Generalised Fourier and Toeplitz results for Rational Orthonormal Bases

Brett Ninness*
Department of Electrical & Computer Engineering and
Centre for Integrated Dynamics and Control (CIDAC)
University of Newcastle, Callaghan 2308, Australia
brett@ee.newcastle.edu.au

Håkan Hjalmarsson†
Department of Sensors, Signals and Systems-Signal Processing,
The Royal Institute of Technology, S-100 44 Stockholm, Sweden.
hakan.hjalmarsson@s3.e.kth.se

Fredrik Gustafsson‡
Department of Electrical Engineering,
Linköping University, S-581 83, Linköping, Sweden.
fredrik@isy.liu.se

Keywords: Orthonormal Bases, Toeplitz Operators, Generalized Fourier Analysis, System Identification.

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.

1 Introduction

Tackling system theoretic problems using orthonormal descriptions has a particularly rich history, going back at least as far as the work of Kolmogorov [5] and Wiener [14] 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 [12].

In examining this latter work, 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

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

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

$$\frac{1}{n}\Gamma_n^\star(\omega)T_n(f)\Gamma_n(\omega) = \sum_{k=-n}^{n}\left(1-\frac{|k|}{n}\right)c_k\, e^{j\omega k} \qquad (2)$$

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

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

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 (2) may be recognised as the Cesàro mean reconstruction of a Fourier series which is known [1], provided f is continuous, to converge uniformly to $f(\omega)$ on $[-\pi,\pi]$.

This latter fact has been exploited by Ljung and co-workers [7, 3, 8, 6] 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

*This work was supported by the Australian Research Council and CIDAC, 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

‡Work partly completed while on leave at CIDAC

order, and observed data length; see [11, 10] for more detail on this point.

Such results have found wide engineering application. However, to derive them, another key ingredient pertaining to the properties of Toeplitz matrices is required. Namely, that asymptotically in size n, Toeplitz matrices possess the algebraic structure [2, 13]

$$T_n(f)T_n(g) \sim T_n(fg) \tag{3}$$

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

$$|A|^2 \triangleq \frac{1}{n}\text{Trace}\{A^\star A\}. \tag{4}$$

The main results of this paper are to extend the results of the convergence of the Cesàro mean (2) and the algebraic structure of Toeplitz matrices (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

$$\mathcal{B}_n(z) \triangleq \frac{\sqrt{1-|\xi_n|^2}}{1-\xi_n z}\prod_{k=0}^{n-1}\left(\frac{z-\overline{\xi_k}}{1-\xi_k z}\right) \tag{5}$$

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

$$[M_n(f)]_{\ell,m} \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathcal{B}_\ell(e^{j\omega})\overline{\mathcal{B}_m(e^{j\omega})}f(\omega)\,\mathrm{d}\omega \tag{6}$$

of Toeplitz matrices is considered, for which it is shown here that a generalisation of (3) still holds, and with the redefinition

$$\Gamma_n^T(\omega) \triangleq [\mathcal{B}_0(e^{j\omega}),\ \mathcal{B}_1(e^{j\omega}),\ \ldots, \mathcal{B}_{n-1}(e^{j\omega})] \tag{7}$$

it is also shown here that a generalisation of the uniform convergence of the Cesàro mean (2) to $f(\omega)$ also holds.

In both cases, the generalisation involves replacing the $1/n$ normalisation appearing in (2) and in the definition of the matrix norm (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 $\{\mathcal{B}_0, \mathcal{B}_1, \cdots, \mathcal{B}_{n-1}\}$.

2 Completeness Properties

The theme of this paper is to examine certain system theoretic issues pertaining to the use of the basis functions (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 one-sided Fourier expansion. As is well known [4], $H_2(\mathbf{T})$ is a Hilbert space when endowed with the inner product

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{j\omega})\overline{g(e^{j\omega})}\,\mathrm{d}\omega \quad f,g \in H_2(\mathbf{T}). \tag{8}$$

That the functions (5) form an orthonormal set in that $\langle \mathcal{B}_n, \mathcal{B}_m\rangle = \delta(n-m) =$ Kronecker delta may easily be shown [9] using the contour integral formulation of the inner product in (8) and Cauchy's residue Theorem.

What must be of central interest if the functions (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 completeness result which has been developed elsewhere, but is presented here for the sake of a self contained presentation.

Theorem 1 (Ninness and Gustafsson [9]).

$$\overline{\text{Span}\,\{\mathcal{B}_k(z)\}_{k\geq 0}} = H_2(\mathbf{T})$$

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 1, to further examine the properties of approximants formed as linear combinations of the basis functions (5), this paper takes the approach of utilising the idea of a 'reproducing kernel' $K_n(z,\mu)$ defined as

$$K_n(z,\mu) \triangleq \sum_{k=0}^{n-1}\mathcal{B}_k(z)\overline{\mathcal{B}_k(\mu)}. \tag{9}$$

In common with the study of orthogonal polynomials [12] via the use of reproducing kernels, a simpler closed form formula for $K_n(z,\mu)$ (called 'Christoffel-Darboux formulae)' is required for the ensuing analysis.

Theorem 2. *<u>Christoffel–Darboux Formula:</u> 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 space spanned by $\{\mathcal{B}_0, \mathcal{B}_1, \cdots, \mathcal{B}_{n-1}\}$ can be expressed as

$$K_n(z,\mu) = \frac{1-\overline{\varphi_n(\mu)}\varphi_n(z)}{1-z\overline{\mu}}. \tag{10}$$

Given a function $f \in H_2(\mathbf{T})$, an obvious way of approximating it in terms of the basis functions $\{\mathcal{B}_0, \mathcal{B}_1, \cdots, \mathcal{B}_{n-1}\}$ is as f_n given by

$$f_n(z) = \arg\min_{g \in X_n} \|f - g\| = \sum_{k=0}^{n-1} \langle f, \mathcal{B}_k \rangle \mathcal{B}_k(z). \tag{11}$$

Provided $\sum(1 - |\xi_k|) = \infty$ holds, then by the completeness theorem 1, the approximation error $\|f_n - f\|$ can be made arbitrarily small for arbitrarily large approximation order n.

A natural question to then ask is how the approximant f_n behaves with respect to other norms, for example the supremum norm on $[-\pi, \pi]$. The purpose of this section is to show that a modified approximant, defined as

$$f_n(\omega) \triangleq \frac{\Gamma_n^\star(\omega) M_n(f) \Gamma_n(\omega)}{K_n(\omega, \omega)} \tag{12}$$

(where Γ_n defined in (7) is an $n \times 1$ vector of general rational orthonormal basis functions (5) and $M_n(f)$ is a generalised Toeplitz matrix as defined in (6) and deriving from the Cesàro (or Fejér) mean of classical Fourier analysis, is supremum norm convergent to f under the same condition of $\sum(1 - |\xi_k|) = \infty$. This result encompasses the classical result for the trigonometric basis by simply setting all the poles $\{\xi_k\}$ to zero.

To analyse the convergence properties of (12), note that by the formulation (9)

$$\Gamma_n^\star(\omega) M_n(f) \Gamma_n(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\sigma) \left| K_n(\omega, \sigma) \right|^2 \, \mathrm{d}\sigma.$$

Therefore, since by the defining property of the reproducing kernel

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |K_n(\omega, \sigma)|^2 \, \mathrm{d}\sigma = \sum_{m=0}^{n-1} |\mathcal{B}_m(e^{j\omega})|^2 \tag{13}$$

then

$$\frac{1}{2\pi} \left| \frac{\Gamma_n^\star(\omega) M_n(f) \Gamma_n(\omega)}{K_n(\omega, \omega)} - f(\omega) \right| =$$

$$\frac{1}{2\pi K_n(\omega, \omega)} \left| \int_{-\pi}^{\pi} [f(\sigma) - f(\omega)] \, |K_n(\omega, \sigma)|^2 \, \mathrm{d}\sigma \right| \tag{14}$$

and so, in analogy with classical Fourier analysis, convergence of the generalised Cesàro mean approximant (12) hinges on a kernel function (in this case depending on the reproducing kernel and being given by $|K_n(\omega, \sigma)|^2 / K_n(\omega, \omega)$) behaving in some sense like the Dirac delta function $\delta(\omega - \sigma)$. Via use of the Christoffel–Darboux formula for the reproducing kernel, it is possible to establish that this 'delta-like' behaviour does in fact occur in the following sense.

Lemma 1. *For any $\rho > 0$ and provided*

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

$$\lim_{n \to \infty} \frac{1}{K_n(\omega, \omega)} \int_{\sigma \notin [\omega - \rho, \omega + \rho]} |K_n(\omega, \sigma)|^2 \, \mathrm{d}\sigma = 0.$$

Use of this 'approximate identity' property allows the following result to be established.

Theorem 3. *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 $|\xi_n| \leq 1 - \delta$ 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 Generalised Toeplitz Matrix Algebra

In applications [7, 3, 8, 6], the consideration of quadratic forms more complicated than (12) 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 [7, 3, 8, 6], 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 (3), the convenient property that $T_n(f) T_n(g) \sim T_n(fg)$ is assured [2, 13] (the meaning of the $\sim$ notation here is as described in conjunction with equation (3)).

The purpose of this section is to establish this same algebraic structure for the generalised Toeplitz matrices defined by (6), the classical results once again arising as the special case of $\xi_k = 0$ in (5).

In presenting this, a definition is required 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 $\forall \omega \in [-\pi, \pi]$

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

With this definition in hand, the following result on the algebraic structure of generalised Toeplitz matrices is available.

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

tem 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 $\prod_{k=1}^{n} A_k = A_1 A_2 \cdots A_n$.

Corollary 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 (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 3 then provides a further corollary representing an extension of the generalised Fourier convergence of Theorem 3.

Corollary 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 (5) satisfy $|\xi_k| \leq 1 - \delta$ for some $\delta > 0$. Then the following limit result holds*

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

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

As a simple but important example of the utility of this corollary, it allows the conclusion that when all the poles $\{\xi_k\}$ are chosen in a closed subset of $\mathbf{D}$ then

$$\lim_{n\to\infty} \frac{\Gamma_n^\star(\omega) M_n^{-1}(f) \Gamma_n(\omega)}{K_n(\omega,\omega)} = \frac{1}{f(\omega)} \tag{15}$$

which has particular relevance to the study of reproducing kernels with respect to weighted inner products [11].

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. The proofs, together with more detailed discussion of the results presented in this short paper are available in [11]. Applications of the results presented here are available in [10].

[1] R. Edwards, *Fourier Series, a Modern Introduction.*, Springer Verlag, 1979.

[2] U. Grenander and G. Szegö, *Toeplitz Forms and their Applications*, Uni. Calif. Press, 1958.

[3] S. Gunnarsson and L. Ljung, *Frequency domain tracking characteristics of adaptive algorithms*, IEEE Trans. Sig. Proc, 37 (1989), pp. 1072–1089.

[4] K. Hoffman, *Banach Spaces of Analytic Functions*, Prentice Hall, 1962.

[5] 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.

[6] L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, 1987.

[7] L.Ljung, *Asymptotic variance expressions for identified black-box transfer function models*, IEEE Trans. Auto. Cont, AC-30 (1985), pp. 834–844.

[8] L.Ljung and B. Wahlberg, *Asymptotic properties of the least squares method for estimating transfer functions and disturbance spectra*, Advances in App. Prob, 24 (1992), pp. 412–440.

[9] B. Ninness and F. Gustafsson, *A unifying construction of orthonormal bases for system identification*, IEEE Trans. Auto. Cont, 42 (1997), pp. 515–521.

[10] B. Ninness, H. Hjalmarsson, and F. Gustafsson, *The fundamental role of general orthonormal bases in system identification*, Tech. Rep. EE9739, Dept. EE & CE, Uni. Newcastle, Australia. Accepted subject to revision, IEEE Trans. Auto. Cont, (1998).

[11] ——, *Generalised Fourier and Toeplitz results for rational orthonormal bases*, Tech. Rep. EE9740, Dept. EE & CE, Uni. Newcastle, Australia. To appear in SIAM J. Cont. & Opt, (1998).

[12] G. Szegö, *Orthogonal Polynomials*, vol. 23, Colloq. publ. - Amer. Math. Soc, New York, 1939.

[13] H. Widom, *Studies in Real and Complex Analysis*, MAA Studies in Math, Prent. Hall, 1965, ch. Toeplitz Matrices.

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