Toeplitz and Circulant Matrices: A Review

Toeplitz and Circulant Matrices: A Review

Robert M. Gray

Department of Electrical Engineering
Stanford University
Stanford CA 94305, USA
rmgray@stanford.edu

the essence of knowledge

Boston – Delft

Foundations and Trends® in Communications and Information Theory

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

A Cataloging-in-Publication record is available from the Library of Congress

The preferred citation for this publication is R.M. Gray, Toeplitz and Circulant Matrices: A Review, Foundation and Trends® in Communications and Information Theory, vol 2, no 3, pp 155–239, 2006

Printed on acid-free paper

ISBN: 1-933019-23-9

Foundations and Trends® in Communications and Information Theory

Volume 2 Issue 3, 2006

Editorial Scope

Foundations and Trends® in Communications and Information Theory will publish survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians
Foundations and Trends® in Communications and Information Theory, 2006, Volume 2, 4 issues. ISSN paper version 1567-2190.ISSN online version 1567-2328.Also available as a combined paper and online subscription.

Foundations and Trends® in
Communications and Information Theory
Vol. 2, No 3 (2006) 155–239

Toeplitz and Circulant Matrices: A Review

Robert M. Gray[1]

[1] *Stanford University, Stanford CA 94305, USA, rmgray@stanford.edu*

Abstract

$$\begin{bmatrix} t_0 & t_{-1} & t_{-2} & \cdots & t_{-(n-1)} \\ t_1 & t_0 & t_{-1} & & \\ t_2 & t_1 & t_0 & & \vdots \\ \vdots & & & \ddots & \\ t_{n-1} & & & \cdots & t_0 \end{bmatrix}$$

The fundamental theorems on the asymptotic behavior of eigenvalues, inverses, and products of banded Toeplitz matrices and Toeplitz matrices with absolutely summable elements are derived in a tutorial manner. Mathematical elegance and generality are sacrificed for conceptual simplicity and insight in the hope of making these results available to engineers lacking either the background or endurance to attack the mathematical literature on the subject. By limiting the generality of the matrices considered, the essential ideas and results can be conveyed in a more intuitive manner without the mathematical machinery required for the most general cases. As an application the results are applied to the study of the covariance matrices and their factors of linear models of discrete time random processes.

Contents

1 Introduction **1**

1.1 Toeplitz and Circulant Matrices 1
1.2 Examples 5
1.3 Goals and Prerequisites 9

2 The Asymptotic Behavior of Matrices **11**

2.1 Eigenvalues 11
2.2 Matrix Norms 14
2.3 Asymptotically Equivalent Sequences of Matrices 17
2.4 Asymptotically Absolutely Equal Distributions 24

3 Circulant Matrices **31**

3.1 Eigenvalues and Eigenvectors 32
3.2 Matrix Operations on Circulant Matrices 34

4 Toeplitz Matrices **37**

4.1 Sequences of Toeplitz Matrices 37
4.2 Bounds on Eigenvalues of Toeplitz Matrices 41
4.3 Banded Toeplitz Matrices 43
4.4 Wiener Class Toeplitz Matrices 48

5 Matrix Operations on Toeplitz Matrices **61**

5.1 Inverses of Toeplitz Matrices 62
5.2 Products of Toeplitz Matrices 67
5.3 Toeplitz Determinants 71

6 Applications to Stochastic Time Series **73**

6.1 Moving Average Processes 74
6.2 Autoregressive Processes 78
6.3 Factorization 81

References **87**

1

Introduction

1.1 Toeplitz and Circulant Matrices

A Toeplitz matrix is an $n \times n$ matrix $T_n = [t_{k,j}; k,j = 0,1,\ldots,n-1]$ where $t_{k,j} = t_{k-j}$, i.e., a matrix of the form

$$T_n = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & \cdots & t_{-(n-1)} \\ t_1 & t_0 & t_{-1} & & \\ t_2 & t_1 & t_0 & & \vdots \\ \vdots & & & \ddots & \\ t_{n-1} & & & \cdots & t_0 \end{bmatrix}. \tag{1.1}$$

Such matrices arise in many applications. For example, suppose that

$$x = (x_0, x_1, \ldots, x_{n-1})' = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}$$

is a column vector (the prime denotes transpose) denoting an "input" and that t_k is zero for $k < 0$. Then the vector

$$
\begin{aligned}
y &= T_n x = \begin{bmatrix} t_0 & 0 & 0 & \cdots & 0 \\ t_1 & t_0 & 0 & & \\ t_2 & t_1 & t_0 & & \vdots \\ \vdots & & & \ddots & \\ t_{n-1} & & & \cdots & t_0 \end{bmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{pmatrix} \\
&= \begin{pmatrix} x_0 t_0 \\ t_1 x_0 + t_0 x_1 \\ \sum_{i=0}^{2} t_{2-i} x_i \\ \vdots \\ \sum_{i=0}^{n-1} t_{n-1-i} x_i \end{pmatrix}
\end{aligned}
$$

with entries

$$y_k = \sum_{i=0}^{k} t_{k-i} x_i$$

represents the the output of the discrete time causal time-invariant filter h with "impulse response" t_k. Equivalently, this is a matrix and vector formulation of a discrete-time convolution of a discrete time input with a discrete time filter.

As another example, suppose that $\{X_n\}$ is a discrete time random process with mean function given by the expectations $m_k = E(X_k)$ and covariance function given by the expectations $K_X(k,j) = E[(X_k - m_k)(X_j - m_j)]$. Signal processing theory such as prediction, estimation, detection, classification, regression, and communcations and information theory are most thoroughly developed under the assumption that the mean is constant and that the covariance is Toeplitz, i.e., $K_X(k,j) = K_X(k-j)$, in which case the process is said to be weakly stationary. (The terms "covariance stationary" and "second order stationary" also are used when the covariance is assumed to be Toeplitz.) In this case the $n \times n$ covariance matrices $K_n = [K_X(k,j); k,j = 0,1,\ldots,n-1]$ are Toeplitz matrices. Much of the theory of weakly stationary processes involves applications of Toeplitz matrices. Toeplitz matrices also arise in solutions to differential and integral equations, spline functions, and problems and methods in physics, mathematics, statistics, and signal processing.

A common special case of Toeplitz matrices – which will result in significant simplification and play a fundamental role in developing more general results – results when every row of the matrix is a right cyclic shift of the row above it so that $t_k = t_{-(n-k)} = t_{k-n}$ for $k = 1, 2, \ldots, n-1$. In this case the picture becomes

$$C_n = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & \cdots & t_{-(n-1)} \\ t_{-(n-1)} & t_0 & t_{-1} & & \\ t_{-(n-2)} & t_{-(n-1)} & t_0 & & \vdots \\ \vdots & & & \ddots & \\ t_{-1} & t_{-2} & & \cdots & t_0 \end{bmatrix}. \tag{1.2}$$

A matrix of this form is called a *circulant* matrix. Circulant matrices arise, for example, in applications involving the discrete Fourier transform (DFT) and the study of cyclic codes for error correction.

A great deal is known about the behavior of Toeplitz matrices – the most common and complete references being Grenander and Szegö [15] and Widom [33]. A more recent text devoted to the subject is Böttcher and Silbermann [5]. Unfortunately, however, the necessary level of mathematical sophistication for understanding reference [15] is frequently beyond that of one species of applied mathematician for whom the theory can be quite useful but is relatively little understood. This caste consists of engineers doing relatively mathematical (for an engineering background) work in any of the areas mentioned. This apparent dilemma provides the motivation for attempting a tutorial introduction on Toeplitz matrices that proves the essential theorems using the simplest possible and most intuitive mathematics. Some simple and fundamental methods that are deeply buried (at least to the untrained mathematician) in [15] are here made explicit.

The most famous and arguably the most important result describing Toeplitz matrices is Szegö's theorem for sequences of Toeplitz matrices $\{T_n\}$ which deals with the behavior of the eigenvalues as n goes to infinity. A complex scalar α is an eigenvalue of a matrix A if there is a nonzero vector x such that

$$Ax = \alpha x, \tag{1.3}$$

in which case we say that x is a (right) eigenvector of A. If A is Hermitian, that is, if $A^* = A$, where the asterisk denotes conjugate transpose, then the eigenvalues of the matrix are real and hence $\alpha^* = \alpha$, where the asterisk denotes the conjugate in the case of a complex scalar. When this is the case we assume that the eigenvalues $\{\alpha_i\}$ are ordered in a nondecreasing manner so that $\alpha_0 \geq \alpha_1 \geq \alpha_2 \cdots$. This eases the approximation of sums by integrals and entails no loss of generality. Szegö's theorem deals with the asymptotic behavior of the eigenvalues $\{\tau_{n,i};\ i = 0, 1, \ldots, n-1\}$ of a sequence of Hermitian Toeplitz matrices $T_n = [t_{k-j}; k, j = 0, 1, 2, \ldots, n-1]$. The theorem requires that several technical conditions be satisfied, including the existence of the Fourier series with coefficients t_k related to each other by

$$f(\lambda) = \sum_{k=-\infty}^{\infty} t_k e^{ik\lambda}; \quad \lambda \in [0, 2\pi] \tag{1.4}$$

$$t_k = \frac{1}{2\pi} \int_0^{2\pi} f(\lambda) e^{-ik\lambda} \, d\lambda. \tag{1.5}$$

Thus the sequence $\{t_k\}$ determines the function f and vice versa, hence the sequence of matrices is often denoted as $T_n(f)$. If $T_n(f)$ is Hermitian, that is, if $T_n(f)^* = T_n(f)$, then $t_{-k} = t_k^*$ and f is real-valued.

Under suitable assumptions the Szegö theorem states that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} F(\tau_{n,k}) = \frac{1}{2\pi} \int_0^{2\pi} F(f(\lambda)) \, d\lambda \tag{1.6}$$

for any function F that is continuous on the range of f. Thus, for example, choosing $F(x) = x$ results in

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \tau_{n,k} = \frac{1}{2\pi} \int_0^{2\pi} f(\lambda) \, d\lambda, \tag{1.7}$$

so that the arithmetic mean of the eigenvalues of $T_n(f)$ converges to the integral of f. The trace $\mathrm{Tr}(A)$ of a matrix A is the sum of its diagonal elements, which in turn from linear algebra is the sum of the eigenvalues of A if the matrix A is Hermitian. Thus (1.7) implies that

$$\lim_{n\to\infty} \frac{1}{n} \mathrm{Tr}(T_n(f)) = \frac{1}{2\pi} \int_0^{2\pi} f(\lambda) \, d\lambda. \tag{1.8}$$

Similarly, for any power s

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\tau_{n,k}^{s}=\frac{1}{2\pi}\int_{0}^{2\pi}f(\lambda)^{s}\,d\lambda. \tag{1.9}$$

If f is real and such that the eigenvalues $\tau_{n,k}\geq m>0$ for all n,k, then $F(x)=\ln x$ is a continuous function on $[m,\infty)$ and the Szegö theorem can be applied to show that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\ln\tau_{n,i}=\frac{1}{2\pi}\int_{0}^{2\pi}\ln f(\lambda)\,d\lambda. \tag{1.10}$$

From linear algebra, however, the determinant of a matrix $T_n(f)$ is given by the product of its eigenvalues,

$$\det(T_n(f))=\prod_{i=0}^{n-1}\tau_{n,i},$$

so that (1.10) becomes

$$\begin{aligned}\lim_{n\to\infty}\ln\det(T_n(f))^{1/n} &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\ln\tau_{n,i}\\ &= \frac{1}{2\pi}\int_{0}^{2\pi}\ln f(\lambda)\,d\lambda. \end{aligned}\tag{1.11}$$

As we shall later see, if f has a lower bound $m>0$, than indeed all the eigenvalues will share the lower bound and the above derivation applies. Determinants of Toeplitz matrices are called *Toeplitz determinants* and (1.11) describes their limiting behavior.

1.2 Examples

A few examples from statistical signal processing and information theory illustrate the the application of the theorem. These are described with a minimum of background in order to highlight how the asymptotic eigenvalue distribution theorem allows one to evaluate results for processes using results from finite-dimensional vectors.

The differential entropy rate of a Gaussian process

Suppose that $\{X_n;\, n = 0,1,\ldots\}$ is a random process described by probability density functions $f_{X^n}(x^n)$ for the random vectors $X^n = (X_0, X_1, \ldots, X_{n-1})$ defined for all $n = 0,1,2,\ldots$. The Shannon differential entropy $h(X^n)$ is defined by the integral

$$h(X^n) = -\int f_{X^n}(x^n) \ln f_{X^n}(x^n)\, dx^n$$

and the differential entropy rate of the random process is defined by the limit

$$h(X) = \lim_{n\to\infty} \frac{1}{n} h(X^n)$$

if the limit exists. (See, for example, Cover and Thomas[7].)

A stationary zero mean Gaussian random process is completely described by its mean correlation function $r_{k,j} = r_{k-j} = E[X_k X_j]$ or, equivalently, by its power spectral density function f, the Fourier transform of the covariance function:

$$\begin{aligned} f(\lambda) &= \sum_{n=-\infty}^{\infty} r_n e^{in\lambda}, \\ r_k &= \frac{1}{2\pi} \int_0^{2\pi} f(\lambda) e^{-i\lambda k}\, d\lambda \end{aligned}$$

For a fixed positive integer n, the probability density function is

$$f_{X^n}(x^n) = \frac{e^{-\frac{1}{2} x^{n\prime} R_n^{-1} x^n}}{(2\pi)^{n/2} \det(R_n)^{1/2}},$$

where R_n is the $n \times n$ covariance matrix with entries r_{k-j}. A straightforward multidimensional integration using the properties of Gaussian random vectors yields the differential entropy

$$h(X^n) = \frac{1}{2} \ln(2\pi e)^n \det R_n.$$

The problem at hand is to evaluate the entropy rate

$$h(X) = \lim_{n\to\infty} \frac{1}{n} h(X^n) = \frac{1}{2}\ln(2\pi e) + \lim_{n\to\infty} \frac{1}{n} \ln\det(R_n).$$

The matrix R_n is the Toeplitz matrix T_n generated by the power spectral density f and $\det(R_n)$ is a Toeplitz determinant and we have immediately from (1.11) that

$$h(X) = \frac{1}{2}\log\left(2\pi e \frac{1}{2\pi}\int_0^{2\pi} \ln f(\lambda)\, d\lambda\right). \tag{1.12}$$

This is a typical use of (1.6) to evaluate the limit of a sequence of finite-dimensional qualities, in this case specified by the determinants of of a sequence of Toeplitz matrices.

The Shannon rate-distortion function of a Gaussian process

As a another example of the application of (1.6), consider the evaluation of the rate-distortion function of Shannon information theory for a stationary discrete time Gaussian random process with 0 mean, covariance $K_X(k,j) = t_{k-j}$, and power spectral density $f(\lambda)$ given by (1.4). The rate-distortion function characterizes the optimal tradeoff of distortion and bit rate in data compression or source coding systems. The derivation details can be found, e.g., in Berger [3], Section 4.5, but the point here is simply to provide an example of an application of (1.6). The result is found by solving an n-dimensional optimization in terms of the eigenvalues $\tau_{n,k}$ of $T_n(f)$ and then taking limits to obtain parametric expressions for distortion and rate:

$$\begin{aligned} D_\theta &= \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \min(\theta, \tau_{n,k}) \\ R_\theta &= \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \max(0, \frac{1}{2}\ln\frac{\tau_{n,k}}{\theta}). \end{aligned}$$

The theorem can be applied to turn this limiting sum involving eigenvalues into an integral involving the power spectral density:

$$\begin{aligned} D_\theta &= \int_0^{2\pi} \min(\theta, f(\lambda))\, d\lambda \\ R_\theta &= \int_0^{2\pi} \max\left(0, \frac{1}{2}\ln\frac{f(\lambda)}{\theta}\right) d\lambda. \end{aligned}$$

Again an infinite dimensional problem is solved by first solving a finite dimensional problem involving the eigenvalues of matrices, and then using the asymptotic eigenvalue theorem to find an integral expression for the limiting result.

One-step prediction error

Another application with a similar development is the one-step prediction error problem. Suppose that X_n is a weakly stationary random process with covariance t_{k-j}. A classic problem in estimation theory is to find the best linear predictor based on the previous n values of X_i, $i = 0, 1, 2, \ldots, n-1$,

$$\hat{X}_n = \sum_{i=1}^{n} a_i X_{n-i},$$

in the sense of minimizing the mean squared error $E[(X_n - \hat{X}_n)^2]$ over all choices of coefficients a_i. It is well known (see, e.g., [10]) that the minimum is given by the ratio of Toeplitz determinants $\det T_{n+1}/\det T_n$. The question is to what this ratio converges in the limit as n goes to ∞. This is not quite in a form suitable for application of the theorem, but we have already evaluated the limit of $\det T_n^{1/n}$ in (1.11) and for large n we have that

$$(\det T_n)^{1/n} \approx \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \ln f(\lambda)\, d\lambda\right) \approx (\det T_{n+1})^{1/(n+1)}$$

and hence in particular that

$$(\det T_{n+1})^{1/(n+1)} \approx (\det T_n)^{1/n}$$

so that

$$\frac{\det T_{n+1}}{\det T_n} \approx (\det T_n)^{1/n} \to \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \ln f(\lambda)\, d\lambda\right),$$

providing the desired limit. These arguments can be made exact, but it is hoped they make the point that the asymptotic eigenvalue distribution theorem for Hermitian Toeplitz matrices can be quite useful for evaluating limits of solutions to finite-dimensional problems.

Further examples

The Toeplitz distribution theorems have also found application in more complicated information theoretic evaluations, including the channel capacity of Gaussian channels [30, 29] and the rate-distortion functions of autoregressive sources [12]. The examples described here were chosen because they were in the author's area of competence, but similar applications crop up in a variety of areas. A Google$^{\mathrm{TM}}$ search using the title of this document shows diverse applications of the eigenvalue distribution theorem and related results, including such areas of coding, spectral estimation, watermarking, harmonic analysis, speech enhancement, interference cancellation, image restoration, sensor networks for detection, adaptive filtering, graphical models, noise reduction, and blind equalization.

1.3 Goals and Prerequisites

The primary goal of this work is to prove a special case of Szegö's asymptotic eigenvalue distribution theorem in Theorem 9. The assumptions used here are less general than Szegö's, but this permits more straightforward proofs which require far less mathematical background. In addition to the fundamental theorems, several related results that naturally follow but do not appear to be collected together anywhere are presented. We do not attempt to survey the fields of applications of these results, as such a survey would be far beyond the author's stamina and competence. A few applications are noted by way of examples.

The essential prerequisites are a knowledge of matrix theory, an engineer's knowledge of Fourier series and random processes, and calculus (Riemann integration). A first course in analysis would be helpful, but it is not assumed. Several of the occasional results required of analysis are usually contained in one or more courses in the usual engineering curriculum, e.g., the Cauchy-Schwarz and triangle inequalities. Hopefully the only unfamiliar results are a corollary to the Courant-Fischer theorem and the Weierstrass approximation theorem. The latter is an intuitive result which is easily believed even if not formally proved. More advanced results from Lebesgue integration, measure theory, functional analysis, and harmonic analysis are not used.

Our approach is to relate the properties of Toeplitz matrices to those of their simpler, more structured special case – the circulant or cyclic matrix. These two matrices are shown to be asymptotically equivalent in a certain sense and this is shown to imply that eigenvalues, inverses, products, and determinants behave similarly. This approach provides a simplified and direct path to the basic eigenvalue distribution and related theorems. This method is implicit but not immediately apparent in the more complicated and more general results of Grenander in Chapter 7 of [15]. The basic results for the special case of a banded Toeplitz matrix appeared in [13], a tutorial treatment of the simplest case which was in turn based on the first draft of this work. The results were subsequently generalized using essentially the same simple methods, but they remain less general than those of [15].

As an application several of the results are applied to study certain models of discrete time random processes. Two common linear models are studied and some intuitively satisfying results on covariance matrices and their factors are given.

We sacrifice mathematical elegance and generality for conceptual simplicity in the hope that this will bring an understanding of the interesting and useful properties of Toeplitz matrices to a wider audience, specifically to those who have lacked either the background or the patience to tackle the mathematical literature on the subject.

References

[1] G. Baxter, "An Asymptotic Result for the Finite Predictor," *Math. Scand.*, vol. 10, pp. 137–144, 1962.

[2] G. Baxter, "A Norm Inequality for a 'Finite-Section' Wiener-Hopf Equation," *Illinois J. Math.*, pp. 97–103, 1962.

[3] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression.* Prentice Hall, Englewood Cliffs, New Jersey, 1971.

[4] A. Böttcher and S. Grudsky, *Toeplitz Matrices, Asymptotic Linear Algebra, and Functional Analysis.* Birkhäuser, 2000.

[5] A. Böttcher and B. Silbermann, *Introduction to Large Truncated Toeplitz Matrices.* Springer, New York, 1999.

[6] W. Cheney, *Introduction to Approximation theory.* McGraw-Hill, 1966.

[7] T. A. Cover and J. A. Thomas, *Elements of Information Theory.* Wiley, New York, 1991.

[8] P. J. Davis, *Circulant Matrices.* Wiley-Interscience, NY, 1979.

[9] D. Fasino and P. Tilli, "Spectral clustering properties of block multilevel Hankel matrices," *Linear Algebra and its Applications*, vol. 306, pp. 155–163, 2000.

[10] R. M. Gray and L. D. Davisson, *An Introduction to Statistical Signal Processing.* Cambridge University Press, London, 2005.

[11] R. M. Gray, "On Unbounded Toeplitz Matrices and Nonstationary Time Series with an Application to Information Theory," *Information and Control*, vol. 24, pp. 181–196, 1974.

[12] R. M. Gray, "Information Rates of Autoregressive Processes," *IEEE Trans. on Info. Theory*, vol. IT-16, no. 4, pp. 412–421, July 1970.

[13] R. M. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Transactions on Information Theory*, vol. 18, pp. 725–730, November 1972.

[14] U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*. Wiley and Sons, NY, 1966. Chapter 1.

[15] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. University of Calif. Press, Berkeley and Los Angeles, 1958.

[16] A. J. Hoffman and H. W. Wielandt, "The variation of the spectrum of a normal matrix," *Duke Math. J.*, vol. 20, pp. 37–39, 1953.

[17] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, New Jersey, 2000.

[18] P. Lancaster, *Theory of Matrices*. Academic Press, NY, 1969.

[19] J. Pearl, "On Coding and Filtering Stationary Signals by Discrete Fourier Transform," *IEEE Trans. on Info. Theory*, vol. IT-19, pp. 229–232, 1973.

[20] C. L. Rino, "The Inversion of Covariance Matrices by Finite Fourier Transforms," *IEEE Trans. on Info. Theory*, vol. IT-16, no. 2, pp. 230–232, March 1970.

[21] J. Rissanen and L. Barbosa, "Properties of Infinite Covariance Matrices and Stability of Optimum Predictors," *Information Sciences*, vol. 1, pp. 221–236, 1969.

[22] W. Rudin, *Principles of Mathematical Analysis*. McGraw-Hill, NY, 1964.

[23] W. F. Trench, "Asymptotic distribution of the even and odd spectra of real symmetric Toeplitz matrices," *Linear Algebra Appl.*, vol. 302–303, pp. 155–162, 1999.

[24] W. F. Trench, "Absolute equal distribution of families of finite sets," *Lin. Alg. Appl.*, vol. 367, pp. 131–146, 2003.

[25] W. F. Trench, "Absolute equal distribution of the spectra of Hermitian matrices," *Lin. Alg. Appl.*, vol. 366, pp. 417–431, 2003.

[26] W. F. Trench, "A note on asymptotic zero distribution of orthogonal polynomials," *Lin. Alg. Appl.*, vol. 375, pp. 275–281, 2003.

[27] W. F. Trench, "Simplification and strengthening of Weyl's definition of asymptotic equal distribution of two families of finite sets," *Cubo A Mathematical Journal*, vol. 06, no. 3, pp. 47–54, 2004.

[28] W. F. Trench, "Absolute equal distribution of the eigenvalues of discrete Sturm–Liouville problems," *J. Math. Anal. Appl.*, in press.

[29] B. S. Tsybakov, "Transmission capacity of memoryless Gaussian vector channels (in Russian)," *Probl. Peredach. Inform.*, vol. 1, pp. 26–40, 1965.

[30] B. S. Tsybakov, "On the transmission capacity of a discrete-time Gaussian channel with filter (in Russian)," *Probl. Peredach. Inform.*, vol. 6, pp. 78–82, 1970.

[31] E. E. Tyrtyshnikov, "Influence of matrix operations on the distribution of eigenvalues and singular values of Toeplitz matrices," *Linear Algebra and its Applications*, vol. 207, pp. 225–249, 1994.

[32] E. E. Tyrtyshnikov, "A unifying approach to some old and new theorems on distribution and clustering," *Linear Algebra and its Applications*, vol. 232, pp. 1–43, 1996.

[33] H. Widom, "Toeplitz Matrices," in *Studies in Real and Complex Analysis*, (J. I.I. Hirschmann, ed.), MAA Studies in Mathematics, Prentice-Hall, Englewood Cliffs, NJ, 1965.

[34] J. H. Wilkinson, "Elementary proof of the Wielandt-Hoffman theorem and of its generalization," Tech. Rep., Stanford University, Department of Computer Science, January 1970. Report Number CS-TR-70-150.