes an application of continued fractions to operator theory, and not the other way around. This application comes in the last chapter of Stone's book [**37**]. In particular he first shows [**37**, p. 531, Theorem 10.23] that reduced Jacobi matrices A_p are essentially self-adjoint. This means [**17, 37**] that their closures $\bar{A}_p$

are self-adjoint, have defect indices (0,0), and hence have only real spectra. There he also identifies the associated distribution function $\rho_p(\lambda) = \|E^{(p)}(\lambda)g_1\|^2$. I used the equivalent notation $\alpha(\lambda) = \langle E_\lambda g, g\rangle$ in (3.6). These distribution functions $\rho_p(\lambda)$ do not have an infinite number of points of strict increase, they have only $p+1$ values. The p jumps therefore correspond to point spectra $\sigma_p(\bar{A}^p)$ which are exactly the roots of a polynomial. The integral [**37**, p. 532]

$$I(l;\rho_p) = \int_{-\infty}^{\infty} \frac{1}{\lambda - l}\, d\rho_p(\lambda) \tag{4.1}$$

is a rational function which can be expressed as a (finite) continued fraction. The integral (4.1) can also be expressed as a series $-\sum_n^\infty c_n/l^{n+1}$ which is analytic at infinity and in which the coefficients are the moments

$$c_n = \int_{-\infty}^{\infty} \lambda^n d\rho_p(\lambda). \tag{4.2}$$

Stone [**37**, pp. 553–583] then goes on to treat any general arbitrary infinite dimensional Jacobi matrix A as a limit of the finite rank Jacobi matrices A_p I have described above. The distribution functions $\rho_p(\lambda)$ converge to the $\rho(\lambda)$ of A except at jumps. There are many interesting details here, and in the sense of the well known saying that "every author writes his book for the last chapter," this application of the spectral theory to Jacobi operators is the crowning application that Stone gives. Although Stone's [**37**] goal in this last section of his last Chapter 10 is to construct all of the spectral theory of Jacobi matrices (his fourth and final application, the other three applications being integral and differential operators) he uses continued fractions only at [**37**, p. 559] where he states: "We may present the facts just established in a somewhat different light by indicating their relation to the theory of infinite continued fractions." In particular he uses continued fractions to determine necessary and sufficient conditions for the operator A to be essentially self-adjoint, i.e., defect indices (0,0), in contrast to the other possibility, defect indices (1,1). I mention that in the latter case of equal but nonzero defect indices, A can have many self-adjoint extensions $\widetilde{A}$, and Stone [**37**] constructs them all.

The second point I would like to make is that Stone here is using the same trick as Stieltjes used in 1884 in his first paper involving continued fractions. This trick is, and I refer to [**6**, pp. 228–231] from which I quote: "The continued fraction was only an intermediate step between the power series and the integral and served as a trick for summing divergent series." Recall that Stieltjes started with the Gaussian quadrature problem $\int_a^b f(t)\,dt \cong \sum_{i=1}^n A_i f(t_i)$ but by his continued fraction formulation he ended up with rational approximations to $\int_a^b (x-t)^{-1} f(t)\,dt$. Stone is working with the resolvent expressions (4.1) of the (assume ρ is absolutely continuous) form $\int_{-\infty}^{\infty} (\lambda - l)^{-1} \rho'(\lambda)\,d\lambda$. More precisely they are, in the case that A is self-adjoint or a self-adjoint extension of a Jacobi operator, and where $E(\lambda)$ is A's spectral family of projection operators,

$$\int_{-\infty}^{\infty} \frac{1}{\lambda - l} d\rho(\lambda) = \langle R_l g_1, g_1 \rangle_\rho \tag{4.3}$$

where $R_l =$ the resolvent operators $(lI - A)^{-1}$ on the range $\mathcal{R}(lI - A)$ even when that is not the whole Hilbert space.

The third point to be made here is that Stone's interest in Jacobi matrices stemmed largely from the fact that he was able to show that all self-adjoint operators H with simple spectra $\sigma(H)$ in a Hilbert space $\mathcal{H}$ are unitarily equivalent to a spectral representation as an operator A in a space $\mathcal{L}^2(\rho)$ and that operator A is a Jacobi operator. See in particular [**37**, Chapter 7, Section 3, Theorems 7.10 and 7.11]. In other words, although Stone's extensive treatment of Jacobi operators [**37**, Chapter 10, Section 4, pp. 530–614] is presented there as his fourth and final "application" of the general operator theory he had developed in the book, a closer reading of the book reveals that he employed in an essential way such matrix representations of operators earlier in the book, viz. Chapter 7; see also Chapter 3. Both because of their own interest and also because I am going to need some of these results from [**37**] later in Section 8, let me here delineate a few salient facts that Stone establishes.

A key issue is that of simple spectra and how that indirectly relates to a spanning basis and eventually relates to a particular choice of polynomials. More important is the existence of a cyclic vector for A, and more important than that is the choice of cyclic vector.

Given a self-adjoint operator H, let $\mathcal{M}$ denote $\overline{sp\,\{\phi_n\}}$ where ϕ_n are the eigenvectors of H, and let $\mathcal{N}$ denote $\mathcal{M}^\perp$. Modern parlance, e.g., [**17**], calls $\mathcal{M} \equiv \mathcal{H}_\rho$, the point spectrum subspace. Because H is self-adjoint, the whole Hilbert space $\mathcal{H}$ is the orthogonal sum of $\mathcal{M}$ and $\mathcal{N}$. Stone considers 3 cases [**37**, p. 247] that recur throughout the book: (1) no point spectrum, (2) some point spectrum, (3) all point spectrum. The terminology here is mine, for the details see [**37**]. Then simple spectrum, whether H has point or continuous spectrum, is defined on page 275. The main point established then is [**37**, Theorem 7.9], that a self-adjoint operator H has simple spectrum if and only if there exists a cyclic vector for H: some $f \neq 0$ such that $\mathcal{M}(f) = \mathcal{H}$. Here $\mathcal{M}(f)$ denotes the closed span of all $\mathcal{L}^2$ functions $F(H)$ applied to the vector f. One can conveniently think of $\mathcal{M}(f)$ as $\overline{sp\{f, Hf, H^2 f, \dots\}}$ in most cases. Theorem 7.12 establishes the matrix A corresponding to H, where the elements a_{mn} of A are given by

$$a_{mn} = \langle Hg_n, g_n\rangle = \int_{-\infty}^{\infty} \lambda G_n(\lambda)\overline{G_m(\lambda)}d\rho(\lambda) \tag{4.4}$$

where $\rho(\lambda) = \|E(\lambda)f\|^2$, the sequence $\{g_n\}$ depends on the choice of cyclic vector f, and $\{G_n\}$ are polynomials corresponding to $\{g_n\}$. Note that $\rho(\lambda) = \langle E_\lambda f, f\rangle$ is a characteristic function $\alpha(\lambda)$ in my (3.6) above.

The $\{g_n\}$ are determined from f as follows, and this is important. One goes back to Chapter 5, page 166, and you use the separability of the Hilbert space $\mathcal{H}$ to select a countable sequence $\{f_n\}$ whose linear span $sp\{f_n\}$ is dense in $\mathcal{D}(H)$ and moreover whose images $sp\{(H - iI)^{-1}f_n\}$ and $\{sp(H + iI)^{-1}f_n\}$ are also dense in $\mathcal{D}(H)$. That one can do the latter needs the operator H to be self-adjoint, H just symmetric will not do. Then one considers the combined countable set $\{\{f_n\}, \{(H - iI)^{-1}f_n\}, \{(H + iI)^{-1}f_n\}\}$ and let $\mathcal{M}$ denote the linear span of this set. Clearly $\mathcal{M}$ is a dense subspace, as also are the defect subspaces $\mathcal{R}(H \pm iI)$. Then the $\{g_n\}$ of Theorem 7.12 are a complete orthonormal set (called $\{\phi_n\}$ by Stone [**37**, p. 166]) obtained by Gram-Schmidt on $\mathcal{M}$ (or some other orthonormalization procedure on $\mathcal{M}$) so that $sp\{g_n\} = \mathcal{M}$.

The key result for Jacobi matrices is now [**37**, Theorem 7.13]. For H with simple spectrum, one can choose f and the $\{g_n\}$ such that $G_n(\lambda)$ is a polynomial of degree $n - 1$ for $n = 1, 2, 3, \dots$ and such that A is a tridiagonal Jacobi matrix

$$A = \begin{bmatrix} a_1 & b_1 & O & \cdots \\ \bar{b}_1 & a_2 & b_2 & \cdots \\ O & \bar{b}_2 & a_3 & \cdots \\ \vdots & & & \end{bmatrix} \tag{4.5}$$

with a_n real and $b_n \neq 0$. Theorem 7.14 establishes the converse: given a Jacobi matrix A, one can define a corresponding self-adjoint operator H which has simple spectrum. It is worth noting that the $\{g_n\}$ again play a critical role here. In particular one takes g_1 to be a cyclic vector, i.e., $\mathcal{M}(g_1) = \mathcal{H}$. Then with $\rho(\lambda) = \langle E(\lambda)g_1, g_1\rangle$ one obtains the spectral correspondence (4.4) where the polynomials $G_n(\lambda)$ are of degree $n-1$ and are determined by the recurrence relations

$$\begin{aligned} G_1(\lambda) &= 1, \quad G_2(\lambda) = (\lambda - a_1)/\bar{b}_1 \\ G_n(\lambda) &= \frac{(\lambda - a_{n-1})G_{n-1}(\lambda) - b_{n-2}G_{n-2}(\lambda)}{\bar{b}_{n-1}} \end{aligned} \tag{4.6}$$

I will use these facts in Section 8.

5. Classic inverse problems: Drums, potentials. Inverse problems are very important in science; they are often more difficult than their associated "direct problems," and often they are therefore not even treated in beginning treatises. For example, a CAT-scan of your head using magnetic resonance machines, or an ultrasound inspection of an interior part of your body, are inverse problems: determining shape from electromagnetic reflections. Radar is an inverse problem. Two of the most famous inverse problems in mathematics are the [**28**] classical "can you hear the shape of a drum?" and the [**9**] quantal "can you determine the interaction potential?" Another related inverse problem is the classical, i.e., wave equation, inverse scattering acoustic problem, e.g., determining underwater shapes, which I shall not discuss at all [**30**].

One may easily formulate moment problems such as those discussed in Section 3 as inverse problems: given the (3.7) 'measurements' $l_\alpha(f)$, how many 'test functions' f must one use to determine the weight function α? Regarding (3.5), how many moments, i.e., how many elements r_n of the Riesz basis $\{r_n\}$ are needed to determine α? As is well known, following the breakthrough of Stieltjes, continued fractions became a key tool to formulate necessary and sufficient conditions for the unique determination of α from specified function classes.

Even the simplest classical Fourier series expansion may be regarded as an (albeit, an easy one) inverse problem. Given the moments, i.e., the Fourier sine coefficients,

$$c_n = \frac{2}{\pi}\int_0^\pi (\sin nx)\alpha(x)\,dx \quad n = 1, 2, 3, \ldots, \tag{5.1}$$

we know we can then determine $\alpha(x)$ uniquely among all functions in $\mathcal{L}^2(0,\pi)$ up to $\mathcal{L}^2$ equivalence. That we can actually then construct $\alpha(x)$ then of course requires a series expansion of $\alpha(x)$ in terms of the orthogonal eigenfunctions $\phi_n(x) = \sin nx$ of the Sturm-Liouville problem

$$-u''(x) = \lambda u(x), \quad 0 < x < \pi; \quad u(0) = u(\pi) = 0 \tag{5.2}$$

Not so easy are the classic inverse problems [**28, 8**].

When Kac [**28**] first posed his problem, he showed that one can "hear" the area of a compact convex drum. That would correspond to hearing the lowest frequency standing wave. Let us suppose that we can hear all of the eigenfrequencies emanating from a vibrating drum. Let us call two domains (drums) isospectral if they produce the same (reduced wave equation) eigenvalues, and let us call them isometric if they are geometrically congruent. Then the question stated mathematically is: are two isospectral domains necessarily isometric? It turned out that the answer was no, see, e.g., [**14**] and citations therein.

The more important quantum inverse scattering problem [**9**] concerns Schrodinger equations and S matrices. I refer to [**9**] for all details, of which there are many. If one considers only radially symmetric interaction potentials $V(r)$, then the reduced Schrodinger equation is

$$\frac{d^2}{dr^2}\psi_l(k,r) + \left(k^2 - \frac{l(l+1)}{r^2}\right)\psi_l(k,r) = V(r)\psi_l(k,r) \tag{5.3}$$

for the partial wave function $\psi_l(k,r)$ at a given angular momentum l. The eigenvalues are the wave numbers k^2. Under certain assumptions restricting the class of potentials V to be allowed, the problem was eventually solved to some extent by Gel'fand and Levitan in 1951. Jost and Kohn, Levinson, Krein, Prosser, Marchenko and others also made contributions to the solution. It turned out that one needed to know the phase shift for all scattering energies. I mention that an easier discretized problem, in which one replaces the diffusion term by a first order centered difference, was solved by Case and Kac [**8**] using orthogonal polynomials. I refer the reader to [**9**] for further details.

Thus in both of these famous inverse problems, knowing just the frequencies, i.e., the spectrum, was not enough. One also needs to know the corresponding eigenfunctions, and sometimes more. Although at first such difficulties may put one off, on the other hand such difficulty indicates an underlying richness of structure to be discovered and used. This point will be important for Section 7 so let me make a couple of related observations here. In moment problems considered as inverse problems we were only trying to discover a function (or measure) from its moments with respect to some given basis. In the vibrating drum problem [**28**] we assumed that we already knew the operator (wave equation) and the boundary conditions (Dirichlet type) and we were trying to discover the underlying geometry. Technically speaking [**18**] that geometry is part of the operator so we can say: we were trying to discover part of an operator, given the rest of it. In the quantum inverse-scattering problem [**9**] again we were trying to discover part of an operator, i.e., the interaction potential $V(r)$ in (5.3). The problem I will pose in Section 7 may also be regarded as such an inverse problem. In the next section I give the motivating example.

6. Wavelet systems: Time operators. Wavelets form a modern theory of a certain class of orthogonal functions. This is a large, still evolving, subject and in no way can I adequately describe it here. See [**11, 29, 38**] among many excellent recent books. Also I mention the paper [**13**] on frames. From one point of view, wavelets evolved from frames. The paper [**13**] is a good one for the frame theory; it also connects to certain aspects of moment theory for not necessarily independent expansion bases $\{r_n\}$.

An entirely different subject is that of ergodic theory, more specifically, Kolmogorov dynamical systems. These are stochastic processes representing chaotic dynamics. Again I make no attempt to describe this subject here. Some years ago [**20, 32, 19**] we developed a theory of Time operators for these processes. About 10 years ago we realized that Time operators could be intimately related to wavelets. Recently [**2, 3**], after considerable delay, we have published these results. I gave an earlier preliminary presentation in [**16**], and all necessary background references to the above subjects may be found in the book [**16**].

The basic idea can be presented here in terms of the Haar basis

$$\begin{aligned} \phi_1(t) &= \chi_{[0,1]}(t) \\ \phi_{2^n+j}(t) &= 2^{n/2}[\chi_{[0,1]}(2^{n+1}t - 2j + 2) - \chi_{(0,1]}(2^{n+1}t - 2j + 1)] \end{aligned} \tag{6.1}$$

where $j = 1, 2, \ldots, 2^n$ and $n = 0, 1, 2, \ldots$. As is well-known the Haar basis (6.1) is a basis for all $\mathcal{L}^p(0,1)$ spaces, $1 \le p < \infty$, and it is an orthonormal basis for $\mathcal{L}^2(0,1)$, where we wish to consider it. See [**18**] for some related history on Haar's other contributions to mathematics. The Haar basis is historically important in both wavelet theory and Kolmogorov dynamical systems.

To describe our results [**2, 3, 16**] we need now the notion of a multi-resolution analysis from wavelet theory. A wavelet multi-resolution analysis [**11, 29, 38**] of square integrable functions $\mathcal{L}^2(\mathbf{R})$ is a sequence of Hilbert subspaces $\mathcal{H}_n$ of $\mathcal{L}^2(\mathbf{R})$ with the properties (MRA)

$$\begin{aligned} &1) \quad \mathcal{H}_n \subset \mathcal{H}_{n+1} \\ &2) \quad \bigcap_n \mathcal{H}_n = \{0\} \\ &3) \quad \bigcup_n \mathcal{H}_n \quad \text{is dense in } \mathcal{L}^2(\mathbf{R}) \\ &4) \quad f(x) \in \mathcal{H}_n \Longleftrightarrow f(2x) \in \mathcal{H}_{n+1} \\ &5) \quad \text{There exists } \phi \in \mathcal{H}_0 \text{ such that the set of integer translates} \\ &\qquad \phi(x-n) \text{ is an orthonormal basis for } \mathcal{H}_0. \end{aligned} \tag{6.2}$$

The function ϕ is usually called the scaling function of the wavelet MRA. Then a corresponding wavelet function ψ may always be obtained in standard ways. For example, the Haar wavelet ψ is obtained from ϕ by the differencing $\psi(x) = \phi(2x) - \phi(2x-1)$.

The following result has been shown in [**2, 3, 16**].

Theorem 6.1. *Any wavelet multi-resolution analysis on $\mathcal{L}^2(\mathbf{R})$ is actually a bilateral shift of scalings on $\mathcal{L}^2(\mathbf{R})$ with countable infinite multiplicity such that the wavelet subspace W_0 is a wandering generating subspace for the shift.*

Proof. (Outline). By the multi-resolution property 4) the space $\mathcal{H}_{n+1}$ is the image of the space $\mathcal{H}_n$ under the unitary transformation

$$Vf(x) = \sqrt{2}\, f(2x). \tag{6.3}$$

The operators V^n provide a unitary representation on $\mathcal{L}^2(\mathbf{R})$ of the discrete scaling maps $x \to 2x$. The multi-resolution properties 1), 2), 3) then are the conditions for V to be a bilateral shift. Such unitary shifts have absolutely continuous spectrum of uniform multiplicity. Thus the wavelet space

$$W_0 = \mathcal{H}_1 \ominus \mathcal{H}_0 = V\mathcal{H}_0 \ominus \mathcal{H}_0 \tag{6.4}$$

is exactly the wandering generating subspace of the shift operator V. W_0 is infinite dimensional because the functions $\psi(x-n)$, $n \in \mathbf{Z}$, form an orthonormal basis of it.

Theorem 6.1 enables us to now define a time operator for Haar's system (6.1) and, more generally, for any wavelet multi-resolution analysis. We recall that Time operators are defined in statistical physics as follows ([**33, 20, 32, 19, 16, 2, 3**], chronologically, see also the citations therein). Given a unitary evolution group U_t, $t \in \mathbf{R}$, on a separable Hilbert space, an internal time operator for U_t is a self-adjoint operator T with dense domain $\mathcal{D}(T)$ on which

$$U_{-t}TU_t = T + tI. \tag{6.5}$$

The time operator T allows us to attribute the average age $\langle f, Tf\rangle$ to the states f in $\mathcal{D}(T)$. The average internal age of the evolved state $U_t f$ advances in step with the external clock time

$$\langle U_t f, TU_t f\rangle = \langle f, Tf\rangle + t. \tag{6.6}$$

Time operators are canonically conjugate to the self adjoint generators L of the unitary group $U_t = e^{-iLt}$:

$$[L, T] = -iI. \tag{6.7}$$

For discrete parameter $n \in \mathbf{Z}$ the relations (6.5), (6.6) become, respectively,

$$U^{-n}TU^n = T + nI \tag{6.8}$$

$$\langle U^n f, TU^n f\rangle = \langle f, Tf\rangle + n. \qquad \square \tag{6.9}$$

From these facts we may obtain

Theorem 6.2. *The time operator T of statistical physics is the natural operator for which the Haar System* (6.1) *is its eigenbasis. The natural setting of Haar's wavelet is that of an eigenbasis canonically conjugate to the Haar detail refinement.*

Proof. (Outline). Let P_n be the projections of $\mathcal{L}^2(\mathbf{R})$ onto the approximating subspace $\mathcal{H}_n$ of the Haar wavelet. Then the wavelet MRA properties 1), 2), 3) imply that

$$E_n = P_{n+1} - P_n \tag{6.10}$$

is a spectral resolution of a self-adjoint operator q. The scaling operator V from (6.3) and the projections P_n clearly satisfy the imprimitivity relation

$$P_{n+1} = VP_nV^{-1}. \tag{6.11}$$

If we now define

$$T = \sum_{n\in\mathbf{Z}} nE_n \tag{6.12}$$

with the E_n from (6.10), then from (6.11) we obtain

$$V^{-n}TV^n = T + nI \tag{6.13}$$

which is the imprimitivity relation (6.8). The action of the operator T on a function f in terms of wavelets is:

$$Tf(x) = \sum_{n\in\mathbf{Z}} n \sum_{m\in\mathbf{Z}} \langle \psi_{nm}, f\rangle \psi_{nm}(x) \tag{6.14}$$

where $\psi_{n,m}$ are the Haar wavelets. For each n the Haar wavelets $\psi_{n,m}$ are an orthonormal basis of the wavelet subspace

$$\mathcal{W}_n = \mathcal{H}_{n+1} \ominus \mathcal{H}_n. \tag{6.15}$$

By (6.12) the Haar wavelets are also the natural eigenbasis of T. ◻

It should be noted that in the above considerations T is not defined in an ad hoc manner but rather T is determined from the Haar system (and, more generally, from any wavelet multiresolution analysis) in exactly the same way as the time operator is determined in statistical physics [**32**]. For the Haar wavelet the meaning of increasing age is increasing detail refinement as captured by the finer scalings in the Haar wavelet approximations.

It should be remembered [**18**] that Haar's thesis in 1909 was, among other tasks, to construct a complete orthonormal system which did not come from a Sturm-Liouville second order differential equation boundary value problem. Although the time operator is not a Sturm-Liouville differential operator, nonetheless we would like to clarify the relation of the time operator of Haar's system to differential equations.

Theorem 6.3. *The time operator of the Haar system is a* (*position*) *multiplication operator. It is naturally equivalent under appropriate transform to a first order* (*momentum*) *differential operator.*

Proof. (Outline). In the continuous parameter case Time operators may be seen [**3, 16**] to be unitarily equivalent to position operators q in a Schrodinger representation

$$[p, q] = -iI \tag{6.16}$$

of the Weyl commutation relations. In this representation we recall that

$$qf(x) = xf(x), \quad pf(x) = -i\frac{d}{dx}(x) \tag{6.17}$$

are self-adjoint operators with dense domains in $\mathcal{L}^2(\mathbf{R})$. These two operators are interchanged within the canonical commutation relations by

Fourier transform. In the discrete parameter case, T is canonically conjugate to the generator of dilations, $D = ix(d/dx) - \alpha I$. These two operators may be canonically interchanged by Fourier-Mellin transform. $\square$

Time operators were first formulated by Pauli [**33**] for quantum mechanics, but they were rejected by him because their spectrum was the whole real line, which is unacceptable according to the semi-bounded nature of the energy Hamiltonians of quantum mechanics. However, for the Liouville Hamiltonians of statistical mechanics [**32**], this is not an objection. Our own interest in Time operators began in [**20**] where we established a fundamental connection between regular stochastic processes and the canonical commutation relations of quantum mechanics. As Pauli [**33**] knew, once you have the commutation relations, you may define time operators. Regular stationary processes are closely related to Kolmogorov dynamical systems and thus so are time operators [**32, 19**]. From this perspective we were later led to formulate the time operators of all wavelets. See [**2**] for a survey of all of this work.

7. General inverse problem: Structure operators. The above discussion now allows me to formulate the following very general inverse problem: what is the most natural linear operator to a given mathematical structure? I am assuming of course that such an operator is *a priori* not known, i.e., we did not build our mathematical structure originally from the operator. I stated an example of this question in a more limited sense in the introduction to this paper: given a basis $\{\phi_n\}$, when is there a "natural" associated linear operator A? The point is that A should also incorporate all the mathematical structure which led to or underlies the $\{\phi_n\}$, in this example.

Let me elaborate how the Time operator embodies 'naturally' virtually every property of a wavelet. Properties (1)–(3) correspond to its spectral projections. In particular the wavelet subspaces $\mathcal{W}_n = \mathcal{H}_{n+1} \ominus \mathcal{H}_n$ are its age eigenspaces. The very important dilation Property (4) corresponds to increasing time, which for wavelets means increasing detail discernment. Moreover the dilation group unitary representation V may be seen to be the exponentiation of a momentum operator, to which T is Fourier-equivalent. Property (5) through T spreads the scale (time) refinement of Property (4) to all parts of the next subspace.

Thus the new inverse problem I pose goes well beyond just asking if a basis and its moments generate an operator. That would be one possibility, but we may entertain also the possibility that other fundamental mathematical structures may generate an operator, or more importantly, those structures should necessarily be represented by any operator which answers the inverse problem question. For example, in addition to the wavelet multi-resolution properties (1) through (5), the time operator necessarily had to satisfy the group-theoretic imprimitivity conditions represented by canonical commutation relations, e.g., properties (6.5) and (6.7).

Within the context of continued fractions, orthogonal functions, and rational approximations, what are the important mathematical structures that one would want to select or discern as necessary to determine a linear operator A which 'naturally' represents them? Starting with a set of orthogonal polynomials $\{\phi_n\}$, if we regard those as being eigenfunctions of a candidate operator A, we also will need, either directly or indirectly, the corresponding eigenvalues, e.g., the spectral frequencies A would represent. For many of the known classical polynomials, e.g. those of Fourier, Legendre, Jacobi and so on, the 'natural' operator A is already known: it is the second order Sturm-Liouville differential operator L which supplies those eigenfrequencies [**18**]. If one wants to restrict the candidate $\{A\}$ class to bounded operators, one could arrive at $A = L^{-1}$ via L's integral equation inverse. So we should assume that the $\{\phi_n\}$ we started with do not come from any such known operators. In addition to the $\{\phi_n\}$, or even independently of them, what structures in a continued fraction could determine a 'natural' operator A? Remember that the Stieltjes transforms, as a truncated continued fraction, became a rational approximation which in the limit converged to a singular integral operator which may be regarded as an operator resolvent $(A - \lambda I)^{-1}$. Another hint could come through the use of the three term recurrence relations important and common to the rational partial sums of continued fractions, Sturm-Liouville theory, and tridiagonal, e.g., Jacobi, matrices. Also one may want to require some analyticity properties of A if one starts from analytic continued fractions. I mention that such analytic properties may exhibit themselves in an indirect way. For example, many Jacobi and Toeplitz operators A have absolutely continuous spectrum, e.g., filling out the unit disk.

We may regard the latter as a manifestation of an underlying analytic structure.

8. Operator of a continued fraction. Here I want to partially answer the inverse problem for structure operator A for continued fractions. That is, the given structure will be a continued fraction only. I will lean heavily on the results of Stone [**37**]. As I indicated in Section 4, much of the needed theory is already in [**37**]. In view of what I have already presented in Section 4, it will be no surprise that I will conclude here that the natural structure operators A for continued fractions are Jacobi matrices. However, there are a few interesting "wrinkles" to this correspondence which I want to mention here. Also I have not fully exploited some of the "associated" structure details, such as those of the associated polynomial or Padé structures. Perhaps I or others will do so elsewhere.

From what was already presented in Section 4, we know that self-adjoint operators H with simple spectrum on a separable Hilbert space H may be unitarily associated with Jacobi matrices A as in (4.5). Moreover by [**37**, Theorem 10.23] each reduced Jacobi matrix A_p has associated with it the rational function

$$I(z;\rho_p) = \int_{-\infty}^{\infty} \frac{1}{\lambda - z}\, d\rho_p(\lambda) \tag{8.1}$$

which can be expressed as the continued fraction

$$I_p(z) = \frac{1\quad|}{|a_1 - z} - \frac{|b_1|^2|}{|a_2 - z} - \cdots - \frac{|b_{p-1}|^2|}{|a_p - z}. \tag{8.2}$$

As I outlined in Section 4, Stone [**37**] used the reduced Jacobi matrices, and hence, implicitly, reduced continued fractions (8.2), to develop the spectral theory for arbitrary self-adjoint operators H with simple spectra, and hence, implicitly, for arbitrary associated Jacobi matrices A with simple spectra. The step to infinite continued fractions is made at [**37**, p. 559]. For any given Jacobi matrix A of form (4.5) we may associate the (Jacobi) continued fraction

$$I(z) = \frac{|b_0|^2|}{|a_1 - z} - \frac{|b_1|^2|}{|a_2 - z} - \cdots - \frac{|b_n|^2|}{|a_{n+1} - z} - \cdots \tag{8.3}$$

where $b_0 = 1$. But notice that these continued fractions are not one-to-one with the Jacobi matrices. There are an infinite number of Jacobi matrices to choose from if one wants one of them to be the structure operator A representing the continued fraction in the sense of the general inverse problem I posed in the preceding section. Of course one could just specify some prescription such as requiring the b_i to be real and positive. One could now bring in the known one to one relationship between Jacobi continued fractions and monic orthogonal polynomials, e.g., see [**10**, pp. 85–86], and moment functionals, e.g., see [**10**, pp. 19–21]. One could recast the given continued fraction in terms of its Padé rational approximants and then use the matrix formulation of the latter [**5**, p. 131] to bring in Nuttal's formula and Rutishauser's continued fractions. I would rather not go into all of that here so I would just conclude this paragraph by the assertion: given a continued fraction of the form (8.3), a natural induced structure operator for it is a Jacobi matrix, which is determined "up to phase" by the continued fraction coefficients.

In particular, then, here we do have another example of an induced structure operator for the general inverse problem posed in Section 7. It is not a very rich example for the simple reason that the inducing continued fraction is by itself not a very rich object.

There is another "wrinkle," no doubt related, which I want to bring out. It is also of interest in its own right. In Section 4 I showed that all of Stone's constructions depended on the existence of a cyclic vector f. In particular, one can choose f so that the polynomials $G_n(\lambda)$ are of degree $n-1$ and satisfy the recursion of (4.6). How did Stone do that? The details take us back to [**37**, pp. 282–285]. Stone approximates bounded continuous functions on an infinite interval (he is working in the spectral representation space $\mathcal{L}^2(\rho)$ now) by functions of the form

$$f(\lambda) = e^{-\alpha\lambda^2} P(\lambda) \tag{8.4}$$

where $P(\lambda)$ is a polynomial. In other words, he is approximating in terms of Hermite polynomials. His constructions at this key point are specifically predicated upon his earlier work [**37**, p. 284, footnote to an Annals of Mathematics paper]. The existence of a cyclic vector g is guaranteed by the simple spectrum hypothesis but is otherwise left arbitrary. Then an improved cyclic vector f is constructed by means of the choice $f = e^{-H^2} g$. This guarantees that f is in the operator domain core $\cap_n \mathcal{D}(H^n)$. The spectral representation correspondence is

$$\lambda^n e^{-\lambda^2} \leftrightarrow H^n e^{-H^2}. \tag{8.5}$$

The important orthonormal sets $\{g_n\}$ and $\{G_n\}$ are then obtained, and I quote directly [**37**, p. 285]:

> From the sequence $\{H^n f\}$ we now form an orthonormal set $\{g_n\}$, $g_n = G_n(H)f$, by means of the process described in Theorem 1.13. When the details of that process are scrutinized with regard to this application, it is found that $G_n(\lambda)$ is a polynomial of degree $n-1$. The orthonormal set $\{g_n\}$ is evidently complete.

The reference to Theorem 1.13 is just to orthonormalization. However, in this key passage Stone is being quite terse. But it is clear that his Jacobi matrix obtained in Theorem 7.13 was obtained in terms of a Hermite polynomial approximation starting from an arbitrary cyclic vector.

As a third "wrinkle" of Stone's approach [**37**] to spectral theory, it is interesting that he chose to use Jacobi matrices (tridiagonal) in his spectral representation of H in the case of continuous (simple) spectra. See [**37**, Theorem 7.17, p. 295] in his treatment of reduction to principal axes. That is, although he also obtains the "fully diagonal" principal axes spectral representation theorem [**37**, Theorem 7.18] favored by current books on spectral theory, he is willing to accept a tridiagonal "principal axis" canonical form in order to get explicit matrix representation of H. It is because of that choice that I am able to make the connections presented in this paper.

Finally, as mentioned above, by itself a continued fraction or a Jacobi matrix is a relatively barren object. It is where they come from that matters. For example, knowing the physical or mathematical moments that gave the former, or the physical or mathematical polynomials that gave the latter, is more informative and therefore more interesting. Best would be to know the unique originating physical or mathematical self-adjoint operator H. But even that is not the full story. As I have shown in this section, the essence is in the particular cyclic vector. All five of the roughly one-to-one related mathematical structures, namely, continued fractions, orthogonal functions, rational approximants, moments, and self-adjoint operators with simple spectra, become more than just "unitarily" equivalent when you are able to compute from a specific cyclic vector. That was the approach taken by Stone. It is also the approach of wavelets.

9. Principal conclusions. My original question "what is the spectral theory (of continued fractions)" led me, after considerable literature perusal, to the treatment by Stone [**37**]. I have long known this important early exposition of general operator theory, but I had never paid attention to Stone's last section Section 4 of his last Chapter X, which comprises pages 530–614 of the book. His goal stated there (page 530) is to "study ... the most general Jacobi matrix ... also obtain the means for solving important problems in the theory of continued fractions and the theory of moments." I have analyzed his contributions in Section 4 and in Section 8. As I concluded there, Stone's principal goal was to establish that every self-adjoint operator H with simple spectrum could be represented by a Jacobi matrix operator A. His second goal in this regard was to develop the spectral theory of general Jacobi matrices A even in the case that the associated symmetric operator H is only essentially self-adjoint or more generally just a symmetric operator with equal defect indices. His discussion of the moment problem is only minor. His interest in continued fractions is chiefly to use them as a convergence tool when treating singular integrals, as Stieltjes did.

Secondly, I have posed a new general inverse problem, which I may state here as that of finding the induced structure operator. This operator A is to naturally represent all provided spectral information and furthermore also should represent insofar as possible the key mathematical or physical structures which underlie that spectral data. Our recent result [**3**] for the time operator as the natural induced structure operator of any wavelet system was given as an example in Section 6. In Section 8 I interpret Jacobi matrices as associated by Stone [**37**] to a continued fraction as another example of natural induced structure operator.

Acknowledgments. I would like to thank Olaf Njåstad for his interest and useful suggestions for the improvement of the preliminary draft of this paper which I circulated at the conference. Also Professor

Njåstad pointed out to me two related papers [**22, 23**] in which self-adjoint operator theory is applied to the strong Hamburger moment problem. In [**23**] Stone's criteria [**37**] for self-adjointness of operators generated by Jacobi matrices is extended to Laurent-Jacobi matrices.

In Section 3 I was overly brief in describing the classical view of spectrum $\sigma(\alpha)$ as the set of points of increase of a real valued distribution function. I am indebted to a referee for bringing to my attention the following literature. In [**41, 36**] one finds the more general perspective of spectrum in terms of measures and their support, real or complex. In [**31, 7**] orthogonality and moment problems are discussed in a more general framework, including Sobolev orthogonal polynomials and Riesz bases. In [**12, 24, 34**] classical problems are discussed from an operator theoretic perspective.

REFERENCES

1. N.I. Akhiezer and M.G. Krein, *Some questions in the theory of moments*, Amer. Math. Soc., Providence, RI, 1962.

2. I. Antoniou and K. Gustafson, *Wavelets and stochastic processes*, Math. Comput. Simulation **49** (1999), 81–104.

3. ———, *The Time operator of wavelets*, Chaos, Solitons Fractals **11** (2000), 443–452.

4.R. Askey, *Gabor Szegö*: *Collected papers*, Vol. 3, Birkhäuser, Boston, (1982), 866–869.

5. C. Brezinski, *Padé-type approximation and general orthogonal polynomials*, Birkhäuser, Basel, 1980.

6. ———, *History of continued fractions and Padé approximants*, Springer, Berlin, 1990.

7. A. Bultheel, P. Gonzalez-Vera, E. Hendriksen and O. Njåstad, *Orthogonal rational functions*, Cambridge Univ. Press, Cambridge, 1999.

8. K.M. Case and M. Kac, *A discrete version of the inverse scattering problem*, J. Math. Phys. **14** (1973), 594–603.

9. K. Chadan and P.C. Sabatier, *Inverse problems in quantum scattering theory*, Springer, New York, 1977.

10. T.S. Chihara, *An introduction to orthogonal polynomials*, Gordon and Breach, London, 1978.

11. C.K. Chui, *An introduction to wavelets*, Academic Press, Boston, 1992.

12. J. Dombrowski, *Orthogonal polynomials and functional analysis*, in *Orthogonal polynomials*: *Theory and practice* (P. Nevai, ed.), Kluwer Academic, Dordrecht, (1990), 147–160.

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

14. C. Gordon, D. Webb and S. Wolpert, *Isospectral plane domains and surfaces via Riemannian orbifolds*, Invent. Math. **110** (1992), 1–22.

15. W. Gragg, *Matrix interpretations and applications of the continued fraction algorithm*, Rocky Mountain J. Math. **4** (1974), 213–225.

16. K. Gustafson, *Lectures on computational fluid dynamics, mathematical physics, and linear algebra*, World Scientific, Singapore, 1997.

17. ———, *Operator spectral states*, Computers Math. Appl. **34** (1997), 467–508.

18. ———, *Partial differential equations and Hilbert space methods*, Dover Publ., New York, 1999.

19. K. Gustafson and K. Goodrich, *Kolmogorov systems and Haar systems*, Colloq. Math. Soc. János Bolyai, vol. 49, North-Holland, Amsterdam, 1985, pp. 401–416.

20. K. Gustafson and B. Misra, *Canonical commutation relations of quantum mechanics and stochastic regularity*, Lett. Math. Phys. **1** (1976), 275–280.

21. K. Gustafson and D. Rao, *Numerical range*, Springer, Berlin, 1997.

22. E. Hendriksen, *The strong Hamburger moment problem and self-adjoint operators in Hilbert space*, J. Comput. Appl. Math. **19** (1987), 79–88.

23. E. Hendriksen and C. Nijhuis, *Laurent-Jacobi matrices and the strong Hamburger moment problem*, Acta Appl. Math. **61** (2000), 119–132.

24. E.K. Ifantis and P.N. Panagopoulos, *Limit points of eigenvalues of truncated tridiagonal operators*, J. Comput. Appl. Math. **133** (2001), 413–422.

25. W.B. Jones and W.J. Thron, *Continued fractions*, Addison-Wesley, Reading, MA, 1980.

26. ———, *Orthogonal Laurent polynomials and Gaussian quadrature*, in *Quantum mechanics in mathematics, chemistry, and physics* (K. Gustafson and W.P. Reinhardt, eds.), Plenum Press, New York, (1981), 449–455.

27.W.B. Jones and O. Njåstad, *Orthogonal Laurent polynomials and strong moment theory*: *A survey*, J. Comput. Appl. Math. **105** (1999), 51–99.

28. M. Kac, *Can one hear the shape of a drum?*, Amer. Math. Monthly **73** (1966), 1–23.

29. G. Kaiser, *A friendly guide to wavelets*, Birkhäuser, Boston, 1994.

30. P. Lax and R. Phillips, *Scattering theory*, Academic Press, New York, 1967.

31. F. Marcellan, M. Alfaro and M. Rezola, *Orthogonal polynomials on Sobolev spaces: old and new directions*, J. Comput. Appl. Math. **48** (1993), 113–131.

32. B. Misra, I. Prigogine and M. Courbage, *From deterministic dynamics to probabilistic descriptions*, Phys. A **98** (1979), 1–26.

33. W. Pauli, *Die allgemeinen Prinzipien der Wellenmechanik*, in *Handbuch der Physik* (S. Flugge, ed.), Springer, Berlin, 1958, pp. 1–168.

34. D. Sarason, *Moment problems and operators in Hilbert space*, in *Moments in mathematics* (H.J. Landau, ed.), Amer. Math. Soc., Providence, RI, 1987, pp. 54–70.

35. A. Sri Ranga, *On a recurrence formula associated with strong distributions*, SIAM J. Math. Anal. **21** (1990), 1135–1148.

36. H. Stahl and V. Totik, *General orthogonal polynomials*, Cambridge Univ. Press, Cambridge, 1992.

37. M.H. Stone, *Linear transformations in Hilbert space*, Amer. Math. Soc., Providence, RI, 1932.

38. G. Strang and T. Nguyen, *Wavelets and filter banks*, Wellesley-Cambridge Press, Wellesley, MA, 1996.

39. G. Szegö, *Orthogonal polynomials*, Amer. Math. Soc., Providence, RI, 1975.

40. ———, *An outline of the history of orthogonal polynomials*, in *Orthogonal expansions and their continuous analogues* (D. Haimo, ed.), Feffer and Simons, London, 1967, pp. 3–11.

41. W. Van Assche, *Asymptotics for orthogonal polynomials*, Springer, Berlin, 1987.

42. E.T. Whittaker and G.N. Watson, *A course of modern analysis*, Cambridge Univ. Press, Cambridge, 1927.

Department of Mathematics, University of Colorado, Boulder, CO 80309
E-mail address: gustafs@euclid.colorado.edu