GENERALISED FOURIER AND TOEPLITZ RESULTS FOR RATIONAL ORTHONORMAL BASES

BRETT NINNESS*, HåKAN HJALMARSSON†, AND FREDRIK GUSTAFSSON‡

Abstract. This paper provides a generalisation of certain classical Fourier convergence and asymptotic Toeplitz matrix properties to the case where the underlying orthonormal basis is not the conventional trigonometric one, but a rational generalisation which encompasses the trigonometric one as a special case. These generalised Fourier and Toeplitz results have particular application in dynamic system estimation theory. Specifically, the results allow a unified treatment of the accuracy of least squares system estimation using a range of model structures, including those that allow the injection of prior knowledge of system dynamics via the specification of fixed pole or zero locations.

AMS subject classifications. 42C15, 93E12, 47B35, 93E24, 60G35

Key words. stochastic processes, prediction theory, system identification.

1. Introduction. In the area of applied mathematics, a fundamental idea is that of approximating or exactly expressing solutions by expanding them in terms of orthogonal basis functions. Well known classical examples are Fourier analysis, solutions of the wave equation and Schrödingers equation in terms of (respectively) Legendre and Laguerre orthogonal polynomials, and solutions of self-adjoint operator equations such as Sturm-Liouville systems in terms of the orthogonal eigenfunctions of the operator. More recently, particularly for the solution of signal processing and other system theoretic problems there has been an explosion of interest in the development and use of a wide class of new orthogonal bases called 'Wavelets' [6, 4].

Indeed, tackling system theoretic problems using orthonormal descriptions has a particularly rich history, going back at least as far as the work of Kolmogorov [20] and Wiener [45] who exploited them in developing their now famous theory on the prediction of random processes. In that work, the orthonormal basis was the trigonometric one, but as was shown by Szegö there is great utility in re-expressing the problem with respect to another orthonormal basis that is adapted to the random process; namely a basis of polynomials orthogonal to a given positive function f which is the spectral density of the process [39, 10]. Such polynomials are called 'Szegö polynomials'.

This latter approach derives its utility from the fact that the n'th order Szegö polynomial is in fact the mean-square best order n one step ahead predictor of the random process [12, 39]. By exploiting the orthonormality of the basis to derive what is called a 'Christoffel–Darboux' formula for the 'Reproducing Kernel' associated with the Szegö polynomial basis, theoretical analysis of this predictor is greatly facilitated.

* This work was supported by the Australian Research Council and the Centre for Integrated Dynamics and Control (CIDAC), The University of Newcastle, N.S.W. 2308, Australia. This author is with CIDAC and the Department of Electrical & Computer Engineering, The University of Newcastle, Australia. email:brett@ee.newcastle.edu.au or FAX: +61 2 49 21 69 93. Part of this work was completed while visiting the Department of Sensors, Signals and Systems-Automatic Control, The Royal Institute of Technology, Stockholm, Sweden.

† Work partly completed while on leave at CIDAC from the Department of Sensors, Signals and Systems-Signal Processing, The Royal Institute of Technology, S-100 44 Stockholm, Sweden. email:hakan.hjalmarsson@s3.e.kth.se or FAX: +46 8 790 7329

‡ Work partly completed while on leave at CIDAC from the Department of Electrical Engineering, Linköping University, S-581 83, Linköping, Sweden. email:fredrik@isy.liu.se or FAX: +46 013 282622

For example, it was by this means that Szegö was able to derive his famous formula

$$\sigma^2 = \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log f(\omega)\, d\omega\right\}$$

for the asymptotic in order n variance σ^2 of the prediction error associated with a spectral density f.

As well, use of the Christoffel–Darboux formula provides a recursive in n formula for the Szegö polynomials [39, 10], and this in turn allows a computationally efficient means for calculating predictors. This recursive formula is of course the famous Levinson recursion, which was developed independently of Szegö's work by exploiting the properties of Toeplitz matrices [22, 33]. In practice, the so–called 'reflection coefficients' required in the Levinson recursions are calculated by the Schur algorithm [36], originally proposed by Schur [37] as a means for testing whether or not a function is bounded positive real (or 'Caratheodory' as it is known in some literature). Here again orthonormal bases and Toeplitz matrices arise since another test for positive realness involves testing for the positive definiteness of the Toeplitz matrix formed from the Fourier co-efficients of the function [38].

These several links between Toeplitz matrices and orthonormal bases arise since (subject to some regularity conditions) the ℓ,m'th element of any $n \times n$ symmetric Toeplitz matrix may be denoted as $T_n(f)$ and expressed using the orthonormal trigonometric basis $\{e^{j\omega n}\}$ as

$$[T_n(f)]_{\ell,m} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\omega\ell}e^{-j\omega m} f(\omega)\, d\omega \tag{1.1}$$

for some positive function f. By recognising this, certain quadratic forms of Toeplitz matrices that arise naturally in the frequency domain analysis of least-squares estimation problems may instead be conveniently rewritten as

$$\frac{1}{n}\Gamma_n^\star(\omega)T_n(f)\Gamma_n(\omega) = \sum_{k=-n}^{n}\left(1-\frac{|k|}{n}\right) c_k\, e^{j\omega k} \tag{1.2}$$

where $\cdot^\star$ denotes 'conjugate transpose' and,

$$c_k \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\omega)e^{-j\omega k}\, d\omega$$

is the k'th Fourier co-efficient of f with $\Gamma_n(\omega)$ an $n \times 1$ vector defined as

$$\Gamma_n^\star(\omega) \triangleq [1, e^{-j\omega}, e^{-j2\omega}, \cdots, e^{-j(n-1)\omega}].$$

The right hand side of (1.2) may be recognised as the Cesàro mean reconstruction of a Fourier series which is known [9], provided f is continuous, to converge uniformly to $f(\omega)$ on $[-\pi, \pi]$.

This latter fact has been exploited by Ljung and co-workers [25, 27, 13, 26, 46, 23] who, reminiscent of Szegö's approach of examining the asymptotic in order n nature of predictors, have provided asymptotic in model order results describing the variability of the frequency response of least-squares system estimates in such a way as to elucidate how they depend on excitation and measurement noise spectral densities, model order, and observed data length; see [32] for more detail on this point.

Such results have found wide engineering application; see for example [2, 11, 24, 14]. However, to derive them, another key ingredient pertaining to the properties of Toeplitz matrices is required. Namely, that asymptotically in size n, Toeplitz matrices posses the algebraic structure [12, 44]

$$T_n(f)T_n(g) \sim T_n(fg) \tag{1.3}$$

where f and g are any continuous positive functions, and for $n \times n$ matrices A_n and B_n, the notation $A_n \sim B_n$ means that $\lim_{n\to\infty} |A_n - B_n| = 0$ where $|\cdot|$ is the Hilbert–Schmidt matrix norm defined by

$$|A|^2 \triangleq \frac{1}{n}\mathrm{Trace}\{A^\star A\}. \tag{1.4}$$

The main results of this paper are to extend the results of the convergence of the Cesàro mean (1.2) and the algebraic structure of Toeplitz matrices (1.3) to more general cases wherein the underlying orthonormal basis is not the trigonometric one, but a generalisation of it. More specifically, this paper studies the use of the basis functions $\mathcal{B}_n(z)$ given by

$$\mathcal{B}_n(z) \triangleq \frac{\sqrt{1-|\xi_n|^2}}{1-\xi_n z}\prod_{k=0}^{n-1}\left(\frac{z-\overline{\xi_k}}{1-\xi_k z}\right) \tag{1.5}$$

where the $\{\xi_k\}$ may be chosen (almost) arbitrarily inside and (in some cases) on the boundary of the open unit disc $\mathbf{D} \triangleq \{z \in \mathbf{C} : |z| < 1\}$ ($\mathbf{C}$ is the field of complex numbers). These functions $\{\mathcal{B}_n\}$ are orthonormal on the unit circle $\mathbf{T} = \{z \in \mathbf{C} : |z| = 1\}$, and the trigonometric basis is a special case of them if all the $\{\xi_k\}$ are chosen as zero. Using them, a generalisation

$$[M_n(f)]_{\ell,m} \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathcal{B}_\ell(e^{j\omega})\overline{\mathcal{B}_m(e^{j\omega})}f(\omega)\,\mathrm{d}\omega \tag{1.6}$$

of Toeplitz matrices is considered, for which it is shown here that a generalisation of (1.3) still holds, and with the redefinition

$$\Gamma_n^T(\omega) \triangleq [\mathcal{B}_0(e^{j\omega}),\ \mathcal{B}_1(e^{j\omega}),\ \dots,\mathcal{B}_{n-1}(e^{j\omega})] \tag{1.7}$$

it is also shown here that a generalisation of the uniform convergence of the Cesàro mean (1.2) to $f(\omega)$ also holds.

In both cases, the generalisation involves replacing the $1/n$ normalisation appearing in (1.2) and in the definition of the matrix norm (1.4) with a frequency dependent term $K_n(\omega,\omega)$ which is the reproducing kernel associated with the linear space spanned by the basis functions $\{\mathcal{B}_0, \mathcal{B}_1, \cdots, \mathcal{B}_{n-1}\}$.

Indeed, this reproducing kernel is the key to the results presented here. Classical derivations of Cesàro summability and Toeplitz matrix results rely heavily on the algebraic structure of the trigonometric basis. Namely, that $e^{j\omega n}e^{j\omega m} = e^{j\omega(n+m)}$. In the cases considered here, since $\mathcal{B}_n\mathcal{B}_m \neq \mathcal{B}_{n+m}$ this algebraic structure is lost and pre-existing analysis techniques are not applicable. Instead, motivated by Szegö's approach to the study of orthogonal polynomials, this paper exploits a closed form expression for the appropriate reproducing kernel.

The utility of the new results presented here is that just as the classical Fourier and Toeplitz results have been used by Ljung and co-workers to analyse estimation using finite impulse response (FIR) and certain other rational model structures, the results of this paper can be used to analyse estimation using generalisations of these model structures. As shown in [32] and as commented on in [32], these generalised structures are actually quite common since they implicitly arise whenever the common practice of data pre-filtering is performed.

There is much other work related to the results presented here. The study of the basis functions (1.5) dates back to Malmquist [29] and was taken up by Walsh [43] in the context of complex rational approximation theory, and by other workers [45, 21, 18, 3, 8, 7, 28, 34] for system theoretic applications such as system approximation and network synthesis, including generalisations of Schur and Levinson recursions, Lattice structures, and the concomitant solution of inverse scattering problems.

In the context of system identification, as well as pertaining to the aforementioned work [25, 27, 13, 26, 46, 23] the results of this paper also have close connections with much recent literature examining the use of model structures derived from orthonormal bases. In [19, 5, 17, 40, 42] the use of the so-called 'Laguerre' basis is examined. This basis can be encompassed by the basis (1.5) by fixing all the poles at a common value $\xi_k = \xi \in \mathbb{R}$, ($\mathbb{R}$ denotes the field of real numbers) and with the substitution $z \mapsto 1/z$ so as to accommodate convention in the signal processing and control theory literature. In this case the name 'Laguerre' derives from the ensuing functions being related to the classical Laguerre orthonormal polynomials via a Fourier and bilinear transform [30]. In [41] a generalisation of this Laguerre case is analysed wherein the common value ξ may be complex valued. In [15, 35], these analyses are again generalised to the case where a fixed set of poles $\{\xi_0, \cdots, \xi_r\}$ are cyclically repeated and orthonormal bases are generated with denominators given as $D_p(z) = \prod_{k=0}^{p-1}(z - \xi_k)$ and numerators the Szegö polynomials associated with the weight function $|D_p(e^{j\omega})|^{-2}$. The cyclic repetition of poles arises due to the latter numerator and denominator pair being multiplied by powers of the all-pass function $z^p D_p(1/z)/D_p(z)$ as the number of required basis functions increases beyond p.

In all these works, any analysis of estimation accuracy proceeds by exploiting the restriction on the choice of ξ_k to establish, via a bilinear transform [40, 41, 42], or a multilinear transform (dubbed a 'Hambo' transform) [35] an algebra isomorphism to the trigonometric basis $\{e^{j\omega n}\}$. The utility of this is that the original results of Ljung [27] can then be employed, having been mapped through the isomorphism, to provide quantification of estimation accuracy.

In spite of the elegance of this approach, it suffers several drawbacks which are the motivation for the work at hand. Firstly, the results pertain only to a restricted class of models in which either all the poles $\{\xi_k\}$ are chosen the same [40, 42, 41], or are cyclically repeated from a fixed set [35]. Secondly, and with particular reference to [35], the results are asymptotic not as is the case here to the number of poles $\{\xi_k\}$ chosen, but to the number of times the whole set $\{\xi_0, \cdots, \xi_{p-1}\}$ is repeated. The results in this paper allow the avoidance of these limitation by eschewing a strategy of forcing an algebra isomorphism to the trigonometric case.

The presentation of these ideas is organised as follows. In § 2 following, the analysis begins by establishing that the general orthonormal bases (1.5) fundamental to this paper form a complete set in the Hilbert space $H_2(\mathbf{T})$. In order to study other approximating properties of the basis, a 'Reproducing Kernel' approach is employed, and § 3 is devoted to explaining certain important principles relevant to this framework. Perhaps more importantly, § 3 also contains the derivation of a closed form 'Christoffel–Darboux' type formula for the reproducing kernel. With these results in hand, § 4 then considers generalised Fourier analysis with respect to the basis (1.5), and using the reproducing kernel ideas establishes uniform convergence for generalised Cesàro mean reconstructions.

In fact, because of application demands, something more is derived in that it is shown that for certain frequencies being different, then uniform convergence to zero also ensues. The generalised Cesàro mean reconstruction is defined with respect to a generalised Toeplitz matrix, and § 5 is devoted to the study of the asymptotic algebraic properties of such matrices since as already explained, these properties are of great utility in certain system theoretic applications.

Pertinent to this, § 5 defines a new notion of asymptotic equivalence between matrices, and then uses this to establish that asymptotically, arbitrary products of generalised Toeplitz matrices and their inverses are equivalent to a single generalised Toeplitz matrix with symbol equal to the product of the corresponding symbols and inverse symbols of the matrices in the product. Finally, § 6 provides a summary and concluding perspectives on the work presented here.

2. Completeness Properties. The theme of this paper is to examine certain system theoretic issues pertaining to the use of the basis functions (1.5) for the purposes of describing discrete time dynamic systems. In the sequel only bounded-input, bounded-output stable and causal systems will be of interest, so that it is natural to embed the analysis in the Hardy space $H_2(\mathbf{T})$ of functions $f(z)$ which are analytic on $\mathbf{D}$, square integrable on $\mathbf{T}$, and posses only a one-sided Fourier expansion. As is well known [16], $H_2(\mathbf{T})$ is a Hilbert space when endowed with the inner product

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{j\omega})\overline{g(e^{j\omega})}\,\mathrm{d}\omega = \frac{1}{2\pi j}\oint_{\mathbf{T}} f(z)\overline{g(z)}\,\frac{\mathrm{d}z}{z}, \quad f, g \in H_2(\mathbf{T}). \tag{2.1}$$

That the functions (1.5) form an orthonormal set in that $\langle \mathcal{B}_n, \mathcal{B}_m\rangle = \delta(n-m) =$ Kronecker delta may easily be shown [31] using the contour integral formulation of the inner product in (2.1) and Cauchy's residue Theorem.

What must be of central interest if the functions (1.5) are to be useful in such a system theoretic setting is whether or not linear combinations of them can describe an arbitrary system in $H_2(\mathbf{T})$ to any degree of accuracy. This may be answered in the affirmative by the following closure result.

Theorem 2.1.

$$\overline{\operatorname{Span}\{\mathcal{B}_k(z)\}} = H_2(\mathbf{T})$$

if and only if

$$\sum_{k=0}^{\infty}(1-|\xi_k|) = \infty$$

where here $\overline{X}$ denotes the norm closure of the space X.

3. Reproducing Kernels. Given the completeness result in Theorem 2.1, to further examine the properties of approximants formed as linear combinations of the basis functions (1.5), this paper takes the approach of utilising the ideas of reproducing kernel spaces [1, 39]. Key to this approach is the availability of the following

"Christoffel–Darboux' like closed form formulation of the Reproducing Kernel.

Theorem 3.1. Christoffel–Darboux Formula: *Define the modified Blaschke product*

$$\varphi_n(z) \triangleq \prod_{k=0}^{n-1} \frac{z - \overline{\xi_k}}{1 - \xi_k z}.$$

Then the Reproducing Kernel of the orthonormal system $\{\mathcal{B}_k\}$ can be expressed as

$$K_n(z, \mu) = \frac{1 - \overline{\varphi_n(\mu)}\varphi_n(z)}{1 - z\overline{\mu}}. \tag{3.1}$$

4. Generalised Fourier Series Convergence. This section considers the convergence of generalised Fourier series approximants formed using the general orthonormal basis (1.5). For this purpose it is necessary to first develop a generalisation of the Cesàro mean. This may be accomplished by the definition

$$f_n(\omega) \triangleq \frac{\Gamma_n^\star(\omega) M_n(f) \Gamma_n(\omega)}{K_n(\omega, \omega)} \tag{4.1}$$

where Γ_n defined in (1.7) is an $n \times 1$ vector of general rational orthornormal basis functions (1.5) and $M_n(f)$ is a generalised Toeplitz matrix as defined in (1.6). If all the poles $\{\xi_k\}$ are set to zero, then it is straightforward to verify that the formulation (4.1) reduces (since in this case $M_n(f) = T_n(f)$) to the usual Cesàro mean. The convergence properties of (4.1) are provided by the following theorem.

Theorem 4.1. *Suppose $f(\omega)$ is a continuous not necessarily real-valued function on $[-\pi, \pi]$. Then provided*

$$\sum_{k=0}^{\infty} (1 - |\xi_k|) = \infty$$

the following limit result holds

$$\lim_{n \to \infty} \frac{\Gamma_n^\star(\omega) M_n(f) \Gamma_n(\omega)}{K_n(\omega, \omega)} = f(\omega)$$

uniformly in ω on $[-\pi, \pi]$. Under the strengthened condition that $|\xi_n| \le 1 - \delta$ for some $\delta > 0$ and all n, then for $\mu \neq \omega$

$$\lim_{n \to \infty} \frac{\Gamma_n^\star(\mu) M_n(f) \Gamma_n(\omega)}{K_n(\omega, \omega)} = 0.$$

5. Algebraic Structure of Generalised Toeplitz Matrices. In applications [25, 27, 13, 26, 46, 23], the consideration of quadratic forms more complicated than (4.1) occur. In fact, what is of more interest are forms such as

$$\frac{\Gamma_n^\star(\omega) M_n(f) M_n(g) \Gamma_n(\omega)}{K_n(\omega, \omega)}.$$

In these aforementioned applications [25, 27, 13, 26, 46, 23], the underlying orthonormal basis is the trigonometric one $\{e^{j\omega n}\}$ in which case $M_n(f) = T_n(f)$ is a bona-fide Toeplitz matrix for which classical results are at hand concerning their algebraic structure. Namely, following the notation defined in (1.3), the convenient property that $T_n(f)T_n(g) \sim T_n(fg)$ is assured [12, 44] (the meaning of the $\sim$ notation here is as described in conjunction with equation (1.3)).

The purpose of this section is to establish this same algebraic structure for the generalised Toeplitz matrices defined by (1.6), the classical results once again arising as the special case of $\xi_k = 0$ in (1.5), but at a cost of somewhat weakening the definition of equivalence over that discussed in § 1 to one in which two $n \times n$ matrices A_n and B_n are said to be asymptotically equivalent as $n \to \infty$ with notation $A_n \sim B_n$ as $n \to \infty$ if

$$\lim_{n\to\infty} \frac{\Gamma_n^\star(\omega)[A_n - B_n][A_n - B_n]^\star\Gamma_n(\omega)}{K_n(\omega,\omega)} = 0 \quad ;\forall\omega \in [-\pi,\pi].$$

Note that this refinement of the definition of matrix equivalence makes no difference for the system theoretic applications motivating this paper (see [32] for more detail on this point). With this definition in hand, the following result on the algebraic structure of generalised Toeplitz matrices is available.

Theorem 5.1. *Consider two not necessarily real valued functions f and g of which at least one of them is Lipschitz continuous of order $\varepsilon > 0$ and the other one bounded. Suppose that the poles $\{\xi_k\}$ of the basis functions $\{\mathcal{B}_k\}$ in (1.5) satisfy $|\xi_k| \leq 1 - \delta$ for some $\delta > 0$. Then*

$$M_n(f)M_n(g) \sim M_n(fg) \quad \text{as } n \to \infty$$

with convergence rate faster than $O(\log^4 n/n^{\varepsilon/(\varepsilon+2)})$ as $n \to \infty$.

Again, something more than this result is actually required in system theoretic applications where one is often concerned with multiple products that also contain matrix inverses. Such cases may be handled by the following corollary to the preceding result. In what follows, matrix products are to be interpreted in a left-to-right fashion as $\prod_{k=1}^n A_k = A_1A_2\cdots A_n$.

Corollary 5.1. *Suppose that the family of possibly complex valued functions $\{f_k\}_{k=1}^m$ are all Lipschitz continuous of order $\varepsilon > 0$. Suppose that the poles $\{\xi_k\}$ of the basis functions $\{\mathcal{B}_k\}$ in (1.5) satisfy $|\xi_k| \leq 1 - \delta$ for some $\delta > 0$. Then with $\sigma_k = \pm 1$*

$$\prod_{k=1}^m M_n^{\sigma_k}(f_k) \sim M_n\left(\prod_{k=1}^m f_k^{\sigma_k}\right) \quad \text{as } n \to \infty$$

with convergence rate faster than $O(\log^4 n/n^{\varepsilon/(\varepsilon+2)})$ as $n \to \infty$ and provided the functions $\{f_k\}$ are invertible where required by the values of σ_k.

Combining this corollary with Theorem 4.1 then provides a further corollary representing an extension of the generalised Fourier convergence of Theorem 4.1.

Corollary 5.2. *Suppose that the family of possibly complex valued functions* $\{f_k\}_{k=1}^{m}$ *are all Lipschitz continuous of order* $\varepsilon > 0$. *Suppose that the poles* $\{\xi_k\}$ *of the basis functions* $\{\mathcal{B}_k\}$ *in (1.5) satisfy* $|\xi_k| \leq 1 - \delta$ *for some* $\delta > 0$. *Then the following limit result holds*

$$\lim_{n\to\infty} \frac{1}{K_n(\omega,\omega)} \Gamma_n^\star(\mu) \left(\prod_{k=1}^{m} M_n^{\sigma_k}(f_k) \right) \Gamma_m(\omega) = \begin{cases} \displaystyle\prod_{k=1}^{m} f_k^{\sigma_k}(\omega) & \mu = \omega, \\ 0 & \mu \neq \omega \end{cases}$$

for any $\omega \in [-\pi, \pi]$ *and where* $\sigma_k = \pm 1$ *with the functions* $\{f_k\}$ *assumed to be invertible when required by the values of* σ_k.

6. Conclusion. The purpose of this paper was to consider certain results in the study of Fourier series and Toeplitz matrices that have proved to be key to various system theoretic applications, and expand them to the case where the underlying orthonormal basis is not the classical trigonometric one, but a rational formulation that encompasses the trigonometric basis as a special case. These results, and the ensuing generalisations developed in this paper are summarized in Table 6.1.

One point worth clarifying, is that in system theoretic settings for which these results will be applicable (control, signal processing, system identification) it is more common to associate the complex variable z with a forward time shift, rather than the backward shift association used here. This discrepancy is easily accommodated by simply transforming $z \mapsto 1/z$ in all the results presented here. A different basis function definition will result, which is in accordance with certain so-called Laguerre and Kautz bases studied in the control theory literature. However, the matrices $M_n(f)$ and the associated Fourier reconstruction formulas will be unchanged.

	Classical	Generalised
Basis	$e^{j\omega n}$	$\mathcal{B}_n(e^{j\omega}) \triangleq \dfrac{\sqrt{1-\|\xi_n\|^2}}{1-\xi_n e^{j\omega}} \prod_{k=0}^{n-1} \left(\dfrac{e^{j\omega}-\overline{\xi_k}}{1-\xi_k e^{j\omega}}\right)$
Completeness	$H_2(\mathbf{T})$	$H_2(\mathbf{T})$ provided $\sum(1-\|\xi_n\|)=\infty$
Associated Matrix	Toeplitz Matrix $[T_n(f)]_{k,\ell} = \dfrac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\omega(k-\ell)} f(\omega)\, \mathrm{d}\omega$	Generalised Toeplitz matrix $[M_n(f)]_{k,\ell} = \dfrac{1}{2\pi}\int_{-\pi}^{\pi} \mathcal{B}_k(e^{j\omega})\overline{\mathcal{B}_\ell(e^{j\omega})} f(\omega)\, \mathrm{d}\omega$
Cesàro Mean	$f_n(\omega) = \dfrac{1}{n}\sum_{k,\ell=0}^{n-1} e^{j\omega(\ell-k)}[T_n(f)]_{k,\ell}$	$f_n(\omega) = \sum_{k,\ell=0}^{n-1} \dfrac{\mathcal{B}_k(e^{j\omega})\overline{\mathcal{B}_\ell(e^{j\omega})}}{K_n(\omega,\omega)}[M_n(f)]_{k,\ell}$
Convergence	$\lim_{n\to\infty} \sup_{\omega\in[-\pi,\pi]} \|f(\omega)-f_n(\omega)\| = 0$	$\lim_{n\to\infty} \sup_{\omega\in[-\pi,\pi]} \|f(\omega)-f_n(\omega)\| = 0$
Def. of equivalence $A_n \sim B_n$ as $n\to\infty$	$\lim_{n\to\infty} \|A_n - B_n\| = 0$	$\lim_{n\to\infty} \dfrac{\\|(A_n-B_n)\Gamma_n(\omega)\\|^2}{K_n(\omega,\omega)} = 0$
Algebraic Properties	$T_n(f)T_n(g) \sim T_n(fg)$	$M_n(f)M_n(g) \sim M_n(fg)$
Extensions $\sigma_k = \pm 1$	$\prod_{k=1}^{m} T_n^{\sigma_k}(f_k) \sim T_n\left(\prod_{k=1}^{m} f_k^{\sigma_k}\right)$	$\prod_{k=1}^{m} M_n^{\sigma_k}(f_k) \sim M_n\left(\prod_{k=1}^{m} f_k^{\sigma_k}\right)$

TABLE 6.1
Summary of classical results and their relation to the generalisations derived here.

REFERENCES

[1] N. ARONSZAJN, *Theory of reproducing kernels*, Acta Mathematica, (1950), pp. 337–404.

[2] R. BITMEAD, M. GEVERS, AND V. WERTZ, *Adaptive Optimal Control, The Thinking Man's GPC.*, Prentice Hall, Englewood Cliffs, New Jersey, 1990.

[3] P. W. BROOME, *Discrete orthonormal sequences*, Journal of the Association for Computing Machinery, 12 (1965), pp. 151–168.

[4] C. K. CHUI, *Wavelets:A Tutorial in Theory and Applications*, vol. 2 of Wavelet Analysis and Its Applications, Academic Press, Boston, 1992.

[5] P. R. CLEMENT, *Laguerre functions in signal analysis and parameter identification*, Journal of the Franklin Institute, 313 (1982), pp. 85–95.

[6] I. DAUBECHIES, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[7] P. DEWILDE AND H. DYM, *Schur recursions, error formulas, and convergence of rational estimators for stationary stochastic sequences*, IEEE Transactions on Information Theory, IT-27 (1981), pp. 446–461.

[8] P. DEWILDE, A. VIEIRA, AND T. KAILATH, *On a generalised Szegö–Levinson realization algorithm for optimal linear predictors based on a network synthesis approach*, IEEE Transactions on Circuits and Systems, CAS-25 (1978), pp. 663–675.

[9] R. EDWARDS, *Fourier Series, a Modern Introduction.*, vol. 2 of Graduate texts in mathematics number 64., Springer Verlag, 1979.

[10] L. Y. GERONIMUS, *Orthogonal polynomials: Estimates, asymptotic formulas, and series of polynomials orthogonal on the unit circle and on an interval*, Consultants Bureau, New York, 1961. Authorized translation from the Russian.

[11] M. GEVERS, L. LJUNG, AND P.M.J. VAN DEN HOF, *Asymptotic variance expressions for closed-loop identification and their relevance in identification for control*, in Proceedings of the 11th IFAC Symposium on System Identification, 1997, pp. 1449–1454.

[12] U. GRENANDER AND G. SZEGÖ, *Toeplitz Forms and their Applications*, University of California Press, Berkeley, 1958.

[13] S. GUNNARSSON AND L. LJUNG, *Frequency domain tracking characteristics of adaptive algorithms*, IEEE Transactions on Acoustics, Speech, Signal Processing, 37 (1989), pp. 1072–1089.

[14] L. GUO AND L. LJUNG, *The role of model validation for assessing the size of unmodelled dynamics*, in Proceedings of the 33rd Conference on Decision and Control, IEEE, December 1994, pp. 3894–3899.

[15] P. HEUBERGER, P.M.J. VAN DEN HOF, AND O. BOSGRA, *A generalized orthonormal basis for linear dynamical systems*, IEEE Transactions on Automatic Control, AC-40 (1995), pp. 451–465.

[16] K. HOFFMAN, *Banach Spaces of Analytic Functions*, Prentice-Hall, Inc., New Jersey, 1962.

[17] C. HWANG AND Y.-P. SHIH, *Parameter identification via Laguerre polynomials*, International Journal of Systems Science, 13 (1982), pp. 209–217.

[18] W. H. KAUTZ, *Network synthesis for specified transient response*, Tech. Rep. 209, Massachusetts Institute of Technology, Research Laboratory of Electronics, 1952.

[19] R. KING AND P. PARASKEVOPOULOS, *Digital Laguerre filters*, Circuit Theory and Applications, 5 (1977), pp. 81–91.

[20] A. KOLMOGOROV, *Interpolation und extrapolation von stationären zufälligen folgen*, Bulletin de l'académie des Sciences de U.S.S.R., 5 (1941), pp. 3–14.

[21] Y. W. LEE, *Statistical Theory of Communication*, Wiley, New York, 1960.

[22] N. LEVINSON, *The Wiener RMS (root mean square) error criterion in filter design and prediction*, Journal of Mathematics and Physics, 25 (1947), pp. 261–278.

[23] L. LJUNG, *System Identification: Theory for the User*, Prentice-Hall, Inc., New Jersey, 1987.

[24] L. LJUNG, *Identification in closed loop: Some aspects on direct and indirect approaches*, in Proceedings of the 11th IFAC Symposium on System Identification, 1997, pp. 141–146.

[25] L.LJUNG, *Asymptotic variance expressions for identified black-box transfer function models*, IEEE Transactions on Automatic Control, AC-30 (1985), pp. 834–844.

[26] L.LJUNG AND B. WAHLBERG, *Asymptotic properties of the least squares method for estimating transfer functions and disturbance spectra*, Advances in Applied Probability, 24 (1992), pp. 412–440.

[27] L.LJUNG AND Z.D.YUAN, *Asymptotic properties of black-box identification of transfer functions*, IEEE Transactions on Automatic Control, 30 (1985), pp. 514–530.

[28] P. MÄKILA, *Approximation of stable systems by Laguerre filters*, Automatica, 26 (1990), pp. 333–345.

[29] F. MALMQUIST, *Sur la détermination d'une classe de fonctions analytiques par leurs valeurs dans un ensemble donné de points*, Comptes Rendus du Sixième Congrès des mathématiciens scandinaves (Kopenhagen), (1925), pp. 253–259.

[30] B. NINNESS AND F. GUSTAFSSON, *A unifying construction of orthonormal bases for system identification*, Tech. Rep. EE9432, Department of Electrical and Computer Engineering at the University of Newcastle, August 1994.

[31] ———, *A unifying construction of orthonormal bases for system identification*, IEEE Transactions on Automatic Control, 42 (1997), pp. 515–521.

[32] B. NINNESS, H. HJALMARSSON, AND F. GUSTAFSSON, *The fundamental role of general orthonormal bases in system identification*, Technical Report EE9739, Department of Electrical and Computer Engineering, University of Newcastle, Australia. Accepted subject to revision,

IEEE Transactions on Automatic Control., (1997).

[33] A. Papoulis, *Levinson's algorithm, Wold's decomposition, and spectral estimation*, SIAM Review, 27 (1985), pp. 405–441.

[34] J. R. Partington, *Approximation of delay systems by Fourier-Laguerre series*, Automatica, 27 (1991), pp. 569–572.

[35] P.M.J. Van den Hof, P.S.C. Heuberger, and J. Bokor, *System identification with generalized orthonormal basis functions*, Automatica, 31 (1995), pp. 1821–1834.

[36] P. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*, Marcel Dekker, 1995.

[37] I. Schur, *Bemerkungen zur theorie der beschränkten bilinearformen mit unendlich vielen veränderlichen*, Journal für die Reine und Angewandte Mathematik, 140 (1911), pp. 1–28.

[38] S.Pillai and T.Shim, *Spectrum Estimation and System Identification*, Springer–Verlag, 1993.

[39] G. Szegö, *Orthogonal Polynomials*, vol. 23, Colloquium publications - American Mathematical Society, New York, 1939.

[40] B. Wahlberg, *System identification using Laguerre models*, IEEE Transactions on Automatic Control, AC-36 (1991), pp. 551–562.

[41] B. Wahlberg, *System identification using Kautz models*, IEEE Transactions on Automatic Control, AC-39 (1994), pp. 1276–1282.

[42] B. Wahlberg and E. Hannan, *Parameteric signal modelling using Laguerre filters*, The Annals of Applied Probabililty, 3 (1993), pp. 467–496.

[43] J. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, vol. XX of Colloquium Publications, American Mathematical Society, 1935.

[44] H. Widom, *Studies in Real and Complex Analysis*, MAA Studies in Mathematics, Prentice Hall, Englewood Cliffs, NJ, 1965, ch. Toeplitz Matrices.

[45] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, M.I.T. Press, 1949.

[46] Z. Yuan and L. Ljung, *Black box identification of multivariable transfer functions:Asymptotic properties and optimal input design*, International Journal of Control, 40 (1984), pp. 233–256.